Annals of the
International Society of
Dynamic Games

Michèle Breton
Krzysztof Szajowski
Editors

Advances in Dynamic Games

Theory, Applications, and Numerical Methods for Differential and Stochastic Games

Dedicated to the
Memory of Arik A. Melikyan

Birkhäuser

Birkhäuser

Annals of the International Society of Dynamic Games
Volume 11

Series Editor
Tamer Başar

Michèle Breton
Krzysztof Szajowski
Editors

Advances in Dynamic Games

Theory, Applications, and Numerical Methods for Differential and Stochastic Games

Dedicated to the Memory of Arik A. Melikyan

Birkhäuser

Editors

Michèle Breton
Department of Management Science
HEC Montréal
3000 Chemin de la Côte Sainte-Catherine
Montréal, Canada H3T 2A7
michele.breton@hec.ca

Krzysztof Szajowski
Institute of Mathematics
and Computer Science
Wrocław University of Technology
Wybrzeże Wyspiańskiego 27
PL-50-370 Wrocław
Poland
krzysztof.szajowski@pwr.wroc.pl

ISBN 978-0-8176-8088-6 e-ISBN 978-0-8176-8089-3
DOI 10.1007/978-0-8176-8089-3
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2010938475

Mathematics Subject Classification (2010): 91A15, 91A22, 91A23, 91A24, 91A25, 91A80

Printed on acid-free paper

www.birkhauser-science.com

This volume of the Annals of Dynamic Games
is dedicated to the memory of
Arik Artavazdovich Melikyan

Foreword

While these Annals were in preparation, on April 6th, 2009 in Moscow, Arik Artavazdovich Melikyan left us.

An outstanding mathematician, member of the International Society of Dynamic Games and best friend, he has very unexpectedly left us in a deep mourning.

His basic contribution to the theory of pursuit and evasion games lies in new approaches for finding the solutions of Bellman–Isaacs equations and constructing new singular surfaces.

His scientific achievements were highly regarded in Russia, which he always considered as his motherland, and where he was elected a corresponding member of the Russian Academy of Sciences. In his second mother country, Armenia, he was also a foreign member of the National Academy of Sciences.

He was a good friend, an optimistic and cheerful person, and as such he will remain in our thoughts.

St. Petersburg, April 2009 *Leon Petrosyan*

Contents

Foreword vii

Preface xiii

Acknowledgements xv

A Tribute to Arik Artavazdovich Melikyan xvii

Testimonies xix

Arik A. Melikyan (10.05.1944–04.06.2009) xxv

Bibliography of Works by Arik A. Melikyan xxix

The Long History of Arik's Conjecture xxxvii
Pierre Bernhard

Optimal Control Problem for a Dynamical System on an Ellipsoid xxxix
Arik A. Melikyan and Fumiaki Imado

Contributors xliii

Part I Dynamic-Differential Games: Theoretical Developments

On a Structure of the Set of Differential Games Values 3
Yurii Averboukh

Existence and Uniqueness of Disturbed Open-Loop Nash Equilibria for Affine-Quadratic Differential Games 25
Teresa-Paula Azevedo-Perdicoúlis and Gerhard Jank

Fields of Extremals and Sufficient Conditions for a Class of Open-Loop Variational Games 41
Dean A. Carlson and George Leitmann

Riemann–Liouville, Caputo, and Sequential Fractional Derivatives in Differential Games 61
Arkadii Chikrii and Ivan Matychyn

The Mixed Zero-Sum Stochastic Differential Game in the Model with Jumps 83
Saïd Hamadène and Hao Wang

Discontinuous Value Function in Time-Optimal Differential Games 111
Liudmila Kamneva

On Grid Optimal Feedbacks to Control Problems of Prescribed Duration on the Plane 133
Nina N. Subbotina and Timofey B. Tokmantsev

Sub- and Super-optimality Principles and Construction of Almost Optimal Strategies for Differential Games in Hilbert Spaces 149
Andrzej Święch

Part II Pursuit Evasion Games

Extremal Aiming in Problems with Unknown Level of Dynamic Disturbance 167
Sergei A. Ganebny, Sergey S. Kumkov, Valerii S. Patsko, and Sergei G. Pyatko

A Generalization of the Multi-Stage Search Allocation Game 189
Ryusuke Hohzaki

Linear Differential Game with Two Pursuers and One Evader 209
Stéphane Le Ménec

Homicidal Chauffeur Game: History and Modern Studies 227
Valerii S. Patsko and Varvara L. Turova

Collision Avoidance Strategies for a Three-Player Game 253
Sriram Shankaran, Dušan M. Stipanović, and Claire J. Tomlin

Part III Evolutionary Games

Imitative Behavior in a Two-Population Model 275
Elvio Accinelli, Juan Gabriel Brida, and Edgar J. Sanchez Carrera

ESS, Population Games, Replicator Dynamics: Dynamics and Games if not Dynamic Games 291
Pierre Bernhard

A Markov Decision Evolutionary Game for Individual Energy Management 313
Yezekael Hayel, Hamidou Tembine, Eitan Altman, and Rachid El-Azouzi

Mutual Mate Choice with Multiple Criteria 337
David M. Ramsey

Part IV Cooperative Games

The Shapley Value in Cooperative Differential Games with Random Duration 359
Ekaterina Shevkoplyas

Dynamically Consistent Cooperative Solutions in Differential Games with Asynchronous Players' Horizons 375
David W.K. Yeung

Part V Applications and Numerical Approaches

Measure-Valued Solutions to a Harvesting Game with Several Players 399
Alberto Bressan and Wen Shen

Optimal Distributed Uplink Channel Allocation: A Constrained MDP Formulation 425
Vivek S. Borkar and Joy Kuri

Binomial Approximations for Barrier Options of Israeli Style 447
Yan Dolinsky and Yuri Kifer

A Game of International Climate Policy Solved by a Homogeneous Oracle-Based Method for Variational Inequalities 469
Laurent Drouet, Alain Haurie, Jean-Philippe Vial, and Marc Vielle

Optimal Stopping of a Risk Process with Disruption and Interest Rates 489
Elżbieta Z. Ferenstein and Adam Pasternak-Winiarski

Examples in Dynamic Optimal Taxation 509
Mikhail Krastanov and Rossen Rozenov

On a Discounted Inventory Game 521
Heinz-Uwe Küenle

Time-Consistent Emission Reduction in a Dynamic Leader-Follower Game 537
Yulia Pavlova

Fish Wars with Changing Area for a Fishery 553
Anna Rettieva

Preface

Modern game theory has evolved enormously since its inception in the 1920s from the work of Borel (1921, 1924, 1927) and von Neumann (1928). Dynamic game theory branched from the pioneering work on differential games by R. Isaacs, L.S. Pontryagin and his school, and from seminal papers on extensive form games by Kuhn and on stochastic games by Shapley.

Since these early development days, dynamic game theory has had a significant impact in such diverse disciplines as applied mathematics, economics, systems theory, engineering, operations research, biology, ecology, and the environmental sciences. Modern dynamic game theory now relies on wide-ranging mathematical and computational methods, and possible applications are rich and challenging.

This edited volume focuses on various aspects of dynamic game theory, providing state-of-the-art information on recent theoretical developments and new application domains and examples. Most of the selected papers are based on presentations at the 13th International Symposium on Dynamic Games and Applications held in Wrocław, Lower Silesia, Poland at the beginning of Summer 2008. The symposium is held every two years under the auspices of the International Society of Dynamic Games (ISDG).

The papers selected to appear in the Annals cover a variety of topics ranging from theory to applications in biology, ecology, engineering, economics, and finance. The list of contributors consists of both well-known and young researchers from all over the world. Every paper has gone through a stringent reviewing process.

While we were in the middle of the review process, our fellow editor and good friend Arik Artavazdovich Melikyan suddenly passed away. This volume is dedicated to him as a tribute to his contribution in the field of dynamic games.

The volume is divided into five parts. The first part contains eight papers devoted to theoretical developments in differential games and general dynamic games, including new numerical methods. Part II contains five papers on pursuit/evasion games, including an historical perspective on the homicidal chauffeur game, collision avoidance and search games, and guidance problems. Part III is devoted to evolutionary games, with four papers on stable strategies, social interactions, mating, and telecommunication. Part IV contains two papers on cooperative games, addressing the problems of dynamic consistency and imputations when the horizon of the game is different or not known with certainty. Finally, Part V contains nine papers devoted to various applications of dynamic games, covering modeling, solutions, and numerical approaches. Applications range from management of fisheries and environmental agreements to insurance, option pricing and taxation, supply chain management, and channel allocation.

The editors are indebted to many colleagues involved in the editorial process. Special thanks go to Valerii S. Patsko, Andrei R. Akhmetzhanov, and Naira Hovakimyan, who helped us recover the editorial work of Arik after his death, put together his work, and collected testimonies from his friends.

Our warmest thanks go to the large number of referees for the papers submitted to the Annals. Without their important contribution this volume would not have been possible.

HEC Montréal, Canada, WUT, Wrocław, Poland *Michèle Breton*
June 2010 *Krzysztof Szajowski*

Acknowledgements

The selection process of the contributions to this volume started during the 13th International Symposium on Dynamic Games and Applications in Wrocław. The conference was cosponsored by the following scientific and academic institutions: Committee of Mathematics, Polish Academy of Science (PAS); Institute of Mathematics and Computer Science, Wrocław University of Technology; Institute of Mathematics, PAS; Faculty of Mathematics, Computer Science, and Econometrics, University of Zielona Góra; Faculty of Mathematics, Informatics, and Mechanics, University of Warsaw; and GERAD, HEC, Montréal, Canada.

Our warmest thanks go to the referees of the papers. Without their important contribution this volume would not have been possible.

List of Reviewers

Andrei R. Akhmetzhanov	Rabah Amir	Lukasz Balbus	Eric Baurdoux
Hassan Benchekroun	Moez Bennouri	Pierre Bernhard	Alejandro Caparrós-Gass
Arkadii A. Chrikrii	Robert C. Clark	Erik Ekström	Jacob Engwerda
Mauricio Falcone	Tom Ferguson	Robbert J. Fokkink	Sergey A. Ganebny
Andrey Garnaev	Talat Genc	Beniamin Goldys	Onesimo Hernandez-Lerma
Anna Jaśkiewicz	Uwe Jensen	Ljudmila V. Kamneva	Anatolii F. Kleimenov
Michael Kopel	Vseslav A. Korneev	Peter Kort	Jacek Krawczyk
Sergei S. Kumkov	Stéphane Le Ménec	Ngo van Long	Cloud Makasu
Patrice Marcotte	Robert J. Marks II	Guiomar Martín-Herrán	Toshihiro Matsumura
Vladimir V. Mazalov	Jacek Miękisz	Teruhisa Nakai	Andrea Pascucci
Bozenna Pasik-Duncan	Valerii S. Patsko	Szymon Peszat	Leon A. Petrosjan
Odile Pourtallier	Marc Quincampoix	Evgeny Y. Rubinovich	Wojciech Sarnowski
Oana Silvia Serea	Anne Souquières	Łukasz Stettner	Andrzej Święch
Mitsushi Tamaki	Corinne Touati	Varvara Turova	Piotr Wiecęk
Andrzej Wieczorek	David Yeung	Jiongmin Yong	Georges Zaccour

A Tribute to Arik Artavazdovich Melikyan

1. Testimonies
2. Short Biography
3. Publications
4. *Pierre Bernhard*, The Long History of Arik's Conjecture
5. *Arik A. Melikyan and Fumiaki Imado*, Optimal Control Problem for a Dynamical System on an Ellipsoid

Testimonies

Arik Artavazdovich Melikyan died suddenly on April 6, 2009, in the Institute for Problems in Mechanics of the Russian Academy of Science (IPMech), where he worked for almost 40 years.

Having graduated with honors from Moscow Institute of Physics and Technology, Arik began to work at IPMech. Here he received his Ph.D. and Dr. of Sciences degrees, was awarded the State Prize of Russia for his research, was elected a Corresponding Member of the Russian Academy of Sciences, and became Head of the Laboratory of Control of Mechanical Systems.

I became acquainted with Arik when he was a bright student vividly interested in the theory of differential games, which was, at that time, a new and fascinating field of research. I was a supervisor of his Ph.D. thesis devoted to the new topic of differential games with incomplete information. Later, I worked with Arik for a long time, sometimes more tightly, sometimes on different subjects, but we always kept in close professional and human contact.

Arik was, of course, a brilliant scientist who obtained original and important results in differential games, the theory of the Hamilton–Jacobi equation, and optimal control.

He was also a remarkable personality: very kind and humane, who was loved and respected by everybody who knew him. We will never forget Arik.

Professor Felix Chernousko
Academician of the Russian Academy of Sciences
Director of the Institute for Problems in Mechanics
Russian Academy of Sciences, Moscow

◇◇◇◇◇◇◇

Arik Melikyan was my Ph.D. advisor from 1988 to 1992 in Moscow. During those years I used to interact quite often with his family both in Moscow and in Yerevan. We had a lot of cozy tea evenings with his wife Karina, his children Zara and Artem in their Moscow apartment; and also warm interactions with his mother Shushanik, his brother Gagik, and his sister Gayane in Yerevan. Ever since then we maintained a close friendship with Arik and his family members. So, my memories of Arik are rich, diverse, and colorful.

As his former Ph.D. student, in this small write-up I would like to place the emphasis on the wisdom that Arik shared with me and most probably with his other graduate students as well. He had a sense of humor and a sharp wit very unique to him. I was always left with an impression that there was no challenge for him — he knew the best way out of every situation. He could easily build relationships between math and life, between science and music, between art and politics. He could lend comfort and calmness, be on your side and the opposite side simultaneously, find the right path out, and the right words to say. Arik was not just a great advisor; that might have been a lot easier to achieve. With his serious attitude to the quality of technical work and the significance of foundational work and open problems in science, he always made hard fundamental problems look much simpler. In fact, only after many years of completing my thesis with him, I realized the significance of my Ph.D. contribution to the canvas of fundamental mathematics. According to Arik, we were having fun with "pretty mathematics." Prettiness was the major driver for him in every pursuit. Arik always seemed to be happy and content with life. He never complained. He knew how to appreciate every minute of life, how to cherish friendship, and how to be a friend. Arik's viewpoint of the world and his philosophy of life are still an integral part of my thinking in my everyday routine.

God bless his soul.

Naira Hovakimyan
Professor and Schaller faculty scholar
Mechanical Science and Engineering
University of Illinois at Urbana-Champaign

◇◇◇◇◇◇◇

Using the philosophy of antagonistic differential games, A.A. Melikyan developed his method of singular characteristics for nonlinear first-order partial differential equations. Within the framework of differential games, the singular characteristics are regarded as "optimal" motions by virtue of some control actions of players.

Having spoken to experts in mathematical physics, Arik always looked for the opportunity of a "non-game" interpretation of the singular characteristics. For example, equivocal singular surfaces are specific for differential games. But are such surfaces known and met in problems describing physical processes related to liquid or gas flows? If yes, then what is the physical interpretation of such singular surfaces?

I think that Arik's book on singular characteristics essentially extends the classic theory and will find many interested readers in the future.

Valerii S. Patsko
Institute of Mathematics and Mechanics
Ural Branch of Russian Academy of Sciences
Ekaterinburg, Russia

◇◇◇◇◇◇◇

In autumn of 1971 Felix Chernousko from the Soviet Academy of Sciences asked me to be an examiner for the Ph.D. candidacy thesis of Arik Melikyan called "On control problems and differential pursuit games with incomplete information." The basic idea was to define a discrete set of observations necessary to solve the control or pursuit problem of a particular kind. This was a new problem setting, which seemed to also have practical importance. And I agreed. This was the start of our friendly and fruitful relationship. I often visited the seminar of Chernousko, where Arik was also very active, and Arik visited St. Petersburg many times to present his new results in differential games in our university seminar. Since at that time I was interested in simple pursuit games with geometrical constraints, we had many topics for discussion. Arik was open minded, friendly, optimistic, and had a very kind personality. He was also religious, which was uncommon for Soviet times. He regularly visited the Armenian Apostolic Church of Surb Harutyun in Moscow and, when in St. Petersburg, the Church of Surb Katarine on Nevsky Prospekt. He truly loved his native country the Soviet Union (later Russia), and of course his motherland, Armenia. I remember an episode. I think it was in 1991. I went to Firenze for a Game Theory Conference and on the way there stopped in Venice for six hours to visit the St. Lazars Island, where in 1717 the Armenian monastery was founded by Abate Mchitar. I was very much impressed by this visit and, meeting Arik during the conference, told him about it. At that time the financial situation in the Soviet Union was very far from good and especially scientists were in the worst condition. But in spite of that, Arik changed his return plans in order to visit the monastery on his way back to Moscow. In winter of 2008 Arik and I were elected foreign members of the National Academy of Sciences of the Republic of Armenia. I had hoped to meet him during the General Assembly of the Academy in April 2009, but things went in a different way. My family and I will always remember Arik as an honest scientist and good friend.

Leon Petrosjan

◇◇◇◇◇◇◇

Arik was my colleague and friend for more than 30 years. For many years we shared one room in the Institute for Problems in Mechanics of the Russian Academy of Sciences. I had the pleasure to work with him, to enjoy his exclusive benevolence, sincere interest in scientific ideas and achievements of his colleagues, and his rare ability to be sincerely glad about success and depressed about the failure of other persons. He had a healthy sense of humor, loved and knew literature and the arts. He loved life in all its diversity.

Nikolai Bolotnik

◇◇◇◇◇◇◇

Warm, gentle, caring, kind, a great mathematician, and always the same are the words that I would use to describe Arik. I certainly lost somebody who had become a dear friend.

Dušan M. Stipanović

◇◇◇◇◇◇◇

One of the last times I met Arik, I told him I was worried that all of Armenia's population will soon leave the country, given the well-known difficulties. He smiled and told me that this would probably happen if the world were linear, but possibly not if it were nonlinear. I shall miss my meetings with him.

Berç Rustem

◇◇◇◇◇◇◇

It was really sad to hear the news of Arik's sudden death at his laboratory. He was almost the same age as I am, and I am afraid he overworked himself. Although we had communicated with each other since 1990, we became close friends when we were invited to Kiev Technical University by Prof. Arkadii Chikrii in June 2002. At the conference I was very interested in his talk "Differential Games on Curved Surfaces." I invited him to Shinshu University in April 2006. During his five weeks' stay, I brought him to many universities and sightseeing spots. I introduced him to many professors, who were all impressed by his deep knowledge and activities in many fields. He was particularly pleased with the varieties of Japanese cooking and the blooming cherry blossoms. We soaked in a hot spring together, drank, and sang Russian songs in a bar. He often said he would come to Japan again with his wife. Alas, it has become an impossible dream. Now I can only pray for his peaceful rest.

Professor Fumiaki Imado
Shinshu University, Japan

◇◇◇◇◇◇◇

Arik Melikyan became, to me, not simply a supervisor of my studies, but a great teacher. I was always surprised at his wisdom and irrepressible energy.

Georgy Naumov
Former student of Arik Melikyan

◇◇◇◇◇◇◇

Arik had a very valuable skill: the ability to influence people. We first met while organizing the 13th ISDG symposium in Wrocław. He brought a different perspective to our discussions, and we learned quickly to respect each other's views. He was always quiet, and in appearance was perhaps unremarkable, yet he had an incredible strength of character. He gave us all so much more than he took, sharing his experience, skills, and knowledge.

I last saw Arik in Delft after the retirement party for G.J. Olsder. He was walking down the old streets, commenting on the habits of local students, planning future meetings, and clearly enjoying life. No one expected that we would never again talk to him, see him, or hear his voice over Skype... We have lost an exceptional person who will be remembered with great warmth and respect by me and by all his colleagues.

Krzysztof Szajowski

◇◇◇◇◇◇◇

I became a student of Arik Melikyan in 2004, during the last year of my undergraduate studies. I knew Arik for five years, and I thank my fate that it gave me a chance to work under the supervision of this great man. Arik was a kind person, always very polite; he was a gentleman with his own style of life. For me, work with Arik was like a game, very joyful and interesting. Unfortunately, God decided to take him away from us very early.

Andrei Akhmetzhanov
Former student of Arik Melikyan

◇◇◇◇◇◇◇

Arik Melikyan was not just a great scientist, but also a good senior friend of mine. His contributions and ideas influenced and broadened my research interests, and also educated a wide audience of young generation of Armenian scientists. During the last few years, Arik Melikyan visited his home country, Armenia, more often than in prior years, and we had the opportunity of numerous meetings, discussions, and exchanges of ideas. Arik Melikyan was well respected in Armenia: recently he was elected as a foreign member of the National Academy of Sciences of Armenia. We were looking forward to hosting him and honoring him here during the week of April 18–25 as a part of the celebrations in honor of foreign members of the Academy, but an unexpected event took away this opportunity. We will always cherish our memories of Arik as a close friend, a sincere patriot, and a great scientist.

Vahan Avetisyan
Leading research scientist
Institute of Mechanics
Armenian Academy of Sciences

◇◇◇◇◇◇◇

Arik was my great friend for many years. We met at FizTech (Moscow Institute of Physics and Technology) in 1966, and from the very first meeting I remember a strong impression of a unique mixture of kindness, humor, and fairness that made me deeply interested in becoming Arik's friend. We met many times during my six years of studies in Dolgoprudny. Yet later in 1972, when I started my Ph.D. studies at the Institute of Problems in Mechanics, we became colleagues in a rather small group conducting research on optimal control of uncertain dynamical systems and became close friends. No wonder, Arik was the witness at my wedding in 1976. Since then we could be separated by a long distance, but our friendship was as strong as it was in our young years. In twenty years of living and working in Atlanta, I made no visit to Moscow without meeting Arik and his wonderful family. Needless to say, our daughters became friends as well. Many times Arik came to Atlanta as a Visiting Professor at the School of Mathematics at Georgia Tech. Our common interest in information and control led to the Fogarty Award, a US grant to a joint USA-Russian team for developing algorithms for analysis of DNA sequences carrying genetic information. Through the years, Arik was constantly present in my world; I felt his support, saw his smile, heard his wise and warm words. I cannot believe that Arik is no longer with us. I will always remember him as my dear friend.

Mark Borodovsky, Ph.D.
Director, Center for Bioinformatics and Computational Genomics at Georgia Tech
Regents' Professor, Wallace H. Coulter Department of Biomedical Engineering
Georgia Tech and Emory University
Georgia Tech Division of Computational Science and Engineering
Atlanta, Georgia, USA

◇◇◇◇◇◇◇

I met Arik Melikyan for the first time at the Institute of Problems in Mechanics of the Russian Academy of Sciences in 1969. My first impression was of an extremely charming and handsome young man, who had biblical, noble, unbelievable beauty and charm, reminding me of old Assyrian bas-reliefs. My first impression was true: Arik Melikyan was indeed noble and handsome all his life. He was a great friend. With Arik you could share the most sacred secrets. He was always very patient, diplomatic, balanced; he never made hasty conclusions. He had a purity that was inspiring to others.

The first problem that Arik tackled was the problem of two boats chasing each other around an island. It was a well-known problem from Isaacs' book. In some sense, this defined the subject of his research interests throughout his life. Arik enjoyed dealing with real-life problems that were difficult from a mathematical point of view and without trivial solutions. Committing himself to that path, Arik made several fundamental discoveries about the structure of the solutions to the Hamilton–Jacobi–Bellman–Isaacs equation, which constitutes the basis for the modern theory of differential games. His successful career is the result of sincere dedication to science and tremendous diligence.

The death of a person presents a property of reverse perspective, in the sense that all the private and minor points tend to disappear, while the main features surface and have the potential of lasting forever. The main feature for Arik was his unparalleled nobleness and dignity.

Alexander Bratus

◇◇◇◇◇◇◇

Arik A. Melikyan (10.05.1944–04.06.2009)

Arik Melikyan, a prominent scientist in the area of differential games, optimal control, and related areas of the theory of differential equations, a Corresponding Member of the Russian Academy of Sciences, a recipient of the Russian State Prize, a foreign member of Armenian National Academy of Sciences, died on April 6, 2009. He was only 64 years old.

Arik Melikyan is well known for his fundamental and long lasting contributions to the theory of optimal control and differential games. He generalized the classical concept of regular differential games and extended the mathematical tools to a substantially broader class of game problems, calling these problems regular in a broader sense. Arik is well known for generalization of the method of regular characteristics for first-order PDEs to handle non-smooth value functions and/or initial conditions. In differential games, quite often the value function of the underlying Hamilton–Jacobi–Bellman–Isaacs (HJBI) equation is not smooth. In such cases, the phase portrait for the complete state space cannot be covered by optimal paths using the regular method of characteristics. Arik Melikyan was the first to obtain the algorithms for construction of optimal phase portraits in differential games with non-smooth value functions by exploring the differential geometric structure of the phase space associated with non-uniqueness of minimal geodesic lines connecting the players. This led to analytical description of a large class of singular surfaces and provided the opportunity for extending the classical method of characteristics to games with non-smooth value functions. The Method of Singular Characteristics by Arik Melikyan is widely used today not just in the analysis of pursuit-evasion games in the presence of obstacles, but also in the investigation of propagation of tangential discontinuities in some physical media and in processing visual information for the purpose of feedback. Arik also significantly contributed to the theory of control of uncertain systems and developed game-theoretic type algorithms for optimal search for objects moving in a bounded area using a computer vision system with a restricted aperture. Arik Melikyan is the author of two monographs and more than 150 papers in Russian and international scientific journals.

Arik Melikyan was born on October 5, 1944 in Yerevan, Armenia. His parents – Artavazd and Shushanik – were high school biology teachers. His brother Gagik is an engineer, and his sister Gayane is a chemist. Arik graduated from high school in 1961, Magna cum Laude (Golden Medal in Russia), and was accepted to the Department of Cybernetics in Yerevan Polytechnic Institute. Being an outstanding student, he sought the highest possible quality of education available at the time in the former USSR. In 1963 he moved to Moscow to continue his education in the Department of Aeromechanics and Applied Mathematics at the Moscow Institute of Physics and Technology (MIPT). Arik graduated from MIPT, magna cum laude, in 1969 as a physicist-engineer, specializing in flight dynamics and control. From his early childhood he proved to be an extremely bright person. Arik was not just a gifted mathematician; as a young student he was one of the top athlete-runners in Armenia. He studied music and played several instruments (piano, accordion). He enjoyed learning languages and spoke several of them. Along with his major interest in mathematics he had strong interests in journalism, literature, and the arts in general. He also took classes in the Department of Journalism at the Lomonosov Moscow State University, publishing interviews with prominent figures of the time. He also wrote short stories for a major Russian literary newspaper. He was actively involved in student life – organizing concerts of well-known singers and actors for the students and professors of MIPT, and writing for the student newspaper. He was a very generous and kind person: as a student and later as a mature professional, he was always taking time to help friends, colleagues, and students. His friends and colleagues liked him for his easy and joyful character, his respect for others, his great intelligence and sense of humor, and his love of people and his profession.

Upon his graduation from the Moscow Institute of Physics and Technology, Arik became a Ph.D. student at this institute and continued his research in mathematical theory of control under the scientific supervision of Felix Chernousko. In 1972 he defended his Ph.D. thesis on differential games with incomplete information. After that he worked at the Institute for Problems in Mechanics of the USSR Academy of Sciences (currently the A.Yu. Ishlinskii Institute for Problems in Mechanics of the Russian Academy of Sciences). His entire scientific career was associated with this institute; he grew from a junior research fellow to become the head of a laboratory. In 1986, he defended the thesis "Game Problems of Dynamics" and received his Doctor-of-Science (habilitation) degree. In 1998, he was awarded the State Prize of the Russian Federation in the Field of Science and Technology. In 2003, he was elected a Corresponding Member of the Russian Academy of Sciences.

Arik Melikyan performed great organizational work for the community: he was a member of the Russian National Committee on Theoretical and Applied Mechanics and of the Executive Committee of the International Society of Dynamic Games, a member of the editorial boards of the *Journal of Computer and Systems Science International* and *International Game Theory Review.* He was a member of various Program Committees and Organizational Committees of numerous Russian and international conferences on the mathematical theory of control and differential games. As a member of the Editorial Board of the *Journal of Computer and Systems Science International,* Arik Melikyan was noted for his benevolent and attentive attitude to authors and his high standards with regard to the quality and language of papers accepted for publication.

In addition to being an outstanding scientist, Arik was a great teacher and mentor. He always considered teaching on par with his research and enjoyed his professorship at MIPT. Arik was a great friend of his graduate students. He used to work with them long hours, and he maintained his friendship with his students throughout his life. He was a professor of the Moscow Institute of Physics and Technology, where he delivered original courses on the calculus of variations, optimal control, and differential games. Arik was also a member of the Educational and Methodological Council of the Department of Aerophysics and Space Exploration. His lectures were always a great success among students.

Arik was not only an exceptional scientist and a gifted teacher but a wonderful person. He was exclusively benevolent and responsive, helped others to find the way out of complicated situations. He was optimistic with a good sense of humor, able to solve difficult life problems. Arik Melikyan will forever live in the memory of all people who knew him.

He is survived by his loving wife Karina, daughter Zara, and son Artem.

Bibliography of Works by Arik A. Melikyan

Books

1. Chernousko, F.L., Melikyan, A.A.: Igrovye zadachi upravleniya i poiska (Game-Theoretical Control and Search Problems). Nauka, Moscow (1978) (in Russian)
2. Melikyan, A.A.: Generalized Characteristics of First Order PDEs. Applications in Optimal Control and Differential Games. Birkhäuser, Boston (1998)

Theses

3. Melikyan, A.A.: Some Problems of Optimal Control and Differential Games with Incomplete Information. PhD Thesis, Acad. Sci. USSR, Institute for Problems in Mechanics, Moscow (1972) (in Russian)
4. Melikyan, A.A.: Some Problems of Optimal Control and Differential Games with Incomplete Information. Author's summary of PhD Thesis, Moscow Institute of Physics and Technology, Moscow (1972) (in Russian)
5. Melikyan, A.A.: Construction of Optimal Motions in Game Problems of Dynamics. Doctoral Thesis, Acad. Sci. USSR, Institute for Problems in Mechanics, Moscow (1985) (in Russian)
6. Melikyan, A.A.: Construction of Optimal Motions in Game Problems of Dynamics. Author's summary of Doctoral Thesis, Acad. Sci. USSR, Ural Sci. Center, Institute of Mathematics and Mechanics, Sverdlovsk (1985) (in Russian)

Papers in Peer-Reviewed Journals

7. Melikian, A.A., Chernousko, F.L.: Certain minimax control problems with incomplete information. J. Appl. Math. Mech. **35(6)**, 907–916 (1971). Translated from Prikl. Mat. Mekh. **35(6)**, 952–961 (1971) (in Russian)
8. Melikjan, A.A.: Differential games with information gaps. Izv. Akad. Nauk Armjan. SSR, Ser. Matematika **7(4)**, 300–308 (1972) (in Russian)
9. Melikyan, A.A., Chernousko, F.L.: On differential games with variable conditions of information. Dokl. Akad. Nauk SSSR **203(1)**, 46–49 (1972) (in Russian)
10. Melikian, A.A.: On minimal observations in a game of encounter. J. Appl. Math. Mech. **37(3)**, 407–414 (1973). Translated from Prikl. Mat. Mekh. **37(3)**, 426–433 (1973) (in Russian)

The bibliography was collected by Valerii S. Patsko and Alexander E. Utemov.

11. Melikjan, A.A.: The minimal number of moments of observations in a model approach game. Izv. Akad. Nauk Armjan. SSR, Ser. Matematika **9(3)**, 242–247 (1974) (in Russian)
12. Melikian, A.A.: On optimum selection of noise intervals in differential games of encounter. J. Appl. Math. Mech. **39(2)**, 195–203 (1975). Translated from Prikl. Mat. Mekh. **39(2)**, 207–215 (1975) (in Russian)
13. Korneev, V.A., Melikian, A.A.: Differential game with impulse control of one player. J. Appl. Math. Mech. **41(2)**, 345–349 (1977). Translated from Prikl. Mat. Mekh. **41(2)**, 354–357 (1977) (in Russian)
14. Melikjan, A.A.: Game value in a linear differential approach game. Dokl. Akad. Nauk SSSR **237(3)**, 521–524 (1977) (in Russian)
15. Korneev, V.A., Melikian, A.A.: Impulse-differential encounter game. J. Appl. Math. Mech. **42(2)**, 220–228 (1978). Translated from Prikl. Mat. Mekh. **42(2)**, 210–218 (1978) (in Russian)
16. Chernousko, F.L., Melikjan, A.A.: Minimax problems of control, observation and search. Teoret. Prilozhna Mekh. **9(3)**, 16–24 (1978) (in Russian)
17. Melikyan, A.A., Chernousko, F.L.: The choice of instants of observation in a linear game of approach. Izv. Akad. Nauk SSSR, Ser. Tekhn. Kibernet. **1**, 11–16 (1979) (in Russian)
18. Melikian, A.A.: Game problem of the design of a multi-impulse motion correction. J. Appl. Math. Mech. **44(5)**, 555–560 (1980). Translated from Prikl. Mat. Mekh. **44(5)**, 783–791 (1980) (in Russian)
19. Anchev, A., Melikyan, A.A.: Optimal reorientation of a satellite on a circular orbit. Izv. Akad. Nauk SSSR, Mekh. Tverd. Tela **6**, 37–42 (1980) (in Russian)
20. Melikyan, A.A.: Optimal interaction of two pursuers in a game problem. Izv. Akad. Nauk SSSR, Ser. Tekhn. Kibernet. **2**, 49–56 (1981) (in Russian)
21. Melikyan, A.A.: The Cauchy problem with an unknown boundary for a first-order partial differential equation. Soviet Math. Dokl. **24(2)**, 268–273 (1981). Translated from Dokl. Akad. Nauk SSSR **260(2)**, 288–292 (1981) (in Russian)
22. Melikyan, A.A.: Necessary conditions for optimality on a surface of discontinuity of a certain type in a differential game. Izv. Akad. Nauk SSSR, Ser. Tekhn. Kibernet. **4**, 10–18 (1981) (in Russian)
23. Korneyev, V.A., Melikyan, A.A.: Optimal synthesis in a game problem of pursuit with a small parameter. Izv. Akad. Nauk SSSR, Ser. Tekhn. Kibernet. **3**, 52–58 (1982) (in Russian)
24. Vishnevetskii, L.S., Melikian, A.A.: Optimal pursuit on a plane in the presence of an obstacle. J. Appl. Math. Mech. **46(4)**, 485–490 (1982). Translated from Prikl. Mat. Mekh. **46(4)**, 613–620 (1982) (in Russian)
25. Melikyan, A.A.: Construction of weak discontinuities in optimal control problems and differential games. Engrg. Cybernetics **22(1)**, 129–133 (1984). Translated from Izv. Akad. Nauk SSSR, Ser. Tekhn. Kibernet. **1**, 45–50 (1984) (in Russian)
26. Melikyan, A.A., Ovseevich, A.I.: Hamiltonian systems with a specified invariant manifold and some of their applications. J. Appl. Math. Mech. **48(2)**, 140–145 (1984). Translated from Prikl. Mat. Mekh. **48(2)**, 205–213 (1984) (in Russian)
27. Melikyan, A.A., Sokolov, B.N.: The construction of universal surfaces in extremal problems of dynamics. J. Appl. Math. Mech. **49(2)**, 144–149 (1985). Translated from Prikl. Mat. Mekh. **49(2)**, 191–198 (1985) (in Russian)
28. Korneev, V.A., Melikyan, A.A., Titovskiy, I.N.: Stabilization of aircraft glide path in wind disturbances in the minimax formulation. Izv. Akad. Nauk SSSR, Ser. Tekhn. Kibernet. **3**, 132–139 (1985) (in Russian)
29. Melikyan, A.A., Prosyanyuk, V.V., Sokolov, B.N., Yakovlev, V.Ya.: Optimization of the arrangement of forming protrusions on the surface of a turning roller. Mashinovedenie **4**, 64–70 (1986) (in Russian)
30. Grudev, A.I., Melikyan, A.A.: Modelling of dynamics of a processing section. Izv. Akad. Nauk SSSR, Ser. Tekhn. Kibernet. **3**, 32–40 (1987) (in Russian)
31. Grudev, A.I., Melikyan, A.A.: Simulation modelling of flexible automated systems. Izv. Akad. Nauk SSSR, Ser. Tekhn. Kibernet. **4**, 119–130 (1988) (in Russian)

32. Melikyan, A.A.: Minimax control problem in the case of incomplete information about target point position. Soviet J. Comput. Systems Sci. **27(5)**, 121–127 (1989). Translated from Izv. Akad. Nauk SSSR, Ser. Tekhn. Kibernet. **2**, 111–118 (1989) (in Russian)
33. Melikyan, A.A.: The problem of time-optimal control with search for a target point. J. Appl. Math. Mech. **54(1)**, 1–7 (1990). Translated from Prikl. Mat. Mekh. **54(1)**, 3–11 (1990) (in Russian)
34. Melikyan, A.A.: A positional method for the analysis of singular motions in extremal problems in dynamics. Soviet Math. Dokl. **41(2)**, 205–209 (1990). Translated from Dokl. Akad. Nauk SSSR **311(1)**, 22–26 (1990) (in Russian)
35. Melikyan, A.A., Ovakimyan, N.V.: Singular trajectories in a problem of simple pursuit on a manifold. J. Appl. Math. Mech. **55(1)**, 42–48 (1991). Translated from Prikl. Mat. Mekh. **55(1)**, 54–62 (1991) (in Russian)
36. Melikyan, A.A., Ovakimyan, N.V.: A simple pursuit-and-evasion game on a two-dimensional cone. J. Appl. Math. Mech. **55(5)**, 607–618 (1991). Translated from Prikl. Mat. Mekh. **55(5)**, 741–751 (1991) (in Russian)
37. Melikyan, A.A.: Optimal guaranteeing control of a dynamical system with search for the target point. Soviet Math. Dokl. **43(1)**, 193–197 (1991). Translated from Dokl. Akad. Nauk SSSR **316(4)**, 815–819 (1991) (in Russian)
38. Bocharov, V.Yu., Melikyan, A.A.: Guaranteed speed problem with target point search in a polygonal uncertainty domain. Soviet J. Comput. Systems Sci. **30(1)**, 61–69 (1992). Translated from Izv. Akad. Nauk SSSR, Ser. Tekhn. Kibernet. **1**, 82–92 (1992) (in Russian)
39. Melikyan, A.A., Ovakimyan, N.V.: A differential game of simple approach on manifolds. J. Appl. Math. Mech. **57(1)**, 47–57 (1993). Translated from Prikl. Mat. Mekh. **57(1)**, 41–51 (1993) (in Russian)
40. Melikyan, A.A.: Necessary conditions for optimality of a singular surface in the form of synthesis for a nonlinear controlled system. J. Comput. Systems Sci. Internat. **32(2)**, 134–139 (1994). Translated from Izv. Akad. Nauk, Ser. Tekhn. Kibernet. **1**, 158–163 (1993) (in Russian)
41. Melikyan, A.A.: Necessary optimality conditions for a singular surface in the form of synthesis. J. Optim. Theory Appl. **82(2)**, 203–217 (1994)
42. Melikyan, A.A.: Games of simple pursuit and approach on the manifolds. Dynam. Control **4(4)**, 395–405 (1994)
43. Golovko, A.S., Melikyan, A.A.: The problem of the optimal speed of response when searching for the target point in a circular region of uncertainty. J. Comput. Systems Sci. Internat. **33(5)**, 42–48 (1995). Translated from Izv. Akad. Nauk, Ser. Tekhn. Kibernet. **4**, 40–47 (1994) (in Russian)
44. Korneev, V.A., Melikyan, A.A.: Construction of generalized solution to two-dimensional Hamilton–Jacobi equation by the method of characteristics. Izv. Akad. Nauk, Teor. Sist. Upr. **6**, 168–177 (1995) (in Russian)
45. Melikyan, A.A.: Singular characteristics of first-order partial differential equations. Russian Acad. Sci. Dokl. Math. **54(3)**, 831–834 (1996). Translated from Dokl. Akad. Nauk **351(1)**, 24–28 (1996) (in Russian)
46. Melikyan, A., Pourtallier, O.: Games with several pursuers and one evader with discrete observations. Nova J. Math. Game Theory Algebra **6(1)**, 81–95 (1996)
47. Klepfish, B.R., Melikyan, A.A.: Optimization of the shape of a disk cam in a planar mechanism. Mechanics of Solids **31(5)**, 9–13 (1996). Translated from Izv. Akad. Nauk, Mekh. Tverd. Tela **5**, 13–18 (1996) (in Russian)
48. Melikyan, A.A., Ovakimyan, N.V., Harutiunian, L.L.: Games of simple pursuit and approach on a two-dimensional cone. J. Optim. Theory Appl. **98(3)**, 515–543 (1998)
49. Shmatkov, A.M., Melikyan, A.A., Chernousko, F.L., Borodovsky, M.: Finding prokariotic genes by the "frame-by-frame" algorithm: targeting gene starts and overlapping genes. Bioinformatics **15(11)**, 874–886 (1999)
50. Borodovsky, M.Yu., Melikyan, A.A., Chernousko, F.L., Shmatkov, A.M.: Application of a hidden Markov model and dynamic programming for gene recognition in DNAs of bacteria

and viruses. J. Comput. Systems Sci. Internat. **39(1)**, 155–161 (2000). Translated from Izv. Akad. Nauk, Teor. Sist. Upr. **1**, 163–169 (2000) (in Russian)

51. Bolotnik, N.N., Melikyan, A.A.: Optimal control of a second order dynamic system with impacts. J. Comput. Systems Sci. Internat. **39(6)**, 63–73 (2000). Translated from Izv. Akad. Nauk, Teor. Sist. Upr. **6**, 47–57 (2000) (in Russian)
52. Hovakimyan, N., Melikyan, A.: Geometry of pursuit-evasion on second order rotation surfaces. Dynam. Control **10(3)**, 297–312 (2000)
53. Melikyan, A.A.: Propagation of weak discontinuities of the solutions to the multiple integral variational problem. In: Proc. Steklov Inst. Math., suppl. 2: Problems of Control in Dynamic Systems, S113–S124 (2000). Translated from Trudy Instituta Matematiki i Mekhaniki UrO RAN **6(2)**, 446–459 (2000) (in Russian)
54. Aldakimov, Yu.V., Melikyan, A.A., Naumov, G.V.: Mode reconstruction in a one-parameter family of optimal control problems. J. Appl. Math. Mech. **65(3)**, 387–394 (2001). Translated from Prikl. Mat. Mekh. **65(3)**, 400–407 (2001) (in Russian)
55. Melikyan, A.A.: Equations of motion on the phase constraint surface of an optimal control problem in the form of singular characteristics. Differential Equations **38(11)**, 1581–1588 (2002). Translated from Differents. Uravneniya **38(11)**, 1490–1497 (2002) (in Russian)
56. Melikyan, A.A.: Structure of the value function in pursuit-evasion games on the surfaces of revolution. Cybernet. Systems Anal. **38(3)**, 444–452 (2002). See also Kibernet. Sistem. Anal. **3**, 155–163 (2002)
57. Korneev, V.A., Melikyan, A.A.: Construction of weak discontinuities of the solution of a variational problem in a domain. Izv. Akad. Nauk, Teor. Sist. Upr. **1**, 47–60 (2003) (in Russian)
58. Melikyan, A.: Some properties of the Bellman–Isaacs equation for the games on surfaces of revolution. Int. Game Theory Rev. **6(1)**, 171–179 (2004)
59. Melikyan, A.A.: Primary strategies of simple pursuit in differential games on two-sided plane figures. J. Appl. Math. Mech. **68(4)**, 545–554 (2004). Translated from Prikl. Mat. Mekh. **68(4)**, 611–622 (2004) (in Russian)
60. Melikyan, A., Bernhard, P.: Geometry of optimal paths around focal singular surfaces in differential games. Appl. Math. Optim. **52(1)**, 23–37 (2005)
61. Melikyan, A.A.: Singular characteristics in the boundary conditions of the Hamilton–Jacobi equation. Doklady Mathematics **72(2)**, 723–727 (2005). Translated from Dokl. Akad. Nauk **404(4)**, 451–455 (2005) (in Russian)
62. Korneev, V.A., Melikyan, A.A.: Numerical investigation of weak waves in the solution to the variational wave equation. J. Comput. Systems Sci. Internat. **44(1)**, 58–69 (2005). Translated from Izv. Akad. Nauk, Teor. Sist. Upr. **1**, 63–74 (2005) (in Russian)
63. Zhelnin, Yu.N., Melikyan, A.A., Utemov, A.Ye., Chernousko, F.L.: Bringing a non-linear manoeuvring object to the optimal position in the shortest time. J. Appl. Math. Mech. **69(2)**, 163–172 (2005). Translated from Prikl. Mat. Mekh. **69(2)**, 179–190 (2005) (in Russian)
64. Melikyan, A.A.: Necessary optimality conditions for different phase portraits in a neighborhood of a singular arc. In: Proc. Steklov Inst. Math., Vol. 255, suppl. 2: Control, Stability, and Inverse Problems of Dynamics, S126–S139 (2006). Translated from Trudy Instituta Matematiki i Mekhaniki UrO RAN **12(2)**, 129–141 (2006) (in Russian)
65. Akhmetzhanov, A.R., Melikyan, A.A.: A linear differential game on a plane with a minimum-type functional. J. Appl. Math. Mech. **71(2)**, 234–246 (2007). Translated from Prikl. Mat. Mekh. **71(2)**, 264–276 (2007) (in Russian)
66. Melikyan, A., Akhmetzhanov, A., Hovakimyan, N.: On initial value and terminal value problems for Hamilton–Jacobi equation. Systems Control Lett. **56(11–12)**, 714–721 (2007)
67. Akhmetzhanov, A.R., Melikyan, A.A., Fradkov, A.L.: The optimality of the velocity-gradient method in the problem of controlling the escape from a potential well. J. Appl. Math. Mech. **71(6)**, 809–818 (2007). Translated from Prikl. Mat. Mekh. **71(6)**, 899–909 (2007) (in Russian)

68. Day, M.V., Melikyan, A.A.: Simple singularities for Hamilton–Jacobi equations with max-concave Hamiltonians and generalized characteristics. J. Optim. Theory Appl. **138(2)**, 155–174 (2008)
69. Melikyan, A., Olsder, G.J., Akhmetzhanov, A.: Parameter analysis of a zero-sum differential game approximating a nonzero-sum game. Int. Game Theory Rev. **10(4)**, 437–459 (2008)
70. Melikyan, A., Olsder, G.J.: Boundary singularities and characteristics of Hamilton–Jacobi equation. Journal of Dynamical and Control Systems **16(1)**, 77–99 (2010)
71. Melikyan, A.A.: Singular boundary characteristics of Hamilton–Jacobi equation. J. Appl. Math. Mech. **74(2)**, 144-153 (2010). Translated from Prikl. Mat. Mekh. **74(2)**, 202–215 (2010) (in Russian)
72. Stipanović, D.M., Melikyan, A., Hovakimyan, N.: Guaranteed strategies for nonlinear multi-player pursuit-evasion games. Int. Game Theory Rev. **12(1)**, 1–17 (2010)

Papers in Peer-Reviewed Collections

73. Chernousko, F.L., Melikyan, A.A.: Some differential games with incomplete information. In: Marchuk, G.I. (ed.) Optimization Techniques (IFIP Technical Conference, July 1-7, 1974, Novosibirsk, USSR), Lect. Notes in Computer Sci., Vol. 27, pp. 445–450. Springer-Verlag, Berlin (1974)
74. Melikyan, A.A.: Invariant manifolds of Hamiltonian systems in the theory of singularities of the Bellman–Isaacs equation. In: Matrosov, V.M., Ertegov, V.D. (eds.) Ustoichivost' dvizheniya (Stability of motion), pp. 227–230. Nauka, Sibirsk. Otdel., Novosibirsk (1985) (in Russian)
75. Melikyan, A.A.: The problem of time-optimal control with the search of the goal point. In: Sebastian, H.-J., Tammer, K. (eds.) System Modelling and Optimization, Lect. Notes in Control and Inform. Sci., Vol. 143, pp. 348–357. Springer, Berlin (1990)
76. Melikyan, A.A.: The method of characteristics for constructing singular paths and manifolds in optimal control and differential games. In: Hamalainen, R.P., Ehtamo H.K. (eds.) Differential Games - Developments in Modelling and Computation, Lect. Notes in Control and Inform. Sci., Vol. 156, pp. 81–90. Springer, Berlin (1991)
77. Melikyan, A.A.: Time-optimal control problem with goal point uncertainty. In: Kurzhanski, A.B., Veliov, V.M. (eds.) Modeling Techniques for Uncertain Systems, Progr. Systems Control Theory, Vol. 18, pp. 205–215. Birkhäuser, Boston (1994)
78. Melikyan, A.A.: Singular paths in differential games with simple motion. In: Basar, T., Haurie, A. (eds.) Advances in Dynamic Games and Applications, Ann. Internat. Soc. Dynam. Games, Vol. 1, pp. 125–135. Birkhäuser, Boston (1994)
79. Melikyan, A.A., Pourtallier, O.: Games with several pursuers and one evader with discrete observations. In: Petrosyan, L.A., Mazalov, V.V. (eds.) Game Theory and Applications, Vol. 2, pp. 169–184. Nova Sci. Publ., New York (1996)
80. Melikyan, A.A., Shinar, J.: Identification and construction of singular surfaces in pursuit-evasion games. In: Filar, J.A., Gaitsgory, V., Mizukami, K. (eds.) Advances in Dynamic Games and Applications, Ann. Internat. Soc. Dynam. Games, Vol. 5, pp. 151–175. Birkhäuser, Boston (2000)
81. Melikyan, A.A.: Singular characteristics of first-order PDEs in optimal control and differential games. J. Mat. Sci. **103(6)**, 745–755 (2001). Translated from Itogi Nauki i Tekhniki, Ser. Sovrem. Mat. Pril., Temat. Obz., Vol. 64, pp. 179–196. VINITI, Moscow (1999) (in Russian)
82. Melikyan, A.A.: Geometry of pursuit-evasion games on two-dimensional manifolds. In: Jorgensen, S., Quincampoix, M., Vincent, T.L. (eds.) Advances in Dynamic Games Theory: Numerical Methods, Algorithms, and Applications to Ecology and Economics, Ann. Internat. Soc. Dynam. Games, Vol. 9, pp. 173–194. Birkhäuser, Boston (2007)
83. Stipanović, D.M., Melikyan, A., Hovakimyan, N.: Some sufficient conditions for multi-player pursuit-evasion games with continuous and discrete observations. In: Bernhard, P., Gaitsgory, V., Pourtallier, O. (eds.) Advances in Dynamic Games and Their Applications. Analitical and Numerical Developments, Ann. Internat. Soc. Dynam. Games, Vol. 10, pp. 133–145. Birkhäuser, Boston (2009)

Scientific Reports

84. Melikyan, A.A., Ovakimyan, N.V.: Game problem of simple pursuit on the manifolds. Preprint no. 489, Institute for Problems in Mechanics, USSR Academy of Sciences, Moscow, 1991. 44 pp. (in Russian)
85. Melikyan, A.A., Ovakimyan, N.V.: Differential games of simple pursuit and approach on manifolds. Preprint, Institute of Mechanics, Armenian Academy of Sciences, Yerevan, 1993. 44 pp.
86. Melikyan, A.A., Botkin, N.D., Turova, V.L.: Propagation of disturbances in inhomogeneous anisotropic media. Preprint no. 181, Stiftung Caesar, Bonn, 2004. 25 pp.

Preprints and Proceedings of Conferences

87. Melikyan, A.: Singular paths in differential games with simple motion. In: Preprint Volume of the 5th International Symposium on Dynamic Games and Applications, 10 pp. Université de Genève (1992)
88. Melikyan, A.A., Shinar, J.: Identification and construction of singular surfaces in pursuit-evasion games. In: Breton, M., Zaccour, G. (eds.) Preprint Volume of the 6th International Symposium on Dynamic Games and Applications, pp. 517–529. GERAD, Ecole des H.E.C. (1994)
89. Melikyan, A.A.: What is Kelley condition from the point of view of synthesis. In: Shinar, J., Kirillova, F.M. (eds.) A Postprint Volume of the 10th IFAC Workshop "Control Applications of Optimization", pp. 47–50. Pergamon Press, London (1996)
90. Melikyan, A.A.: Equations for singular focal paths in dynamic games. In: Filar, J.A., Gaitsgory, V., Imado, F. (eds.) Proceedings of the 7th International Symposium on Dynamic Games and Applications, Vol. 2, pp. 690–695. Kanagawa (1996)
91. Melikyan, A.A., Pourtallier, O.: Games with several pursuers and one evader with discrete observations. In: Filar, J.A., Gaitsgory, V., Imado, F. (eds.) Proceedings of the 7th International Symposium on Dynamic Games and Applications, Vol. 2, pp. 672–689. Kanagawa (1996)
92. Melikyan, A.A.: Singular characteristics in optimal control and differential games. In: Huijberts, H.J.C., Nijmeijer, H., van der Schaft, A.J., Scherpen, J.M.A. (eds.) Proceedings of the 4th IFAC Symposium on Nonlinear Control Systems Design, Vol. 1, pp. 100–102. Elsevier-Pergamon, Oxford (1998)
93. Melikyan, A.A.: Games of pursuit on two-sided plane with a hole. In: Preprint Volume of the 8th International Symposium on Dynamic Games and Applications, pp. 394–396. Maastricht (1998)
94. Melikyan, A.A., Korneev, V.A.: Construction algorithm for non-smooth game value using singular characteristics. In: Preprint Volume of the 8th International Symposium on Dynamic Games and Applications, pp. 397–401. Maastricht (1998)
95. Melikyan, A.A.: Singular characteristics in dynamic programming. In: Proceedings of American Control Conference, Vol. 4, pp. 2857–2862. San Diego (1999)
96. Bolotnik, N.N., Melikyan, A.A.: Time-optimal control of a vibroimpact system with one degree of freedom. In: Chernousko, F.L., Fradkov, A.L. (eds.) Proceedings 2nd Intern. Conf. "Control of Oscillations and Chaos", Vol. 2, pp. 249–252. St. Petersburg (2000)
97. Hovakimyan, N., Melikyan, A.: Bifurcations in pursuit-evasion games on one-parameter family of surfaces. In: Nowak, A.S., Szajowski, K. (eds.) Proceedings of the Ninth International Symposium on Dynamic Games and Applications, Vol. 1, pp. 247–262. University of South Australia, Adelaide (2000)
98. Melikyan, A.A., Naumov, G.V.: Phase portrait bifurcation in one-parameter family of optimal control problems. In: Preprints of the 5th IFAC Symposium "Nonlinear Control Systems", pp. 1286–1291. St. Petersburg (2001)
99. Bolotnik, N.N., Melikyan, A.A.: Optimal control of a mechanical system with impacts. In: Preprints of the 5th IFAC Symposium "Nonlinear Control Systems", pp. 1590–1594. St. Petersburg (2001)

100. Melikyan, A.A., Naumov, G.V.: Phase portrait bifurcation in one-parameter family of optimal control problems. In: Kurzhanski, A.B., Fradkov, A.L. (eds.) Proceedings of the 5th IFAC Symposium "Nonlinear Control Systems", Vol. 3, pp. 1205–1210. Elsevier, Oxford (2002)
101. Bolotnik, N.N., Melikyan, A.A.: Optimal control of a mechanical system with impacts. In: Kurzhanski, A.B., Fradkov, A.L. (eds.) Proceedings of the 5th IFAC Symposium "Nonlinear Control Systems", Vol. 3, pp. 1505–1510. Elsevier, Oxford (2002)
102. Melikyan, A.A.: Construction of singular surfaces in multiple integral variational problem. In: Camacho, E.F., Basanez, L., de la Puente, J.A. (eds.) Proceedings of the 15th Triennial World Congress of the IFAC, Vol. C: Stability and Nonlinear Systems, pp. 465–470. Barcelona (2002)
103. Melikyan, A.A.: Pursuit-evasion games on the surfaces of revolution. In: Petrosjan, L.A., Zenkevich, N.A. (eds.) Proceedings of the Tenth International Symposium on Dynamic Games and Applications, Vol. 2: International Society of Dynamic Games, pp. 617–622. Saint-Petersburg State University, Inst. Chem., St. Petersburg (2002)
104. Melikyan, A., Hovakimyan, N.: Some variation calculus problems in dynamic games on 2D surfaces. In: Proc. Vol. Intern. Congr. of Math., Satellite Conf. on Game Theory and Appl., pp. 513–522. Quingdao Publ. House, Quingdao (2002)
105. Melikyan, A., Hovakimyan, N.: Some variational calculus problems in dynamic games on 2D surfaces. In: Petrosyan, L., Yeung, D. (eds.) Proceedings ICM Millennium Lectures on Games, pp. 287–296. Birkhäuser, Berlin, Heidelberg, New York (2003)
106. Melikyan, A., Hovakimyan, N., Ikeda, Y.: Dynamic programming approach to a minimum distance optimal control problem. In: Proceedings of the 42nd IEEE Conference on Decision and Control, pp. 239–244. Maui, Hawaii (2003)
107. Melikyan, A.: Geometry of Pursuit-evasion Games on 2D Manifolds. In: Vincent, T.L. (ed.) Proceedings of the 11th International Symposium on Dynamic Games and Applications, Vol. 2, pp. 650–669. University of Arizona, Tucson (2004)
108. Melikyan, A., Akhmetzhanov, A.: A linear 2D differential game with a minimum-type cost function. In: Vincent, T.L. (ed.) Proceedings of the 11th International Symposium on Dynamic Games and Applications, Vol. 2, pp. 637–649. University of Arizona, Tucson (2004)
109. Melikyan, A., Akhmetzhanov, A., Hovakimyan, N.: A remark on IVP and TVP non-smooth viscosity solutions to Hamilton–Jacobi equations. In: Proceedings of American Control Conference, Vol. 2, pp. 864–869. Portland (2005)
110. Melikyan, A., Botkin, N., Turova, V.: Modelling of disturbance propagation in anisotropic media using differential game methods. In: Horacek, P., Simandl, M., Zitek, P. (eds.) Preprints of the 16 IFAC World Congress, 6 pp. Prague (2005)
111. Melikyan, A.A., Botkin, N.D., Turova, V.L.: Singularity phenomena in ray propagation in anisotropic media. In: Fradkov, A.L., Churilov, A.N. (eds.) Proceedings of the International Conference "Physics and Control", pp. 373–378. St. Petersburg (2005)
112. Stipanović, D.M., Melikyan, A., Hovakimyan, N.: Some sufficient conditions for multi-player pursuit-evasion games with continuous and discrete observations. In: Proceedings of the 12th International Symposium on Dynamic Games and Applications, 11 pp. Sophia Antipolis (2006)
113. Melikyan, A., Hovakimyan, N., Akhmetzhanov, A.: Application of singular characteristics for image processing. In: Proceedings of the 12th International Symposium on Dynamic Games and Applications, 7 pp. Sophia Antipolis (2006)
114. Melikyan, A.A., Olsder, G.J.: Feedback solution of a non-zero-sum game via reduction to a zero-sum game. In: Proceedings of the 12th International Symposium on Dynamic Games and Applications, 16 pp. Sophia Antipolis (2006)
115. Melikyan, A.A.: Some applications of boundary characteristics of Hamilton–Jacobi equation. In: Proceedings of the IX International Chetayev Conference "Analytical Mechanics, Stability and Control of Motion", Vol. 1, pp. 114–118. Institute for System Dynamics and Control Theory of Siberian Branch of the Russian Academy of Sciences, Irkutsk (2007)

116. Stipanović, D.M., Melikyan, A., Hovakimyan, N.: Guaranteed strategies for nonholonomic players in pursuit-evasion games. In: Melikyan, A.A., Nowak, A.S., Szajowski, K.J. (eds.) 13th International Symposium on Dynamic Games and Applications, 3 pp. Wroclaw (2008)
117. Melikyan, A., Olsder, G.J., Akhmetzhanov, A.: A sufficient condition for the reduction of a nonzero-sum game to a zero-sum one. In: Melikyan, A.A., Nowak, A.S., Szajowski, K.J. (eds.) 13th International Symposium on Dynamic Games and Applications, 14 pp. Wroclaw (2008)

The Long History of Arik's Conjecture

Pierre Bernhard

Long ago, when both Arik Melikyan and I were young – we are of the same age – at an international conference, I was giving a talk on my constructive theories for singular surfaces of two-person zero-sum differential games. I had constructive equations for all the classical menagerie: dispersal lines, universal lines, equivocal lines, switch envelopes... but one piece was missing: I could only account for one-dimensional focal lines in two-dimensional D.G.'s. For higher-dimensional focal manifolds, I could only offer a conjecture, which, if true, would let me propose a constructive theory.

Soon after my talk, Arik came up to me, telling me that he did not believe my conjecture. Contrary to my account, he conjectured that such hyper-surfaces should be traversed by *two non-colinear fields of trajectories*. If he were right, as he soon convinced me, we were left with an intriguing phenomenon, but still no constructive theory. And we lacked an example.

We continued discussing this issue by mail and each time we met for years. Some years later, he had examples of singularities similar to his conjecture in mathematical physics. But we still had no example in D.G.'s, and no certified way of constructing one. We decided to try and reverse engineer a D.G. from one of his examples. But this failed.

In January 2002, we met at the Lyapunov Institute in Moscow. I had in hand a complicated and uncertain example of what might have been an instance of Arik's conjecture. I remember that I had written him a couple of weeks before, telling him that I had a pair of coupled, first-order PDEs, and that his first answer had been, "But nobody does that!" I showed him my proposed example. He quickly agreed that it was indeed an instance of what we had been after for all those years.

Following this, he was able to get a scholarship from the Lyapunov Institute to come and spend some time in Sophia Antipolis, where we would work together. There, he quickly understood why I had these two PDEs, and made a general

P. Bernhard (✉)
University of Nice-Sophia Antipolis and INRIA Sophia Antipolis Méditerranée, France
e-mail: Pierre.Bernhard@sophia.inria.fr

theory of higher-dimensional focal hyper-surfaces, including a constructive theory involving a pair of first-order PDEs. This theory was his, even though it was published in a joint paper [1], because the original example was mine.

Let me add that the original exposition of this example – in mathematical finance – which can be found in a paper at the 2002 ISDG conference [2], was really difficult to understand. With Arik's theory, in our joint paper, it takes less than one page. Subsequently my Ph.D. student Stéphane Thiery showed that by using Arik's conjecture – now, theorem – he could derive the focal surface in that example in a few lines.

Over time, we became good friends, as Arik was such a kind person. I spontaneously turned to him for any question about characteristic theory of Isaacs' equation and many other topics in differential games. We also started a collaboration with his Ph.D. student, Andrei Akhmetzhanov. I was seeking, through Andrei, Arik's opinion on our ongoing work on singular surfaces in Nash equilibria of nonzero-sum games when, on the morning of April 7, I learned from Valerii Patsko of his sudden death. This was a shock and left me, like all his many friends, quite saddened.

References

1. Arik Melikyan and Pierre Bernhard. Geometry of optimal trajectories around a focal singular surface in differential games. *Applied Mathematics and Optimization*, 52(1):23–37, 2005.
2. Pierre Bernhard, Naïma El Farouq, and Stéphane Thiery. An impulsive differential game arising in finance with interesting singularities. In Alain Haurie, Shigeo Muto, Leon A. Petrosjan, and T.E.S. Raghavan, editors, *10th ISDG International Symposium on Dynamic Games and Applications, Saint Petersburg, 2002*, volume 8 of *Annals of the ISDG: Advances in Dynamic Games*, pages 335–363. 2006.

Optimal Control Problem for a Dynamical System on an Ellipsoid*

Arik A. Melikyan† and Fumiaki Imado

Abstract A point mass is considered whose motion is controlled in three-dimensional space by a force of a bounded absolute value. The presence of gravity is assumed. The task of the control is to bring the point from a given initial position to a terminal position with minimal time, while the trajectory is required to be on the surface of the given ellipsoid. Different terminal conditions are considered. The problem is investigated using a state-constraint maximum principle.

1 Introduction

Differential games and optimal control problems for dynamical systems on the surfaces and manifolds are considered in [1–4].

2 Dynamics of a Point Mass on an Ellipsoid

The dynamics are given by the differential equations and control constraints:

$$\ddot{x} = u_1, \quad \ddot{y} = u_2, \quad \ddot{z} = u_3 + g \tag{1}$$
$$u_1^2 + u_2^2 + u_3^2 \le \mu^2, \quad \mu = \frac{F_*}{m}.$$

*Note from the editors: This paper reports on work in progress that was initiated during A. Melikyan's stay in Japan, under the auspices of JSPS. It has not been published except in a report for JSPS.

† Deceased
A.A. Melikyan
Institute for Problems in Mechanics, RAS, Moscow, Russia

F. Imado (✉)
Faculty of Engineering "Measurement and Control System, Department of Mechanical Systems Engineering", Shinshu University 4-17-1 Wakasato Nagano, Nagano 380-8553, Japan
e-mail: imadofu@shinshu-u.ac.jp

Here, m is the mass of the considered point, F_* is the maximal absolute value of the control force, u_i are normalized force components (control parameters), μ is the normalized absolute value of the control vector, and g is the gravity constant. The point mass is restricted to move on the surface of the ellipsoid:

$$R(x, y, z) \equiv \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0. \tag{2}$$

Differentiating twice the function $R(x, y, z)$ along the solutions of the system (1) (until the control parameters explicitly arise, see [2]), one can get the following three necessary conditions:

$$R(x, y, z) = 0, \quad \dot{R} = \langle n, v \rangle = k_1 x \dot{x} + k_2 y \dot{y} + k_3 z \dot{z} = 0 \tag{3}$$

$$\ddot{R} = k_1 \dot{x}^2 + k_2 \dot{y}^2 + k_3 \dot{z}^2 + k_1 x u_1 + k_2 y u_2 + k_3 z (u_3 + g) = 0.$$

Here v is the velocity vector, and $n = (n_1, n_2, n_3)$ is a normal vector to the surface, the gradient of the function R:

$$v = (\dot{x}, \dot{y}, \dot{z}), \quad n = \nabla R = (k_1 x, k_2 y, k_3 z) \quad \left(k_1 = \frac{2}{a^2}, k_2 = \frac{2}{b^2}, k_3 = \frac{2}{c^2} \right).$$

Generally, the dynamical system (1) has six phase coordinates $(x, \dot{x}, y, \dot{y}, z, \dot{z})$ and three control variables, $u = (u_1, u_2, u_3)$. Using the first two equations in (3), one can reduce the number of coordinates to four, and using the last equation one can get a two-dimensional control vector. If, in addition, the control force of maximal absolute value is used (which is often recommended by the maximum principle), i.e., $\mid u \mid = \mu$, then one more control parameter is eliminated, and one gets a scalar control. Thus, the dynamic system on an ellipsoid can be characterized by four coordinates and one control parameter.

3 Dynamics of a Point Mass on a Solid Ellipsoid

When a point mass moves along a solid (stiff) surface then, generally, three forces act on a point: the control force, the external force (gravity), and the reaction force of the surface. In case of frictionless motion, the reaction force is orthogonal to the surface, and the dynamic equations take the form:

$$\ddot{x} = N_1 + u_1, \quad \ddot{y} = N_2 + u_2, \quad \ddot{z} = N_3 + u_3 + g, \tag{4}$$

where $N = \lambda n$. The control force is assumed to be tangent to the surface, which gives the restriction:

$$\langle u, n \rangle = u_1 n_1 + u_2 n_2 + u_3 n_3 = 0. \tag{5}$$

The parameters N_i can be excluded by writing the Lagrange equation of the second kind for these systems with two degrees of freedom. The resulting dynamic equations include four phase variables and two control parameters. For motion with maximal force, one has only one control variable.

4 Isaacs' Model of a Car

The control force here is orthogonal to the point's velocity:

$$\langle u, v\rangle = u_1\dot{x} + u_2\dot{y} + u_3\dot{z} = 0 \tag{6}$$

so that the vector v does not change its length during the motion. Here, the dynamic equation can be reduced to four phase variables and one control variable. A motion with maximal value of the force does not contain any control freedom. For the plane model of a car, when $R = 0$ determines a plane and there is no gravity, it is known that time-optimal trajectories consist of segments of straight lines (the force vanishes and the point performs a free motion) and parts of circles (the force has maximal absolute value). The straight lines are the shortest lines (geodesics) in a plane. This property also takes place for the motion on a general surface, as shown in the next section. However, the motion with maximal force, generally, is a spiral on a surface rather than a closed trajectory.

5 Free Motion on a Surface (Ellipsoid)

Consider the motion of a mass point on a (stiff) surface in the absence of friction, gravity, and other external forces. The dynamic equations have the form:

$$\ddot{x} = N_1, \quad \ddot{y} = N_2, \quad \ddot{z} = N_3, \tag{7}$$

where $N = (N_1, N_2, N_3)$ is the force acting from the surface to the mass point. In the absence of friction, it is a normal force, so that

$$N = \lambda n$$

for some scalar, and the equations take the form:

$$\ddot{x} = \lambda n_1, \ \ \ddot{y} = \lambda n_2, \ \ \ddot{z} = \lambda n_3. \tag{8}$$

One can show that these equations of free motion have the solutions which are the geodesic lines of the surface – the shortest lines connecting two points of the surface. Geodesic lines on the surface are the solutions of the following variational problem, with one restriction:

$$J = \int_{\theta_0}^{\theta_1} \sqrt{(\mathrm{d}x/\mathrm{d}\theta)^2 + (\mathrm{d}y/\mathrm{d}\theta)^2 + (\mathrm{d}z/\mathrm{d}\theta)^2}\mathrm{d}\theta \to \min \tag{9}$$

$$R(x, y, z) = 0,$$

where $(x(\theta), y(\theta), z(\theta)), \theta_0 \leq \theta \leq \theta_1$, is a parametrization of the curve searched for. The classical variation a calculus says that the solution of the above problem is the solution of an unconditioned problem with the integrand:

$$F(x, y, z, \acute{x}, \acute{y}, \acute{z}) = \sqrt{(\mathrm{d}x/\mathrm{d}\theta)^2 + (\mathrm{d}y/\mathrm{d}\theta)^2 + (\mathrm{d}z/\mathrm{d}\theta)^2} + \bar{\mu} R(x, y, z),$$

with some scalar $\bar{\mu} = \bar{\mu}(x, y, z)$. Here the prime sign means differentiation with respect to θ. The Euler system of equations

$$F_x - \frac{\mathrm{d}}{\mathrm{d}\theta} F_{\acute{x}} = 0, \;\; F_y - \frac{\mathrm{d}}{\mathrm{d}\theta} F_{\acute{y}} = 0, \;\; F_z - \frac{\mathrm{d}}{\mathrm{d}\theta} F_{\acute{z}} = 0$$

for such a problem for the surface $R(x, y, z) = 0$ takes the form

$$x'' = \bar{\mu} n_1, \;\; y'' = \bar{\mu} n_2, \;\; z'' = \bar{\mu} n_3. \tag{10}$$

One can see that the solutions of (8) and (10) coincide, the only difference being the parametrization of the curves. In the case of an ellipsoid, these curves are appropriate ellipses.

References

1. Isaacs, R. Differential Games. *Wiley, New York*, 1965.
2. Bryson, A.E., Ho, Yu-Chi “Larry.” Applied Optimal Control. *Washington, DC: Hemisphere*, 1975.
3. Imado, F. Optimal Control Problem on an Ellipsoid, JSPS report, 2006.
4. Melikyan A.A. Generalized Characteristics of First Order PDEs: Applications in Optimal Control and Differential Games. *Birkhauser, Boston*, 1998.

Contributors

Elvio Accinelli Facultad de Economia, UASLP, Mexico, elvio.accinelli@eco.uaslp.mx

Eitan Altman INRIA Sophia Antipolis Méditerranée, 2004, route des lucioles, 06902, Sophia Antipolis, France, eitan.altman@sophia.inria.fr

Yurii Averboukh Institute of Mathematics and Mechanics UrB RAS, S. Kovalevskaya street 16, 620219, GSP-384, Ekaterinburg, Russia, ayv@imm.uran.ru

Teresa-Paula Azevedo-Perdicoúlis ISR-pólo de Coimbra & Departamento Matemática, UTAD, 5000-911 Vila Real, Portugal, tazevedo@utad.pt

Pierre Bernhard University of Nice-Sophia Antipolis and INRIA Sophia Antipolis Méditerranée, France, Pierre.Bernhard@sophia.inria.fr

Alberto Bressan Department of Mathematics, Penn State University, University Park, PA 16802, USA, bressan@math.psu.edu

Juan Gabriel Brida School of Economics, Free University of Bozen, Bolzano, Italy, juangabriel.brida@unibz.it

Vivek S. Borkar School of Technology and Computer Science, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400005, India, borkar@tifr.res.in

Dean A. Carlson American Mathematical Society, Mathematical Reviews, 416 Fourth Street, Ann Arbor, MI 48103, USA, dac@ams.org

Edgar J.S. Carrera Department of Economics, University of Siena, Italy, sanchezcarre@unisi.it

Arkadii Chikrii Glushkov Institute of Cybernetics NASU, 40 Glushkov Prsp 03680, MSP Kiev 187, Ukraine, chikrii@gmail.com

Yan Dolinsky Institute of Mathematics, The Hebrew University, Jerusalem 91904, Israel, yann1@math.huji.ac.il

Laurent Drouet Resource Centre for Environmental Technologies, Public Research Centre Henri Tudor, Technoport Schlassgoart, 66 rue de Luxembourg, L-4002 Esch-sur-Alzette, Luxembourg, laurent.drouet@tudor.lu

Rachid El-Azouzi Laboratoire Informatique d'Avignon/CERI, Université d'Avignon, 84911, Avignon, France, rachid.elazouzi@univ-avignon.fr

Elżbieta Z. Ferenstein Warsaw University of Technology, Faculty of Mathematics and Information Science, Pl. Politechniki 1, 00-661 Warszawa, Poland, efer@mini.pw.edu.pl

Sergei A. Ganebny Institute of Mathematics and Mechanics, Ural Branch of Russian Academy of Sciences, S. Kovalevskaya str. 16, Ekaterinburg, Russia, ganebny@imm.uran.ru

Saïd Hamadène Laboratoire Manceau de Mathématiques, Université du Maine, 72085 Le Mans Cedex 9, France, hamadene@univ-lemans.fr

Alain Haurie ORDECSYS, Chêne-Bougeries, Switzerland, ahaurie@ordecsys.com

Yezekael Hayel Laboratoire Informatique d'Avignon/CERI, Université d'Avignon, 84911, Avignon, France, yezekael.hayel@univ-avignon.fr

Ryusuke Hohzaki Department of Computer Science, National Defense Academy, Hashirimizu, Yokosuka, 239-8686, Japan, hozaki@cc.nda.ac.jp

Fumiaki Imado Faculty of Engineering "Measurement and Control System, Department of Mechanical Systems Engineering", Shinshu University 4-17-1 Wakasato Nagano, Nagano 380-8553, Japan, imadofu@shinshu-u.ac.jp

Gerhard Jank Universidade de Aveiro, 3810-193 Aveiro, Portugal, jank@math2.rwth-aachen.de

Liudmila Kamneva Institute of Mathematics and Mechanics, S. Kovalevskaya str., 16, 620219, GSP-384, Ekaterinburg, Russia, kamneva@imm.uran.ru

Yuri Kifer Institute of Mathematics, The Hebrew University, Jerusalem 91904, Israel, kifer@math.huji.ac.il

Mikhail Krastanov Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Acad. G. Bonchev Str. Bl.8, 1113 Sofia, Bulgaria, krast@math.bas.bg

Heinz-Uwe Küenle Brandenburg University of Technology Cottbus, Department of Mathematics, PF 10 13 44, D-03013 Cottbus, Germany, kuenle@math.tu-cottbus.de

Sergey S. Kumkov Institute of Mathematics and Mechanics, Ural Branch of Russian Academy of Sciences, S. Kovalevskaya str. 16, Ekaterinburg, Russia, kumkov.jr@imm.uran.ru

Joy Kuri Centre for Electronics Design and Technology, Indian Institute of Science, Bangalore 560012, India, kuri@cedt.iisc.ernet.in

George Leitmann Graduate School, University of California at Berkeley, Berkeley, CA 94720, USA, gleit@berkeley.edu

Stéphane Le Ménec EADS/MBDA France, 1 avenue Réaumur, 92358 Le Plessis-Robinson Cedex, France, stephane.le-menec@mbda-systems.com

Ivan Matychyn Glushkov Institute of Cybernetics NASU, 40 Glushkov Prsp 03680, MSP Kiev 187, Ukraine, atychyn@gmail.com

Arik A. Melikyan Institute for Problems in Mechanics, RAS, Moscow, Russia

Adam Pasternak-Winiarski Warsaw University of Technology, Faculty of Mathematics and Information Science, Pl. Politechniki 1, 00-661 Warszawa, Poland, A.Pasternak-Winiarski@mini.pw.edu.pl

Valerii S. Patsko Institute of Mathematics and Mechanics, Ural Branch of Russian Academy of Sciences, S. Kovalevskaya str. 16, Ekaterinburg, Russia, patsko@imm.uran.ru

Yulia Pavlova Economic Research, MTT Agrifood Research Finland, Luutnantintie 13, Helsinki 00410, Finland, yulia.pavlova@mtt.fi

Sergei G. Pyatko Academy of Civil Aviation, Pilotov str., 38, St. Petersburg, 196210, Russia, boss@nita.ru; nita@nita.ru

David M. Ramsey Department of Mathematics and Statistics, University of Limerick, Limerick, Ireland, david.ramsey@ul.ie

Anna Rettieva Institute of Applied Mathematical Research, Karelian Research Center of RAS, Pushkinskaya str. 11, Petrozavodsk, 185910, Russia, annaret@krc.karelia.ru

Rossen Rozenov Bulgarian National Bank, 1 King Alexander Sq., 1000 Sofia, Bulgaria, rossen_rozenov@yahoo.com

Sriram Shankaran General Electric Research Center, Niskayuna, NY, USA, sshankaran@gmail.com

Wen Shen Department of Mathematics, Penn State University, University Park, PA 16802, USA, shen_w@math.psu.edu

Ekaterina Shevkoplyas St. Petersburg State University, Faculty of Applied Mathematics, University pr., 35, Petrodvorets, 198 904 St. Petersburg, Russia, ekaterina.shevkoplyas@gmail.com

Dušan M. Stipanović Department of Industrial and Enterprise Systems Engineering, and Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Illinois, USA, dusan@illinois.edu

Nina N. Subbotina Institute of Mathematics and Mechanics UrB RAS, S. Kovalevskaja str., 16, Ekaterinburg, 620219, Russia, subb@uran.ru

Andrzej Święch School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA, swiech@math.gatech.edu

Hamidou Tembine Laboratoire Informatique d'Avignon/CERI, Université d'Avignon, 84911, Avignon, France, tembineh@yahoo.fr

Timofey B. Tokmantsev Institute of Mathematics and Mechanics UrB RAS, S. Kovalevskaja str., 16, Ekaterinburg, 620219, Russia, tokmantsev@imm.uran.ru

Claire J. Tomlin Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, USA, tomlin@eecs.berkeley.edu

Varvara L. Turova Technische Universität München, Boltzmannstr. 3, 85747 Garching bei München, Germany, turova@ma.tum.de

Jean-Philippe Vial University of Geneva and ORDECSYS, Switzerland, jpvial@ordecsys.com

Marc Vielle Ecole Polytechnique Fédérale de Lausanne and C-ORDEE, Switzerland and Toulouse School of Economics (LERNA), TSE - Manufacture de Tabacs, 21 allées de Brienne - 31000 Toulouse, France, marc.vielle@epfl.ch

Hao Wang Laboratoire Manceau de Mathématiques, Université du Maine, 72085 Le Mans Cedex 9, France
and
School of Mathematics, Shandong University, Jinan 250100, P. R. China, hao.wang.etu@univ-lemans.fr

David W.K. Yeung Department of Business Administration, Hong Kong Shue Yan University, Hong Kong
and
Center of Game Theory, St. Petersburg State University, St. Petersburg, Russia, dwkyeung@hksyu.edu

Part I
Dynamic-Differential Games: Theoretical Developments

1. *Yurii Averboukh*, On a Structure of the Set of Differential Games Values
2. *Teresa-Paula Azevedo-Perdicoúlis and Gerhard Jank*, Existence and Uniqueness of Disturbed Open-Loop Nash Equilibria for Affine-Quadratic Differential Games
3. *Dean A. Carlson and George Leitmann*, Fields of Extremals and Sufficient Conditions for a Class of Open-Loop Variational Games
4. *Arkadii Chikrii and Ivan Matychyn*, Riemann–Liouville, Caputo, and Sequential Fractional Derivatives in Differential Games
5. *Saïd Hamadène and Hao Wang*, The Mixed Zero-Sum Stochastic Differential Game in the Model with Jumps
6. *Liudmila Kamneva*, Discontinuous Value Function in Time-Optimal Differential Games
7. *Nina N. Subbotina and Timofey B. Tokmantsev,* On Grid Optimal Feedbacks to Control Problems of Prescribed Duration on the Plane
8. *Andrzej Święch*, Sub- and Super-optimality Principles and Construction of almost Optimal Strategies for Differential Games in Hilbert Spaces

On a Structure of the Set of Differential Games Values

Yurii Averboukh

Abstract The set of value functions of all-possible zero-sum differential games with terminal payoff is characterized. The necessary and sufficient condition for a given function to be a value of some differential game with terminal payoff is obtained.

1 Introduction

This chapter is devoted to the theory of two-controlled, zero-sum differential games. Within the framework of this theory, the control processes under uncertainty are studied. N.N. Krasovskii and A.I. Subbotin introduced the feedback formalization of differential games [5]. The formalization allows them to prove the existence of the value function.

We characterize the set of value functions of all-possible zero-sum differential games with terminal payoff. The value function is a minimax (or viscosity) solution of the corresponding Isaacs–Bellman equation (Hamilton–Jacobi equation) [7].

One can consider a differential game within usual constraints as a complex of two control spaces, game dynamic and terminal payoff function. The time interval and state space of the game are assumed to be fixed. The following problem is considered: let a locally lipschitzian function $\varphi(t, x)$ be given, do there exist control spaces, a dynamic function and a terminal payoff function such that the function $\varphi(t, x)$ is the value of the corresponding differential game?

2 Preliminaries

In this section, we recall the main notions of the theory of zero-sum differential games. We follow the formalization of N.N. Krasovskii and A.I. Subbotin.

Y. Averboukh (✉)
Institute of Mathematics and Mechanics UrB RAS, S. Kovalevskaya street 16, 620219, GSP-384, Ekaterinburg, Russia
e-mail: ayv@imm.uran.ru

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_1,

Usually in the theory of differential games, the following problem is considered. Let the controlled system

$$\dot{x} = f(t, x, u, v), \;\; t \in [t_0, \vartheta_0], \;\; x \in \mathbb{R}^n, \;\; u \in P, \;\; v \in Q$$

and the payoff functional $\sigma(x(\vartheta_0))$ be given. Here, u and v are controls of the player U and the player V, respectively. The player U tries to minimize the payoff and the player V wishes to maximize the payoff. The problem is to find the value of the game. The value is a function of position.

Suppose that P and Q are finite-dimensional compacts. The function f satisfies the following assumption:

F1. f is continuous.
F2. f is locally lipschitzian with respect to the phase variable.
F3. There exists a constant Λ_f such that for every $t \in [t_0, \vartheta_0]$, $x \in \mathbb{R}^n$, $u \in P$, $v \in Q$ the following inequality holds:

$$\|f(t, x, u, v)\| \leq \Lambda_f(1 + \|x\|).$$

The function $\sigma : \mathbb{R}^n \to \mathbb{R}$ satisfies the following assumption (see [1], [7]):

$\Sigma 1$. σ is locally lipshitzian.
$\Sigma 2$. There exists Λ_σ such that $|\sigma(x)| \leq \Lambda_\sigma(1 + \|x\|)$.

Assumption $\Sigma 1$ guarantees the locally Lipschitzeanity of the value function. Assumption $\Sigma 1$ is often replaced by the condition of continuity of σ.

We use the control design suggested in [5]. Player U chooses the control in the class of counter-strategies, and player V chooses the control in the class of feedback strategies. N.N. Krasovskii and A.I. Subbotin proved that the value function is well defined in this case. Let us denote the value function by $Val^f(\cdot, \cdot, P, Q, f, \sigma)$. A.I. Subbotin proved that the value of the differential game is the minimax solution of the Cauchy problem for Hamilton–Jacobi equation

$$\varphi(\vartheta_0, x) = \sigma(x) \tag{1}$$

$$\frac{\partial \varphi(t, x)}{\partial t} + H(t, x, \nabla\varphi(t, x)) = 0. \tag{2}$$

Here, $\nabla\varphi(t, x)$ means the vector of partial derivatives of φ with respect to the space variables. The function H is called the Hamiltonian of the differential game. It is defined in the following way.

$$H(t, x, s) \triangleq \max_{v \in Q} \min_{u \in P} \langle s, f(t, x, u, v)\rangle. \tag{3}$$

A.I. Subbotin introduced several definitions of minimax solution of the Hamilton–Jacobi equation [7]. He proved that they are equivalent. Also A.I. Subbotin proved that the notion of minimax solution coincides with the notion of viscosity solution (see [1] and [7]).

Below we use one of the equivalent definitions of minimax solution. The function $\varphi(t,x)$ is called a minimax solution of Hamilton–Jacobi equation (2), if for every $(t,x) \in (t_0,\vartheta_0) \times \mathbb{R}^n$ the following conditions are fulfilled:

$$a + H(t,x,s) \le 0 \;\; \forall (a,s) \in D^-_{\mathrm{D}}\varphi(t,x); \tag{4}$$

$$a + H(t,x,s) \ge 0 \;\; \forall (a,s) \in D^+_{\mathrm{D}}\varphi(t,x); \tag{5}$$

Here, we use the notions of nonsmooth analysis [2]. Sets $D^-_{\mathrm{D}}\varphi(t,x)$ and $D^+_{\mathrm{D}}\varphi(t,x)$ are called Dini subdifferential and Dini superdifferential, respectively. They are defined by the following rules.

$$D^-_{\mathrm{D}}\varphi(t,x) \triangleq \Big\{(a,s) \in \mathbb{R} \times \mathbb{R}^n : \\ a\tau + \langle s, g\rangle \le \liminf_{\alpha \to 0} \frac{\varphi(t+\alpha\tau, x+\alpha g) - \varphi(t,x)}{\alpha} \;\; \forall (\tau, g) \in \mathbb{R} \times \mathbb{R}^n\Big\},$$

$$D^+_{\mathrm{D}}\varphi(t,x) \triangleq \Big\{(a,s) \in \mathbb{R} \times \mathbb{R}^n : \\ a\tau + \langle s, g\rangle \ge \limsup_{\alpha \to 0} \frac{\varphi(t+\alpha\tau, x+\alpha g) - \varphi(t,x)}{\alpha} \;\; \forall (\tau, g) \in \mathbb{R} \times \mathbb{R}^n\Big\}.$$

The function φ is locally lipshitzian if σ is lipshitzian. One can prove this statement using definition (M1) of minimax solution [7]. Denote the differentiability set of φ by J, $J \subset (t_0,\vartheta_0) \times \mathbb{R}^n$. By the Rademacher's theorem [3], the measure $([t_0,\vartheta_0] \times \mathbb{R}^n) \setminus J$ is 0, therefore the closure of J is equal to $[t_0,\vartheta_0] \times \mathbb{R}^n$. For $(t,x) \in J$ full derivative of φ is $(\partial\varphi(t,x)/\partial t, \nabla\varphi(t,x))$.

If $D^+_{\mathrm{D}}\varphi(t,x)$ and $D^-_{\mathrm{D}}\varphi(t,x)$ are nonempty simultaneously, then φ is differentiable at (t,x), and $D^+_{\mathrm{D}}\varphi(t,x) = D^-_{\mathrm{D}}\varphi(t,x) = \{(\partial\varphi(t,x)/t, \nabla\varphi(t,x))\}$ [2]. If φ is differentiable at some position (t,x), then equality (2) is valid at the position (t,x) in the ordinary sense.

We have [2]

$$D^-_{\mathrm{D}}\varphi(t,x), D^+_{\mathrm{D}}\varphi(t,x) \subset \partial_{\mathrm{Cl}}\varphi(t,x).$$

Here, $\partial_{\mathrm{Cl}}\varphi(t,x)$ denotes the Clarke subdifferential. Since φ is locally lipshitzian, the Clarke subdifferential can be expressed in the following way [2]:

$$\partial_{\mathrm{Cl}}\varphi(t,x) \\ = \mathrm{co}\left\{(a,s) : \exists \{(t_i,x_i)\}_{i=1}^{\infty} \subset J : \;\; a = \lim_{i\to\infty} \frac{\partial\varphi(t_i,x_i)}{\partial t}, s = \lim_{i\to\infty} \nabla\varphi(t_i,x_i)\right\}.$$

Let us describe the properties of the Hamiltonian.

First, let us introduce a class of real-valued functions. This class will be used extensively throughout this paper. Denote by Ω the set of all even semiadditive functions $\omega : \mathbb{R} \to [0, +\infty)$ such that $\omega(\delta) \to 0$, $\delta \to 0$.

The following conditions are valid with $\Upsilon = \Lambda_f$ (see [7]):

H1. (Sublinear growth condition) for all $(t, x, s) \in \mathbb{R}^n$

$$|H(t, x, s)| \leq \Upsilon \|s\|(1 + \|x\|);$$

H2. For every bounded region $A \subset \mathbb{R}^n$ there exist a function $\omega_A \in \Omega$ and a constant L_A such that for all $(t', x', s'), (t'', x'', s'') \in [t_0, \vartheta_0] \times A \times \mathbb{R}^n$, $\|s'\|, \|s''\| \leq R$ the following inequality holds:

$$\begin{aligned}&\|H(t', x', s') - H(t'', x'', s'')\| \\ &\quad \leq \omega_A(t' - t'') + L_A R\|x' - x''\| + \Upsilon(1 + \inf\{\|x'\|, \|x''\|\})\|s_1 - s_2\|;\end{aligned}$$

H3. H is positively homogeneous with respect to the third variable:

$$H(t, x, \alpha s) = \alpha H(t, x, s)\ \forall (t, x) \in [t_0, \vartheta_0] \times \mathbb{R}^n,\ \forall s \in \mathbb{R}^n\ \forall \alpha \in [0, \infty).$$

3 Main Result

In this section, we study the set of values of all possible differential games. The main result is formulated below.

Denote by COMP the set of all finite-dimensional compacts. Let $P, Q \in$ COMP, denote by DYN(P, Q) the set of all functions $f : [t_0, \vartheta_0] \times \mathbb{R}^n \times P \times Q \to \mathbb{R}^n$ satisfying Conditions F1–F3. The set of functions $\sigma : \mathbb{R}^n \to \mathbb{R}$ satisfying Condition $\Sigma 1$ and $\Sigma 2$ is denoted by TP.

The set of values of differential games is described in the following way.

$$\begin{aligned}\mathrm{VALF} = \{&\varphi : [t_0, \vartheta_0] \times \mathbb{R}^n \to \mathbb{R} : \\ &\exists P, Q \in \mathrm{COMP} \exists f \in \mathrm{DYN}(P, Q) \exists \sigma \in \mathrm{TP} :\ \varphi = Val^f(\cdot, \cdot, P, Q, f, \sigma)\}.\end{aligned}$$

Denote by Lip_B the set of all locally lipschitzian functions $\varphi : [t_0, \vartheta_0] \times \mathbb{R}^n \to \mathbb{R}$ such that $\varphi(\vartheta_0, \cdot)$ satisfies the sublinear growth condition. The set VALF is subset of the set Lip_B, since we consider only locally lipschitzian payoff functions.

Let $\varphi \in \mathrm{Lip}_B$. Our task is to examine the inclusion $\varphi \in$ VALF. Denote the differentiability set of φ by J. For $(t, x) \in J$ set

$$\begin{aligned}E_1(t, x) &\triangleq \{\nabla\varphi(t, x)\}; \\ h(t, x, \nabla\varphi(t, x)) &\triangleq -\frac{\partial\varphi(t, x)}{\partial t}.\end{aligned} \tag{6}$$

Put the following condition.

`Condition (E1).` For any position $(t_*, x_*) \notin J$, and any sequences $\{(t_i', x_i')\}_{i=1}^{\infty}$, $\{(t_i'', x_i'')\}_{i=1}^{\infty} \subset J$ such that $(t_i', x_i') \to (t_*, x_*)$, $i \to \infty$, $(t_i'', x_i'') \to (t_*, x_*)$, $i \to \infty$, the following implication holds:

$$(\lim_{i\to\infty} \nabla\varphi(t_i', x_i') = \lim_{i\to\infty} \nabla\varphi(t_i'', x_i''))$$
$$\Rightarrow (\lim_{i\to\infty} h(t_i', x_i', \nabla\varphi(t_i', x_i')) = \lim_{i\to\infty} h(t_i'', x_i'', \nabla\varphi(t_i'', x_i''))).$$

Condition (E1) is the condition of extendability of the function h. It is a corollary of the continuity of the Hamiltonian.

Let $(t, x) \in [t_0, \vartheta_0] \times \mathbb{R}^n \setminus J$, denote

$$E_1(t,x) \triangleq \{s \in \mathbb{R}^n : \exists\{(t_i, x_i)\} \subset J : \lim_{i\to\infty}(t_i, x_i) = (t,x) \,\&\, \lim_{i\to\infty} \nabla\varphi(t_i, x_i) = s\}.$$

Since φ is locally lipschitzian, the set $E_1(t,x)$ is nonempty and bounded for every $(t,x) \in [t_0, \vartheta_0] \times \mathbb{R}^n \setminus J$.

If $(t,x) \in [t_0, \vartheta_0] \times \mathbb{R}^n \setminus J$ and $s \in E_1(t,x)$, then assumption (E1) yields that the following value is well defined:

$$h(t,x,s) \triangleq \lim_{i\to\infty} h(t_i, x_i, \nabla\varphi(t_i, x_i))$$
$$\forall\{(t_i, x_i)\}_{i=1}^{\infty} \subset J : \lim_{i\to\infty}(t_i, x_i) = (t,x) \,\&\, s = \lim_{i\to\infty} \nabla\varphi(t_i, x_i). \quad (7)$$

The function h is defined on the basis of the Clarke subdifferential of φ at $(t,x) \in (t_0, \vartheta_0) \times \mathbb{R}^n \setminus J$. Indeed, the Clarke subdifferential of φ at $(t,x) \notin J$ is equal to

$$\partial_{\mathrm{Cl}}\varphi(t,x) = \mathrm{co}\{(-h(t,x,s), s) : s \in E_1(t,x)\}. \quad (8)$$

Recall that for any $(t,x) \in (t_0, \vartheta_0) \times \mathbb{R}^n$

$$D_{\mathrm{D}}^{-}\varphi(t,x), D_{\mathrm{D}}^{+}\varphi(t,x) \subset \partial_{\mathrm{Cl}}\varphi(t,x). \quad (9)$$

Denote

$$CJ^{-} \triangleq \{(t,x) \in (t_0, \vartheta_0) \times \mathbb{R}^n \setminus J : D_{\mathrm{D}}^{-}\varphi((t,x)) \neq \varnothing\};$$

$$CJ^{+} \triangleq \{(t,x) \in (t_0, \vartheta_0) \times \mathbb{R}^n \setminus J : D_{\mathrm{D}}^{+}\varphi((t,x)) \neq \varnothing\}.$$

Notice that $CJ^{-} \cap CJ^{+} = \varnothing$.

Define a set $E_2(t,x)$ for $(t,x) \in CJ^{-}$ by the rule:

$$E_2(t,x) \triangleq \{s \in \mathbb{R}^n : \exists a \in \mathbb{R} : (a,s) \in D_{\mathrm{D}}^{-}\varphi((t,x))\} \setminus E_1(t,x).$$

If $(t, x) \in CJ^+$, set

$$E_2(t, x) \triangleq \{s \in \mathbb{R}^n : \exists a \in \mathbb{R} : (a, s) \in D^+_{\mathrm{D}}\varphi((t, x))\} \setminus E_1(t, x).$$

If $(t, x) \in ([t_0, \vartheta_0] \times \mathbb{R}^n) \setminus (CJ^- \cup CJ^+)$, set

$$E_2(t, x) \triangleq \varnothing.$$

The set $E_2(t, x)$ is the complement of $E_1(t, x)$ with respect to the projection of the Dini subdifferential (or superdifferential) at (t, x).

For $(t, x) \in [t_0, \vartheta_0] \times \mathbb{R}^n$ define

$$E(t, x) \triangleq E_1(t, x) \cup E_2(t, x).$$

$E(t, x) \neq \varnothing$ for any $(t, x) \in [t_0, \vartheta_0] \times \mathbb{R}^n$.

Let us introduce the following notations. If $i = 1, 2$, then

$$\mathbb{E}_i \triangleq \{(t, x, s) : (t, x) \in [t_0, \vartheta_0] \times \mathbb{R}^n, \ \ s \in E_1(t, x)\}.$$

Denote

$$\mathbb{E} \triangleq \{(t, x, s) : (t, x) \in [t_0, \vartheta_0] \times \mathbb{R}^n, \ \ s \in E(t, x)\};$$

$$\mathbb{E}^\natural \triangleq \{(t, x, s) : (t, x) \in [t_0, \vartheta_0] \times \mathbb{R}^n, \ \ s \in E^\natural(t, x)\}.$$

Note that $\mathbb{E}^\natural \subset [t_0, \vartheta_0] \times \mathbb{R}^n \times S^{(n-1)}$. Here, $S^{(n-1)}$ means the $(n-1)$-dimensional sphere

$$S^{(n-1)} \triangleq \{s \in \mathbb{R}^n : \|s\| = 1\}.$$

Also, $\mathbb{E} = \mathbb{E}_1 \cup \mathbb{E}_2$.

Note that the function h is defined on $\mathbb{E}_1$. The truth of inclusion $\varphi \in$ VALF depends on the existence of an extension of h to $\mathbb{E}$.

Theorem 1. *The function $\varphi \in \mathrm{Lip}_B$ belongs to the set* VALF *if and only if Condition* (E1) *holds and the function h defined on $\mathbb{E}_1$ by formulas (6) and (7) is extendable to the set $\mathbb{E}$ such that Conditions* (E2)–(E4) *are valid. (Conditions* (E2)–(E4) *are defined below.)*

`Condition (E2).`

- If $(t, x) \in CJ^-$ then for any $s_1, \dots s_{n+2} \in E_1(t, s)$ $\lambda_1, \dots, \lambda_{n+2} \in [0, 1]$ such that $\sum \lambda_k = 1$, $(-\sum \lambda_k h(t, x, s_k), \sum \lambda_k s_k) \in D^-\varphi(t, x)$ the following inequality holds:

$$h\left(t, x, \sum_{k=1}^{n+2} \lambda_k s_k\right) \leq \sum_{k=1}^{n+2} \lambda_k h(t, x, s_k).$$

- If $(t, x) \in CJ^+$ then for any $s_1, \dots s_{n+2} \in E_1(t, s)$ $\lambda_1, \dots, \lambda_{n+2} \in [0, 1]$ such that $\sum \lambda_k = 1$, $(-\sum \lambda_k h(t, x, s_k), \sum \lambda_k s_k) \in D^+\varphi(t, x)$ the following inequality holds:

$$h\left(t, x, \sum_{k=1}^{n+2} \lambda_k s_k\right) \geq \sum_{k=1}^{n+2} \lambda_k h(t, x, s_k).$$

Condition (E2) is an analog of minimax inequalities (4), (5).

`Condition (E3).` For all $(t, x) \in [t_0, \vartheta_0] \times \mathbb{R}^n$:

- If $0 \in E(t, x)$, then $h(t, x, 0) = 0$.
- If $s_1 \in E(t, x)$ and $s_2 \in E(t, x)$ are codirectional (i.e. $\langle s_1, s_2 \rangle = \|s_1\| \cdot \|s_2\|$), then

$$\|s_2\| h(t, x, s_1) = \|s_1\| h(t, x, s_2).$$

This condition means that the function h is positively homogeneous with respect to the third variable.

Let us introduce a function $h^\natural(t, x, s) : \mathbb{E}^\natural \to \mathbb{R}$. Put for $(t, x) \in [t_0, \vartheta_0] \times \mathbb{R}^n$ $\forall s \in E(t, x) \setminus \{0\}$

$$h^\natural(t, x, \|s\|^{-1} s) \triangleq \|s\|^{-1} h(t, x, s).$$

Under Condition (E3), the function $h^\natural$ is well defined.

`Condition (E4).`

- The function $h^\natural$ satisfies the sublinear growth condition: there exists $\Gamma > 0$ such that for any $(t, x, s) \in \mathbb{E}^\natural$

$$h^\natural(t, x, s) \leq \Gamma(1 + \|x\|).$$

- For every bounded region $A \subset \mathbb{R}^n$ there exist $L_A > 0$ and a function $\omega_A \in \Omega$ such that for any $(t', x', s'), (t'', x'', s'') \in \mathbb{E}^\natural \cap [t_0, \vartheta_0] \times A \times \mathbb{R}^n$ the following inequality is fulfilled

$$\begin{aligned} &\|h^\natural(t', x', s') - h^\natural(t'', x'', s'')\| \\ &\quad \leq \omega_A(t' - t'') + L_A \|x' - x''\| + \Gamma(1 + \inf\{\|x'\|, \|x''\|\}) \|s' - s''\|. \end{aligned}$$

Condition (E4) is a restriction of Conditions H1 and H2 to the set $\mathbb{E}$.

The proof of the main theorem is given in Sect. 5. Let us introduce a method of extension of the function h from $\mathbb{E}_1$ to the set $\mathbb{E}$.

Corollary 1. *Let $\varphi \in \mathrm{Lip}_B$. Suppose that h defined on $\mathbb{E}_1$ by formulas (6) and (7) satisfies Condition* (E1)*. Suppose also that the extension of h on $\mathbb{E}_2$ given by the following rule is well defined: $\forall (t, x) \in CJ^- \cup CJ^+$, $s \in E_2(t, x)$*

$$h(t, x, s) \triangleq \sum_{i=1}^{n+2} \lambda_i h(t, x, s_i) \tag{10}$$

for any $s_1, \dots, s_{n+2} \in E_1(t, x)$, $\lambda_1, \dots, \lambda_{n+2}$ *such that* $\sum \lambda_i = 1$ $\sum \lambda_i s_i = s$. *If the function* $h : \mathbb{E} \to \mathbb{R}$ *satisfies Conditions* (E3) *and* (E4)*, then* $\varphi \in$ VALF.

The corollary is proved in Sect. 5.

4 Examples

4.1 Positive Example

Let $n = 2$, $t_0 = 0$, $\vartheta_0 = 1$. Consider the function $\varphi^1(t, x_1, x_2) \triangleq t + |x_1| - |x_2|$. Let us show that $\varphi^1(\cdot, \cdot, \cdot) \in$ VALF. The function φ^1 is differentiable on the set

$$J = \{(t, x_1, x_2) \in (0, 1) \times \mathbb{R}^2 : x_1, x_2 \neq 0\}.$$

If $(t, x_1, x_2) \in J$, then

$$\frac{\partial \varphi^1(t, x_1, x_2)}{\partial t} = 1, \ \nabla \varphi^1(t, x_1, x_2) = (\mathrm{sgn} x_1, -\mathrm{sgn} x_2).$$

Here, $\mathrm{sgn} x$ means the sign of x.

Therefore, if $(\theta, g_1, g_2) \in D^+_{\mathrm{D}} \varphi^1(t, x_1, x_2) \cup D^-_{\mathrm{D}} \varphi^1(t, x_1, x_2)$, then $\theta = 1$.

Let us determine the set $E_1(t, x_1, x_2) \subset \mathbb{R}^n$ and the function $h^1(t, x_1, x_2; s_1, s_2)$ for $(t, x_1, x_2) \in J$ and $(s_1, s_2) \in E_1(t, x_1, x_2)$. The representation of J and the formulas for partial derivatives of φ^1 yield the following representation of $E(t, x_1, x_2)$ and $h^1(t, x_1, x_2)$ for $(t, x_1, x_2) \in J$

$$E_1(t, x_1, x_2) = \{(\mathrm{sgn} x_1, -\mathrm{sgn} x_2)\},$$
$$h^1(t, x_1, x_2; \mathrm{sgn} x_1, -\mathrm{sgn} x_2) = -1.$$

Notice that Condition (E1) for φ^1 is fulfilled. Let $(t, x_1, x_2) \notin J$, then

$$E_1(t, x_1, x_2) = \begin{cases} \{(s_1, -\mathrm{sgn} x_2) : |s_1| = 1\}, \ x_1 = 0, x_2 \neq 0, \\ \{(\mathrm{sgn} x_1, s_2) : |s_2| = 1\}, \ x_1 \neq 0, x_2 = 0, \\ \{(s_1, s_2) : |s_1| = |s_2| = 1\}, \ x_1 = x_2 = 0. \end{cases}$$

If $(t, x_1, x_2) \notin J$, then for $(s_1, s_2) \in E_1(t, x_1, x_2)$ put

$$h^1(t, x_1, x_2; s_1, s_2) = -1.$$

Now let us determine $D^+_{\mathrm{D}} \varphi^1(t, x_1, x_2)$ and $D^-_{\mathrm{D}} \varphi^1(t, x_1, x_2)$ for $(t, x_1, x_2) \notin J$.

Let $x_2 \neq 0$, then

$$D^-_{\mathrm{D}} \varphi^1(t, 0, x_2) = \{(1, s_1, -\mathrm{sgn} x_2) : s_1 \in [-1, 1]\}, \ D^+_{\mathrm{D}} \varphi^1(t, 0, x_2) = \varnothing.$$

Indeed, the function φ^1 has directional derivatives at points $(t, 0, x_2)$ for $x_2 \neq 0$. In addition, the derivative along the vector (τ, g_1, g_2) is

$$d\varphi^1(t, 0, x_2; \tau, g_1, g_2) = \tau + |g_1| - g_2 \mathrm{sgn} x_2.$$

We have,

$$\{(1, s_1, -\mathrm{sgn} x_2) : s_1 \in [-1, 1]\}$$
$$= \{(\theta, s_1, s_2) : (\theta\tau + s_1 g_1 + s_2 g_2) \leq d\varphi^1(t, 0, x_2; \tau, g_1, g_2)\} = D^-_{\mathrm{D}}\varphi^1(t, 0, x_2).$$

Similarly, for $x_1 \neq 0$ we have

$$D^+_{\mathrm{D}}\varphi^1(t, x_1, 0) = \{(1, \mathrm{sgn} x_1, s_2) : s_1 \in [-1, 1]\},\ D^-_{\mathrm{D}}\varphi^1(t, 0, x_2) = \varnothing.$$

Further, $D^+_{\mathrm{D}}\varphi^1(t, 0, 0) = D^-_{\mathrm{D}}\varphi^1(t, 0, 0) = \varnothing$. Naturally, the function φ^1 has directional derivatives at the point $(t, 0, 0)$ and $d\varphi^1(t, x_1, x_2; \tau, g_1, g_2) = \tau + |g_1| - |g_2|$. If $(\theta, s_1, s_2) \in D^+_{\mathrm{D}}\varphi^1(t, 0, 0)$, then $s_1 g_1 \geq |g_1|\ \forall g_1 \in \mathbb{R}$. This yields that $s_1 \geq 1$ and $s_1 \leq -1$. Thus, $D^+_{\mathrm{D}}\varphi^1(t, 0, 0) = \varnothing$. Similarly, $D^-_{\mathrm{D}}\varphi^1(t, 0, 0) = \varnothing$.

Therefore, in this case $CJ^- = \{(t, 0, x_2) \in (0, 1) \times \mathbb{R}^2 : x_2 \neq 0\}$, $CJ^+ = \{(t, x_1, 0) \in (0, 1) \times \mathbb{R}^2 : x_1 \neq 0\}$.

We have

$$E_2(t, x_1, x_2) = \begin{cases} \{(1, s_1, -\mathrm{sgn} x_2) : s_1 \in [-1, 1]\}, & x_1 = 0, x_2 \neq 0, \\ \{(1, \mathrm{sgn} x_1, s_2) : s_1 \in [-1, 1]\}, & x_1 \neq 0, x_2 = 0, \\ \varnothing, & x_1 x_2 \neq 0, \text{ or } x_1 = x_2 = 0. \end{cases}$$

One can use corollary 1 to extend the function h^1 to the set $\mathbb{E}_2$. For $(t, x_1, x_2) \in CJ^- \cup CJ^+$, $s = (s_1, s_2) \in E_2(t, x_1, x_2)$ put $h^1(t, x_1, x_2, s_1, s_2) \triangleq -1$. Since for any $s' = (s'_1, s'_2) \in E_1(t, x_1, x_2)$ $h(t, x_1, x_2, s'_1, s'_2) = -1$, one can suppose that $h^1(t, x_1, x_2, s_1, s_2)$ is determined by (10). Notice that Condition (E3) is fulfilled since for any position (t, x_1, x_2) the set $E(t, x_1, x_2)$ does not contain codirectional vectors as well as vector $(0, 0)$. It is easy to check that Condition (E4) holds.

There exists one more way of extension of the Hamiltonian. Let $\mathscr{Q} \triangleq (\{-1, 1\} \times [-1, 1]) \cup ([-1, 1] \times \{-1, 1\})$. It is easy to check that for any $(t, x_1, x_2) \in [0, 1] \times \mathbb{R}^2$ $\partial_{\mathrm{Cl}}\varphi^1(t, x_1, x_2) \subset \mathscr{Q}$. Put for $(s_1, s_2) \in \mathscr{Q}$ $h(s_1, s_2) \triangleq -1$. The function h is an extension of h^1. The condition of positive homogeneity yields that $h(s_1, s_2) = \min\{-|s_1|, -|s_2|\}$ for $(s_1, s_2) \in \mathbb{R}^2$. One can use the general method of the game reconstruction (see Lemma 2), but here it is possible to guess a control system. Indeed $h(s_1, s_2) = \min\{-|s_1|, -|s_2|\}$ is the Hamiltonian for the control system

$$\begin{cases} \dot{x}_1 = u_0 u_1, \\ \dot{x}_2 = (1 - u_0) u_2, \end{cases}$$

$u_0 \in \{0, 1\}$, $u_1, u_2 \in [-1, 1]$.

4.2 Negative Example

Let $n = 2, t_0 = 0, \vartheta_0 = 1$. Let us show that $\varphi^2(t, x_1, x_2) \triangleq t(|x_1|-|x_2|) \notin$ VALF.

The function $\varphi^2(\cdot,\cdot,\cdot)$ is differentiated on the set $J = \{(t, x_1, x_2) \in (0, 1)\times\mathbb{R}^2 : x_1x_2 \neq 0\}$. We have

$$\frac{\partial\varphi^2(t, x_1, x_2)}{\partial t} = |x_1| - |x_2|,\ \nabla\varphi^2(t, x_1, x_2) = (t \cdot \mathrm{sgn} x_1, -t \cdot \mathrm{sgn} x_2)$$

for $(t, x_1, x_2) \in J$. Thus, for$(t, x_1, x_2) \in J$

$$h^2(t, x_1, x_2, t \cdot \mathrm{sgn} x_1, t \cdot \mathrm{sgn} x_2) = -(|x_1| - |x_2|),$$
$$E_1(t, x_1, x_2) = (t \cdot \mathrm{sgn} x_1, -t \cdot \mathrm{sgn} x_2).$$

Further, if $(t, x_1, x_2) \in J$, $(s_1, s_2) \in E(t, x)$, then $\|(s_1, s_2)\| = t\sqrt{2}$. Thus for $(t, x_1, x_2) \in J$, the following equality is fulfilled

$$E^\natural(t, x_1, x_2) = (\mathrm{sgn} x_1/\sqrt{2}, -\mathrm{sgn} x_2/\sqrt{2}).$$

One can check directly that Condition (E1) holds in this case. Therefore, we may suppose that $h^2(t, x_1, x_2, s_1, s_2)$ is defined on $\mathbb{E}_1$. Here we use formula (7).

Let us introduce the set $\mathbb{E}_0 \subset (0, 1) \times \mathbb{R}^2 \times \mathbb{R}^2$. Put

$$\mathbb{E}_0 \triangleq \{(t, x_1, x_2, t \cdot \mathrm{sgn} x_1, -t \cdot \mathrm{sgn} x_2) : (t, x_1, x_2) \in J\}.$$

By definition of $\mathbb{E}$ we have $\mathbb{E}_0 \subset \mathbb{E}$.

Suppose that there exists extension of the function h^2 satisfying Conditions (E2) and (E3). Hence, the set

$$\mathbb{E}_0^\natural \triangleq \{(t, x_1, x_2, \mathrm{sgn} x_1/\sqrt{2}, -\mathrm{sgn} x_2/\sqrt{2}) : (t, x_1, x_2) \in J\}$$

is a subset of $\mathbb{E}^\natural$. Further, the function $(h^2)^\natural$ is well defined on $\mathbb{E}_0$. In this case,

$$(h^2)^\natural(t, x_1, x_2, \mathrm{sgn} x_1/\sqrt{x_1}, -\mathrm{sgn} x_2/\sqrt{2}) = \frac{|x_1| - |x_2|}{t\sqrt{2}}.$$

Obviously, $(h^2)^\natural$ is unbounded on $(0, 1) \times A \times \mathbb{R}^n \cap \mathbb{E}_0^\natural$. Here, A is any nonempty bounded subset of the set $\{(x_1, x_2) \in \mathbb{R}^n : x_1x_2 \neq 0\})$. Hence, Condition (E4) does not hold for any extension of h^2. Thus $\varphi^2 \notin$ VALF.

5 Construction of the Differential Games Whose Value Coincides with a Given Function

In this section, we prove the statements formulated in Sect. 3.

The proof is based on the following statements.

Lemma 1. *Under Conditions* (E1)–(E4) *the function* $h : \mathbb{E} \to \mathbb{R}$ *can be extended to* $[t_0, \vartheta_0] \times \mathbb{R}^n \times \mathbb{R}^n$ *such that the extension satisfies Conditions* H1–H3.

Lemma 2. *Let a function* $H : [t_0, \vartheta_0] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *satisfy Conditions* H1–H3. *Then there exist sets* $P, Q \in \mathrm{COMP}$ *and a function* $f \in \mathrm{DYN}(P, Q)$ *such that*

$$H(t, x, s) = \max_{v \in Q} \min_{u \in P} \langle s, f(t, x, u, v)\rangle \ \forall (t, x, s) \in [t_0, \vartheta_0] \times \mathbb{R}^n \times \mathbb{R}^n. \tag{11}$$

They are proved in the Appendix.

Proof of the Main Theorem. Necessity. Let $\varphi \in \mathrm{Lip}_B \cap \mathrm{VALF}$. Then by definition of VALF there exist sets $P, Q \in \mathrm{COMP}$, and functions $f \in \mathrm{DYN}(P, Q)$, $\sigma \in \mathrm{TP}$ such that $\varphi = Val^f(\cdot, \cdot, P, Q, f, \sigma)$. Therefore, (see [7]) φ is the minimax solution of the problem (1), (2) with the Hamiltonian defined by (3). Consider the function h defined by formula (6) on J. Note that J means the set of differentiability of φ. We have

$$h(t, x, s) = H(t, x, s),\ (t, x) \in J, s \in E(t, x).$$

Let (t, x) be a position at which the function φ is nondifferentiable, $s \in E_1(t, x)$. Denote

$$\begin{aligned} L\varphi(t, x, s) \triangleq \{a \in \mathbb{R} : \exists \{(t_i, x_i)\}_{i=1}^{\infty} \subset J :& \\ (t, x, s) = \lim_{i\to\infty} (t_i, x_i, \nabla\varphi(t_i, x_i)) \ \&\ a = \lim_{i\to\infty} \partial\varphi(t_i, x_i)/\partial t\}.& \end{aligned}$$

Since $\partial\varphi(t, x)/\partial t = -H(t, x, \nabla\varphi(t, x))$ for $(t, x) \in J$, the continuity H yields that

$$L\varphi(t, x, s) = \{-H(t, x, s)\},\ \ (t, x) \notin J,\ \ s \in E_1(t, x).$$

Thus, the function $h = H$ satisfies Condition (E1). In addition, the function $h(t, x, s) = H(t, x, s)$ is determined by (7) for $(t, x) \notin J$, $s \in E_1(t, x)$. We have that $h = H$ on $\mathbb{E}_1$.

Set the extension of the function h to the set $\mathbb{E}_2$ to be equal to H. Since φ is the minimax solution of the Hamilton–Jacobi equation, we get that for all $(t, x) \in [t_0, \vartheta_0] \times \mathbb{R}^n$ the following inequalities hold

$$a + H(t, x, s) \le 0\ \forall (a, s) \in D^-\varphi(t, x).$$
$$a + H(t, x, s) \ge 0\ \forall (a, s) \in D^+\varphi(t, x).$$

If the function φ is not differentiable at (t,x) and $D^-\varphi(t,x) \cup D^+\varphi(t,x) \neq \varnothing$, then either $(t,x) \in CJ^-$ or $(t,x) \in CJ^+$. Let $(t,x) \in CJ^-$. Consider $\lambda_1, \ldots \lambda_{n+2} \in [0,1]$ and $s_1, \ldots, s_{n+2} \in E_1(t,x)$ such that $\sum \lambda_i = 1$ and

$$\left(-\sum_{k=1}^{n+2} \lambda_k H(t,x,s_k), \sum_{k=1}^{n+2} \lambda_k s_k\right) \in D^-\varphi(t,x).$$

Therefore,

$$-\sum_{k=1}^{n+2} \lambda_k H(t,x,s_k) + H\left(t,x,\sum_{k=1}^{n+2} \lambda_k s_k\right) \leq 0.$$

Similarly, if $(t,x) \in CJ^+$, $\lambda_1, \ldots \lambda_{n+2} \in [0,1]$ and $s_1, \ldots, s_{n+2} \in E_1(t,x)$ satisfy the Conditions $\sum \lambda_i = 1$ and

$$\left(-\sum_{k=1}^{n+2} \lambda_k H(t,x,s_k), \sum_{k=1}^{n+2} \lambda_k s_k\right) \in D^+\varphi(t,x),$$

then the following inequality is fulfilled:

$$-\sum_{k=1}^{n+2} \lambda_k H(t,x,s_k) + H\left(t,x,\sum_{k=1}^{n+2} \lambda_k s_k\right) \geq 0.$$

We get that the function $h = H$ satisfies Condition (E2).

Condition (E3) holds since H is positively homogeneous. Note that $h^\natural(t,x,s) = H(t,x,s)\ \forall (t,x,s) \in \mathbb{E}^\natural$. Since H satisfies Conditions H1 and H2, Condition (E4) is fulfilled also. □

Proof of the Main Theorem. Sufficiency. Consider the function h defined on $\mathbb{E}_1$ by formulas (6) and (7). By the assumption there exists the extension of h to $\mathbb{E}$, which satisfies Conditions (E2)–(E4). By Lemma 1, there exists the function $H : [t_0, \vartheta_0] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ which is extension of h and satisfies Conditions H1–H3. By Lemma 2, there exist compacts $P, Q \in \mathrm{COMP}$ and a function $f \in \mathrm{DYN}(P,Q)$ such that

$$H(t,x,s) = \max_{u \in P} \min_{v \in Q} \langle s, f(t,x,u,v) \rangle. \tag{12}$$

Put $\sigma(x) \triangleq \varphi(\vartheta_0, x)$. Since $\varphi \in \mathrm{Lip}_B$, we get $\sigma \in \mathrm{TP}$. Let us show that $\varphi = Val^f(\cdot,\cdot,P,Q,f,\sigma)$. This is equivalent to the requirement that φ satisfies Conditions (1), (4), and (5).

The boundary condition (1) is valid by the definition of σ. Let us show that φ satisfies Conditions (4) and (5).

If $(t,x) \in J$, then $D^-_{\mathrm{D}}\varphi(t,x) = D^+_{\mathrm{D}}\varphi(t,x) = \{(\partial\varphi(t,x)/\partial t, \nabla\varphi(t,x))\}$

$$\frac{\partial \varphi(t,x)}{\partial t} = -h(t,x,\nabla\varphi(t,x)) = -H(t,x,\nabla\varphi(t,x)).$$

Therefore, for $(t,x) \in J$ inequalities (4) and (5) hold.

Now let $(t, x) \notin J$. By the properties of Clarke subdifferential and the function h (see (8), (9)) it follows that

$$D_{\mathrm{D}}^{-}\varphi(t, x), D_{\mathrm{D}}^{+}\varphi(t, x) \subset \mathrm{co}\{(-h(t, x, s), s) : s \in E_1(t, x)\}. \tag{13}$$

If $(a, s) \in D_{\mathrm{D}}^{-}\varphi(t, x)$ (in this case $(t, x) \in CJ^{-}$), then there exist $\lambda_1, \ldots, \lambda_{n+2} \in [0, 1]$, $s_1, \ldots, s_{n+2} \in E_1(t, x)$ such that $\sum \lambda_k = 1$, $\sum \lambda_k s_k = s$, $-\sum \lambda_k h\ (t, x, s_k) = a$ (see (13)). Using Condition (E2) we obtain

$$h(t, x, s) \leq \sum \lambda_k h(t, x, s_k) = -a.$$

This is equivalent to Condition (4). Similarly, the truth of (5) can be proved. Thus, φ is minimax solution of the problem (2), (1). By [7] and (12), it follows that $\varphi = Val^f(\cdot, \cdot, P, Q, f, \sigma)$. This completes the proof. □

Proof of Corollary 1. Condition (E1) is valid by the assumption. Condition (E2) follows from (13) and Caratheodory theorem. Conditions (E3) and (E4) hold by assumption. Therefore, $\varphi \in$ VALF. □

6 Conclusion

The method of reconstruction of differential games allows us to explore the set of possible values of differential games. The following application seems rather useful. The suggested method can be used for numerical method verification: One can test a computer program using known value and constructed control system.

The objective for the future is to design the methods of reconstruction of concrete classes of differential games. The classes of linear differential games, stationary differential games, and games with simple motions will be considered. These problems are interesting in theory as well as for applications.

Acknowledgements Work is supported by RFBR (grant No 09-01-00436).

Appendix

Proof of Lemma 1. The proof is based on McShane's theorem about extension of range of function [6]. The extension of h is designed by two stages. First, we extend the function $h^{\natural} : \mathbb{E}^{\natural} \to \mathbb{R}$ to $[t_0, \vartheta_0] \times \mathbb{R}^n \times S^{(n-1)}$. Second, we complete a definition by positive homogeneously.

Let us define a function $h^* : [t_0, \vartheta_0] \times \mathbb{R}^n \times S^{(n-1)} \to \mathbb{R}$. The function is designed to be an extension of $h^{\natural}$. To define h^* we design a sequence of sets $\{G_r\}_{r=0}^{\infty}$,

$G_r \subset [t_0, \vartheta_0] \times \mathbb{R}^n \times S^{(n-1)}$, and a sequence of functions $\{h_r\}_{r=0}^{\infty}$, $h_r : G_r \to \mathbb{R}$, possessing the following properties.

(G1) $G_0 = \mathbb{E}^{\natural}$, $h_0 = h^{\natural}$
(G2) $G_{r-1} \subset G_r$ for all $r \in \mathbb{N}$.
(G3) $\bigcup_{r=0}^{\infty} G_r = [t_0, \vartheta_0] \times \mathbb{R}^n \times S^{(n-1)}$;
(G4) for every natural number r the restriction of h_r on G_{r-1} coincides with h_{r-1};
(G5) for any $(t, x, s) \in G_r$ the following inequality is fulfilled:

$$|h_r(t, x, s)| \leq \Gamma(1 + \|x\|),$$

(G6) for every $r \in \mathbb{N}_0$ and every bounded set $A \subset \mathbb{R}^n$ there exist a constant $L_{A,r}$ and a function $\omega_{A,r} \in \Omega$ such that for any $(t', x', s'), (t'', x'', s'') \in G_r \cap [t_0, \vartheta_0] \times A \times S^{(n-1)}$ the following inequality is fulfilled:

$$\begin{aligned} &|h_r(t', x', s') - h_r(t'', x'', s'')| \leq \\ &\quad \leq \omega_{A,r}(t' - t'') + L_{A,r}\|x' - x''\| + \Gamma(1 + \inf\{\|x'\|, \|x''\|\})\|s' - s''\|. \end{aligned} \tag{14}$$

Here, $\mathbb{N}_0 \triangleq \mathbb{N} \cup \{0\}$.

We define the function h^* in the following way: for every $(t, x, s) \in [t_0, \vartheta_0] \times \mathbb{R}^n \times S^{(n)}$ $h^*(t, x, s) = h_l(t, x, s)$. Here l is the least number $k \in \mathbb{N}_0$ such that $(t, x, s) \in G_k$.

Now let us define the sets G_r. If $x \in \mathbb{R}^n$, $j \in \overline{1, n}$, then by x^j denote the j-th coordinate of x. By $\|\cdot\|_*$ denote the following norm of x:

$$\|x\|_* \triangleq \max_{j=\overline{1,n}} |x^j|.$$

If $x \in \mathbb{R}^n$, then

$$\|x\|_* \leq \|x\|. \tag{15}$$

Let $e \in \mathbb{Z}^n$, $a \in [0, \infty)$. Here, $\mathbb{Z}$ means the set of integer numbers. By $\Pi(e, a)$ denote the n-dimensional cube with center at e and length of edge which is equal to a:

$$\Pi(e, a) \triangleq \left\{x \in \mathbb{R}^n : \|e - x\|_* \leq \frac{a}{2}\right\}.$$

If $a \geq 1$, then

$$\mathbb{R}^n = \bigcup_{e \in \mathbb{Z}^n} \Pi(e, a).$$

Order elements $e \in \mathbb{Z}^n$, such that the following implication holds: if $\|e_i\|_* \leq \|e_k\|_*$, then $i \leq k$. Denote

$$M_k(a) \triangleq [t_0, \vartheta_0] \times \Pi(e_k, a) \times S^{(n-1)}.$$

Define the sequence $\{G_r\}_{r=0}^{\infty}$ by the rule:

$$G_0 \triangleq \mathbb{E}^\natural,\ G_k \triangleq G_{k-1} \cup M_k(1)\ \forall k \in \mathbb{N}. \tag{16}$$

We have

$$[t_0, \vartheta_0] \times \mathbb{R}^n \times S^{(n-1)} = \bigcup_{k \in \mathbb{N}_0} G_k.$$

Conditions (G1)–(G3) are fulfilled by definition.

Now we turn to the definition of sequence of functions $\{h_r\}$. Put

$$h_0(t, x, s) \triangleq h^\natural(t, x, s)\ \forall (t, x, s) \in G_0 = \mathbb{E}^\natural.$$

Notice that for $r = 0$ Conditions (G5) and (G6) are fulfilled by (E4).

Now suppose that the function h_{k-1} is defined on G_{k-1} such that Conditions (G5) and (G6) hold with $r = k - 1$. Let us define the function $h_k : G_k \to \mathbb{R}$.

Denote by L_k the constant $L_{A,k-1}$ in Condition (G6) with $A = \Pi(e_k, 3)$. We may assume that

$$L_k \geq \Gamma. \tag{17}$$

By ω_k we denote the function $\omega_{A,k-1}$ with $A = \Pi(e_k, 3)$.

Let $(t, x, s) \in G_k$. For $(t, x, s) \notin M_k(1)$ put $h_k(t, x, s) \triangleq h_{k-1}(t, x, s)$. For $(t, x, s) \in M_k(1)$ put

$$\begin{aligned} h_k(t, x, s) \triangleq \max\{&-\Gamma(1 + \|x\|), \\ &\sup\{h_{k-1}(\tau, y, \xi) - \omega_k(t - \tau) - L_k\|x - y\| \\ &- \Gamma(1 + \|x\|)\|s - \xi\| : (\tau, y, \xi) \in G_{k-1} \cap M_k(3)\}\}. \end{aligned} \tag{18}$$

Let us show that Condition (G4) is fulfilled for $r = k$. This means that $h_k(t, x, s) = h_{k-1}(t, x, s)$ for $(t, x, s) \in G_{k-1} \cap M_k(1)$. We have

$$\begin{aligned} &\sup\{h_{k-1}(\tau, y, \xi) - \omega_k(t - \tau) - L_k\|x - y\| - \Gamma(1 + \|x\|)\|s - \xi\| : \\ &\quad (\tau, y, \xi) \in G_{k-1} \cap M_k(3)\} \geq h_{k-1}(t, x, s) \geq -\Gamma(1 + \|x\|). \end{aligned}$$

Hence,

$$\begin{aligned} h_k(t, x, s) = \sup\{&h_{k-1}(\tau, y, \xi) - \omega_k(t - \tau) - L_k\|x - y\| - \Gamma(1 + \|x\|)\|s - \xi\| : \\ &(\tau, y, \xi) \in G_{k-1} \cap M_k(3)\} \geq h_{k-1}(t, x, s). \end{aligned} \tag{19}$$

Let $\varepsilon > 0$, let $(\tau, y, \xi) \in G_k \cap M_k(3)$ be an element satisfying the inequality

$$h_k(t, x, s) \leq h_{k-1}(\tau, y, \xi) - \omega_k(t - \tau) - L_k\|x - y\| - \Gamma(1 + \|x\|)\|s - \xi\| + \varepsilon. \tag{20}$$

Using (14) with $r = k - 1$ and $A = \Pi(e_k, 3)$, we obtain

$$h_{k-1}(\tau, y, \xi) - h_{k-1}(t, x, s) \leq \omega_k(t - \tau) + L_k\|x - y\| + \Gamma(1 + \inf\{\|x\|, \|y\|\})\|s - \xi\|.$$

This and formula (20) yield the following estimate:

$$h_k(t,x,s) - h_{k-1}(t,x,s) \le \varepsilon.$$

Since ε is arbitrary we obtain that $h_k(t,x,s) \le h_{k-1}(t,x,s)$ for $(t,x,s) \in G_{k-1} \cap M_k(1)$. The opposite inequality is established above (see (19)). Therefore, if $(t,x,s) \in G_{k-1} \cap M_k(1)$, then $h_k(t,x,s) = h_{k-1}(t,x,s)$. Thus, the function h_k is an extension of h_{k-1}.

Moreover, one can prove the following implication: if $(t,x,s) \in G_{k-1} \cap M_k(3)$, then

$$h_{k-1}(t,x,s) = \sup\{h_{k-1}(\tau,y,\xi) - \omega_k(t-\tau) - L_k\|x-y\| - \Gamma(1+\|x\|)\|s-\xi\| : (\tau,y,\xi) \in G_{k-1} \cap M_k(3)\}. \tag{21}$$

Let $(t,x,s) \in G_k \cap M_k(3)$. We say that the sequence $\{(t_i,x_i,s_i)\}_{i=1}^{\infty} \subset G_{k-1} \cap M_k(3)$ realizes the value of $h_k(t,x,s)$, if

$$h_k(t,x,s) = \lim_{i\to\infty}[h_{k-1}(t_i,x_i,s_i) - \omega_k(t-t_i) - L_k\|x-x_i\| - \Gamma(1+\|x\|)\|s-s_i\|].$$

If $h_k(t,x,s) > -\Gamma(1+\|x\|)$, then there exists at least one sequence realizing the value of $h_k(t,x,s)$ (see (18)).

Now we prove that h_k satisfies Condition (G5) for $r = k$. We may consider only triples $(t,x,s) \in M_k(1)$. If $h_k(t,x,s) = -\Gamma(1+\|x\|)$, then the sublinear growth condition holds. Now let $h_k(t,x,s) > -\Gamma(1+\|x\|)$. Let sequence $\{(\tau_i,y_i,\xi_i)\}_{i=1}^{\infty} \subset G_{k-1} \cap M_k(3)$ realize the value of $h_k(t,x,s)$. Using inequality (17), we obtain

$$\begin{aligned} h_{k-1}(\tau_i,y_i,\xi_i) - \omega_k(t-\tau_i) - L_k\|x-y_i\| - \Gamma(1+\|x\|)\|s-\xi_i\| &\le \Gamma(1+\|y_i\|) \\ -L_k\|x-y_i\| \le \Gamma(1+\|x\|) + \Gamma\|x-y_i\| - L_k\|x-y_i\| &\le \Gamma(1+\|x\|). \end{aligned}$$

Consequently (see (17)) Condition (G5) holds for $r = k$.

Let us show that h_k satisfies Condition (G6) for $r = k$. Let A be a bounded subset of $\mathbb{R}^n$, let $(t',x',s'), (t'',x'',s'') \in ([t_0,\vartheta_0] \times A \times S^{(n-1)}) \cap G_k$. We estimate the difference $h_k(t',x',s') - h_k(t'',x'',s'')$.

Let us consider 3 cases.

i. $(t',x',s'),\ (t'',x'',s'') \notin M_k(1)$. Since $h_k(t,x,s) = h_{k-1}(t,x,s)$ for $(t,x,s) \in G_k \setminus M_k(1)$, we have

$$\begin{aligned} &h_k(t',x',s') - h_k(t'',x'',s'') \\ &\quad \le \omega_{A,k-1}(t'-t'') + L_{A,k-1}\|x'-x''\| + \Gamma(1+\inf\{\|x'\|,\|x''\|\})\|s'-s''\|. \end{aligned} \tag{22}$$

ii. $(t', x', s'), (t'', x'', s'') \in M_k(3)$ and at least one triple is in $M_k(1)$. ¿From the definition of h_k, it follows that two subcases are possible.
If $h_k(t', x', s') = -\Gamma(1 + \|x'\|)$, then

$$h_k(t', x', s') - h_k(t'', x'', s'') \leq -\Gamma(1 + \|x'\|) + \Gamma(1 + \|x''\|) \leq \Gamma\|x'' - x'\|.$$

Now let $h_k(t', x', s') > -\Gamma(1 + \|x'\|)$. Let the sequence $\{(t_i, x_i, s_i)\}_{i=1}^{\infty} \subset G_{k-1} \cap M_k(3)$ realize the value of $h_k(t', x', s')$. By (21) for $(t, x, s) = (t', x', s')$ and inequality $\|s'' - s'\|, \|s' - s_i\| \leq 2$ we have

$$\begin{aligned}
&h_{k-1}(t_i, x_i, s_i) - \omega_k(t' - t_i) - L_k\|x' - x_i\| - \Gamma(1 + \|x'\|)\|s' - s_i\| \\
&\qquad - h_k(t'', x'', s'') \\
&\quad \leq h_{k-1}(t_i, x_i, s_i) - \omega_k(t' - t_i) - L_k\|x' - x_i\| - \Gamma(1 + \|x'\|)\|s' - s_i\| \\
&\qquad - h_{k-1}(t_i, x_i, s_i) + \omega_k(t'' - t_i) + L_k\|x'' - x_i\| + \Gamma(1 + \|x''\|)\|s'' - s_i\| \\
&\quad \leq \omega_k(t' - t'') + L_k\|x' - x''\| + \Gamma(1 + \|x''\|)(\|s'' - s_i\| - \|s' - s_i\|) \\
&\qquad + \Gamma(\|x''\| - \|x'\|)\|s' - s_i\| \\
&\quad \leq \omega_k(t' - t'') + L_k\|x' - x''\| + \Gamma(1 + \|x''\|)\|s' - s''\| + 2\Gamma\|x' - x''\| \\
&\quad \leq \omega_k(t' - t'') + (L_k + 4\Gamma)\|x' - x''\| + \Gamma(1 + \inf\{\|x'\|, \|x''\|\})\|s' - s''\|.
\end{aligned}$$

Hence,

$$\begin{aligned}
&h_k(t', x', s') - h_k(t'', x'', s'') \\
&\quad \leq \omega_k(t' - t'') + (L_k + 4\Gamma)\|x' - x''\| + \Gamma(1 + \inf\{\|x'\|, \|x''\|\})\|s' - s''\|.
\end{aligned}$$

iii. One of triples (t', x', s'), (t'', x'', s'') belongs to $M_k(1)$, and another triple does not belong to $M_k(3)$. Therefore, $\|x' - x''\| \geq \|x' - x''\|_* > 1$ (see (15)). Since Condition (G5) for $r = k$ is established above, we have

$$h(t', x', s') - h(t'', x'', s'') \leq 2\Gamma(1 + \sup_{y \in A}\|y\|) \leq 2\Gamma(1 + \sup_{y \in A}\|y\|)\|x' - x''\|. \tag{23}$$

Estimates (22)–(23) yield that if $(t', x', s'), (t'', x'', s'') \in G_k \cap ([t_0, \vartheta_0] \times A \times S^{(n-1)})$, then

$$\begin{aligned}
&h_k(t', x', s') - h_k(t'', x'', s'') \\
&\quad \leq \omega_{A,k}(t' - t'') + L_{A,k}\|x' - x''\| + \Gamma(1 + \inf\{\|x'\|, \|x''\|\})\|s' - s''\|.
\end{aligned}$$

Here, $\omega_{A,k}$ is defined by the rule

$$\omega_{A,k}(\delta) \triangleq \max\{\omega_{A,k-1}(\delta), \omega_k(\delta)\}$$

(one can check directly that $\omega_{A,k} \in \Omega$); the constant $L_{A,k}$ is defined by the rule $L_{A,k} \triangleq \max\left\{L_{A,k-1}, L_k + 4\Gamma, \Gamma(1 + \sup_{y \in A}\|y\|)\right\}$.

Therefore, Condition (G6) is fulfilled for $r = k$.

This completes the designing of the sequences $\{G_r\}_{r=0}^{\infty}$ and $\{h_r\}_{r=0}^{\infty}$ satisfying Conditions (G1)–(G6).

For every $(t, x, s) \in [t_0, \vartheta_0] \times \mathbb{R}^n \times S^{(n-1)}$, there exists a number $k \in \mathbb{N}_0$ such that $(t, x, s) \in G_k$. Put

$$h^*(t, x, s) \triangleq h_k(t, x, s).$$

The value of $h^*(t, x, s)$ does not depend on number k satisfying the property $(t, x, s) \in G_k$. By definition of h_k (see (G5)), we have

$$h^*(t, x, s) \leq \Gamma(1 + \|x\|).$$

Let us prove that for every bounded set $A \subset \mathbb{R}^n$ there exist a function $\omega_A \in \Omega$ and a constant L_A such that for all $(t', x', s'), (t'', x'', s'') \in [t_0, \vartheta_0] \times A \times S^{(n-1)}$ the following estimate is fulfilled

$$\begin{aligned}
&|h^*(t', x', s') - h^*(t'', x'', s'')| \\
&\quad \leq \omega_A(t' - t'') + L_A\|x' - x''\| + \Gamma(1 + \inf\{\|x'\|, \|x''\|\})\|s' - s''\|.
\end{aligned} \tag{24}$$

Indeed, there exists a number m such that

$$A \subset \bigcup_{k=1}^{m} \Pi(e_k, 1).$$

By definition of $\{G_k\}$ (see (16)), we have

$$[t_0, \vartheta_0] \times A \times S^{(n-1)} \subset [t_0, \vartheta_0] \times \left[\bigcup_{k=1}^{m} \Pi(e_k, 1)\right] \times S^{(n-1)} \subset G_m.$$

Put $\omega_A \triangleq \omega_{A,m}$, $L_A \triangleq L_{A,m}$. Since $h^*(t, x, s) = h_m(t, x, s)$ $\forall (t, x, s) \in [t_0, \vartheta_0] \times A \times S^{(n-1)}$, property (G6) for $r = m$ yields that

$$\begin{aligned}
&|h^*(t', x', s') - h^*(t'', x'', s'')| = |h_m(t', x', s') - h_m(t'', x'', s'')| \leq \\
&\quad \leq \omega_{A,m}(t' - t'') + L_{A,m}\|x' - x''\| + \Gamma(1 + \inf\{\|x'\|, \|x''\|\})\|s' - s''\|.
\end{aligned}$$

Thus, inequality (24) is fulfilled.

Now let us introduce a function $H : [t_0, \vartheta_0] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Put

$$H(t, x, s) \triangleq \begin{cases} \|s\| h^*(t, x, \|s\|^{-1}s), & s \neq 0 \\ 0, & s = 0. \end{cases} \tag{25}$$

The function H is an extension of h. Naturally, let $(t, x, s) \in \mathbb{E}$. If $s \neq 0$, then $(t, x, \|s\|^{-1}s) \in \mathbb{E}^{\natural}$. Hence,

$$H(t, x, s) = \|s\| h^*(t, x, \|s\|^{-1}s) = \|s\| h^{\natural}(t, x, \|s\|^{-1}\|s\|) = h(t, x, s).$$

If $s = 0$, then by Condition (E3) we have

$$h(t, x, 0) = 0 = H(t, x, 0).$$

The function H satisfies Condition H2. Let $s_1, s_2 \in \mathbb{R}^n$, $(t, x) \in [t_0, \vartheta_0] \times \mathbb{R}^n$. Let us estimate $|H(t, x, s_1) - H(t, x, s_2)|$. Without loss of generality, it can be assumed that $\|s_1\| \geq \|s_2\|$. If $\|s_2\| = 0$, then

$$\begin{aligned} |H(t, x, s_1) - H(t, x, s_2)| &= |H(t, x, s_1)| \\ &\leq \Gamma(1 + \|x\|)\|s_1\| = \Gamma(1 + \|x\|)\|s_1 - s_2\|. \end{aligned} \tag{26}$$

Now let $\|s_2\| > 0$.

$$\begin{aligned} &|H(t, x, s_1) - H(t, x, s_2)| = \left| \|s_1\| h^* \left(t, x, \frac{s_1}{\|s_1\|}\right) - \|s_2\| h^* \left(t, x, \frac{s_2}{\|s_2\|}\right) \right| \\ &\quad \leq (\|s_1 - s_2\|) \left| h^* \left(t, x, \frac{s_1}{\|s_1\|}\right) \right| + \|s_2\| \left| h^* \left(t, x, \frac{s_1}{\|s_1\|}\right) - h^* \left(t, x, \frac{s_2}{\|s_2\|}\right) \right| \\ &\quad \leq \Gamma(1 + \|x\|)\|s_1 - s_2\| + \|s_2\| \Gamma(1 + \|x\|) \left\| \frac{s_1}{\|s_1\|} - \frac{s_2}{\|s_2\|} \right\| \leq \\ &\quad \leq 2\Gamma(1 + \|x\|)\|s_1 - s_2\|. \end{aligned} \tag{27}$$

To prove the last estimate in (27) we need to show that if $\|s_1\| \geq \|s_2\|$ then

$$\left\| \frac{\|s_2\| s_1}{\|s_1\|} - s_2 \right\| \leq \|s_1 - s_2\|. \tag{28}$$

Let $z \in \mathbb{R}^n$ be codirectional with s_1, let γ be the angle between s_1 and s_2:

$$\cos\gamma = \frac{\langle s_1, s_2 \rangle}{\|s_1\| \cdot \|s_2\|}.$$

Consider a triangle formed by the origin and the terminuses of z and s_2. The lengths of the sides of the triangle are $\|z\|$, $\|s_2\|$ and $\|z - s_2\|$. By the cosine theorem, we have

$$\begin{aligned} \|z - s_2\|^2 &= \|s_2\|^2 + \|z\|^2 - 2\|z\|\|s_2\| \cos\gamma \\ &= \|s_2\|^2(1 - \cos^2\gamma) + (\|z\| - \|s_2\| \cos\gamma)^2. \end{aligned}$$

Hence, the function $\|z - s_2\|$ as a function of $\|z\|$ increases on the region $\|z\| \geq \|s_2\| \cos\gamma$. Since

$$\left\| \frac{\|s_2\| s_1}{\|s_1\|} \right\| = \|s_2\| \leq \|s_1\|,$$

the estimate (28) holds.

Combining estimates (26) and (27), we get

$$|H(t,x,s_1) - H(t,x,s_2)| \le \Upsilon(1+\|x\|)\|s_1 - s_2\|$$
$$\forall (t,x) \in [t_0,\vartheta_0] \times \mathbb{R}^n \ \forall s_1, s_2 \in \mathbb{R}^n.$$

Here $\Upsilon = 2\Gamma$. Using the definition of H (see (25)) and the properties of the function h^* (see (24)), we obtain that the function H satisfies Condition H2.

Notice that for all $(t,x,s) \in [t_0,\vartheta_0] \times \mathbb{R}^n \times \mathbb{R}^n$ the following inequality holds:

$$|H(t,x,s)| \le \Gamma\|s\|(1+\|x\|) \le \Upsilon\|s\|(1+\|x\|).$$

This means that the function H satisfies Condition H1.

The function H is positively homogeneous by definition. □

Proof of Lemma 2. Lemma 2 is close to the result of L.C. Evans and P.E. Souganidis (see [4]) about construction of differential games.

Denote $B \triangleq \{s \in \mathbb{R}^n : \|s\| \le 1\}$. By Condition H2, there exists a real number Υ, such that for all $(t,x,s_1), (t,x,s_2) \in [t_0,\vartheta_0] \times \mathbb{R}^n \times \mathbb{R}^n$ the following estimate holds:

$$|H(t,x,s_1) - H(t,x,s_2)| \le \Upsilon(1+\|x\|)\|s_1 - s_2\|.$$

Therefore,

$$\begin{aligned} H(t,x,s) &= \|s\| H\left(t,x,\frac{s}{\|s\|}\right) \\ &\quad = \|s\| \max_{z\in B}\left[H(t,x,z) - \Upsilon(1+\|x\|)\left\|\frac{s}{\|s\|} - z\right\|\right] = \\ &= \|s\| \max_{z\in B}\min_{y\in B}\left[H(t,x,z) + \Upsilon(1+\|x\|)\left\langle y, \frac{s}{\|s\|} - z\right\rangle\right] \\ &= \|s\| \max_{z\in B}\min_{y\in B}\Big[(H(t,x,z) + \Upsilon(1+\|x\|)) - \Upsilon(1+\|x\|) \\ &\quad + \Upsilon(1+\|x\|)\Big\langle y, \frac{s}{\|s\|} - z\Big\rangle\Big] = \max_{z\in B}\min_{y\in B}[(H(t,x,z) + \Upsilon(1+\|x\|)\|s\| \\ &\quad + \Upsilon(1+\|x\|)\langle y,s\rangle - \Upsilon(1+\|x\|)(1+\langle y,z\rangle)\|s\|] \end{aligned}$$

Since for all $y, z \in B$

$$H(t,x,z) + \Upsilon(1+\|x\|),\ \Upsilon(1+\|x\|)(1+\langle y,z\rangle) \ge 0,$$

it follows that

$$\begin{aligned} H(t,x,s) = &\max_{z\in B}\min_{y\in B}\max_{z'\in B}\min_{y'\in B}[(H(t,x,z) + \Upsilon(1+\|x\|))\langle z', s\rangle \\ &+ \Upsilon(1+\|x\|)\langle y,s\rangle + \Upsilon(1+\|x\|)(1+\langle y,z\rangle)\langle y', s\rangle]. \end{aligned} \tag{29}$$

In formula (29) one can interchange $\min_{y\in B}$ and $\max_{z'\in B}$. Denoting $P = Q = B \times B$, and

$$f(t, x, u, v) \triangleq H(t, x, z)z' + \Upsilon(1 + \|x\|)[z' + y + (1 + \langle y, z\rangle)y'],$$

for $(t, x) \in [t_0, \vartheta_0] \times \mathbb{R}^n$, $u = (y, y')$, $v = (z, z')$ we obtain that (11) is fulfilled. By definition of f, it follows that $f \in \mathrm{DYN}(P, Q)$. □

References

1. Bardi, M., Capuzzo-Dolcetta, I.: Optimal control and viscosity solutions of Hamilton–Jacobi-Bellman equations. With appendices by Maurizio Falcone and Pierpaolo Soravia. Birkhauser Boston, Boston (1997)
2. Demyanov, V.F., Rubinov, A.M.: Foundations of Nonsmooth Analysis, and Quasidifferential Calculus. In Optimization and Operation Research **23**, Nauka, Moscow (1990)
3. Evans, L.C., Gariepy, R.F.: Measure Theory and Fine Properties of Functions. CRC, Boca Raton (1992)
4. Evans, L.C., Souganidis, P.E.: Differential games and representation formulas for solutions of Hamilton–Jacobi-Isaacs Equations. Indiana University Mathematical Journal **33**, 773–797 (1984)
5. Krasovskii, N.N., Subbotin, A.I.: Game-Theoretical Control Problems. Springer, New York (1988)
6. McShane, E.J.: Extension of range of function. Bull. Amer. Math. Soc. **40**, 837–842, (1934)
7. Subbotin, A.I.: Generalized solutions of first-order PDEs. The dynamical perspective. Birkhauser, Boston (1995)

Existence and Uniqueness of Disturbed Open-Loop Nash Equilibria for Affine-Quadratic Differential Games

Teresa-Paula Azevedo-Perdicoúlis and Gerhard Jank

Abstract In this note, we investigate the solution of a disturbed quadratic open loop (OL) Nash game, whose underlying system is an affine differential equation and with a finite time horizon. We derive necessary and sufficient conditions for the existence/uniqueness of the Nash/worst-case equilibrium. The solution is obtained either via solving initial/terminal value problems (IVP/TVP, respectively) in terms of Riccati differential equations or solving an associated boundary value problem (BVP). The motivation for studying the case of the affine dynamics comes from practical applications, namely the optimization of gas networks. As an illustration, we applied the results obtained to a scalar problem and compare the numerical effectiveness between the proposed approach and an usual Scilab BVP solver.

1 Introduction

The problem of solving games with quadratic performance criterion and a linear differential equation defining the constraints has been addressed by many authors (see e.g. [4] and [5] and references therein), since this is often used as a benchmark to evaluate game outcomes/strategies. In [8], a disturbed linear quadratic differential game, where each player chooses his strategy according to a modified Nash equilibrium model under OL information structure, has been analyzed. Conditions for the existence and uniqueness of such an equilibrium were given, and it was also shown how these conditions are related to certain Riccati differential equations.

T.-P. Azevedo-Perdicoúlis (✉)
ISR-pólo de Coimbra & Departamento Matemática, UTAD, 5000-911 Vila Real, Portugal
e-mail: tazevedo@utad.pt

G. Jank
Universidade de Aveiro, 3810-193 Aveiro, Portugal
e-mail: jank@math2.rwth-aachen.de

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_2,

However, some applied problems exist where the underlying system of the game is affine, for instance, in the modeling of the optimization of gas networks as a dynamic game, where the unknown offtakes from the gas net are modeled as a disturbance (see for instance [2]). In [3], the payoffs obtained by the players of a linear disturbed game and the equivalent undisturbed setting have been compared.

In this paper, we consider a quadratic performance criterion for every player, a finite planning horizon, an affine differential equation as constraint, and following [8] use an operator approach.

In [6], affine LQ Nash games equilibria on infinite planning horizon are studied and conditions for existence and uniqueness of an equilibrium are obtained. Moreover, in [5] and [7] it has been remarked that a particular transformation applied to make the game a standard linear quadratic one, which then can be solved by standard methods, cannot be applied since it may corrupt the OL information structure. In our case, also exists a simple transformation of the form

$$x(t) = y(t) + Z(t),$$

where $y(t)$ fulfills a linear differential equation while $x(t)$ solves the IVP (1) and $Z(t)$ is an adequately chosen function depending on the affine term $F(t)$. But, as the aforementioned observation clearly also applies to the finite time horizon situation, independently of a disturbance being present (or not), the OL information structure may be corrupted as well. Therefore, it is the aim of this paper to extend directly the work of Jank and Kun in [8] to the affine case, as well as the affine LQ game stated in [4] to the disturbed situation.

In this note, we generalize the procedure to calculate the equilibrium controls proposed in [8], where the controls are obtained in terms of the solution of certain Riccati differential equations, to the case when the underlying system is affine. The outcome is an interesting procedure from the applications point of view, since only IVPs have to be solved instead of a BVP. We ought to mention that usual routines for solving BVPs cannot be applied to the vectorial case.

The outline is as follows. Section 2 states the problem and introduces the notation. In Sect. 3, sufficient conditions for the existence/uniqueness of an OL equilibrium are investigated. Moreover, more explicit representations for such solutions, in terms of certain Riccati differential equations, are obtained. In Sect. 4, we provide a scalar example, where the equilibrium controls are calculated in two different manners: (a) using the transformation stated in Theorem 4, and (b) solving directly the BVP of Theorem 3. In the Appendix, the approach stated in Theorem 4 is fully described as an algorithm.

2 Preliminaries

We discuss disturbed noncooperative affine differential games with quadratic costs, defined on a finite time horizon $\left[t_0, t_f\right] \subset \mathbb{R}$. This means that the underlying system is affine, with each state being controlled by N players and a disturbance term:

$$
\begin{aligned}
\dot{x} &= A(t)x(t) + \sum_{i=1}^{N} B_i(t)u_i(t) + C(t)w(t) + F(t), \\
x(0) &= x_0 \in \mathbb{R}^n,
\end{aligned} \tag{1}
$$

with $x(t) \in \mathbb{R}^n$, piecewise continuous and bounded functions, $A(t) \in \mathbb{R}^{n\times n}$, $B_i(t) \in \mathbb{R}^{n\times m_i}$, $C(t) \in \mathbb{R}^{n\times m}$, and $F(t) \in \mathbb{R}^n$. Furthermore, $u_i(\cdot) \in \mathscr{U}_i$ denotes the control of the i^{th} player, and $w \in \mathscr{W}$ the disturbance. $\mathscr{U}_i$ and $\mathscr{W}$ denote, respectively, the Hilbert spaces $\mathscr{L}_2^{m_i}\left[t_0, t_f\right]$ and $\mathscr{L}_2^{m}\left[t_0, t_f\right]$. We also set $\mathscr{U} = \prod_{i=1}^{N} \mathscr{U}_i$.

For $i = 1, \ldots, N$, the cost functional has the form:

$$
\begin{aligned}
J_i\left(u_1(\cdot), \ldots, u_N(\cdot), w\right) &= x^T(t_f) K_{if} x(t_f) \\
&+ \int_{t_0}^{t_f} \left(x^T(t) Q_i(t) x(t) + \sum_{j=1}^{N} u_j^T(t) R_{ij} u_j(t) + w^T(t) P_i(t) w(t) \right) \mathrm{d}t
\end{aligned} \tag{2}
$$

with symmetric matrices $K_{if} \in \mathbb{R}^{n\times n}$, and symmetric, piecewise continuous and bounded matrix functions $Q_i(t) \in \mathbb{R}^{n\times n}$, $R_{ij}(t) \in \mathbb{R}^{m_i\times m_j}$, and $P_i(t) \in \mathbb{R}^{m\times m}$, $i, j = 1, \ldots, N$.

We recall next the definition of Nash/worst-case strategy:

Definition 1. We define the Nash/worst-case equilibrium in two stages. Consider $u = (u_1, \ldots, u_N) \in \mathscr{U}$, then:

1. $\hat{w}_i(u) \in \mathscr{W}$ is the *worst-case disturbance from the point of view of the i^{th} player* according to these controls if

$$
J_i\left(u, \hat{w}_i(u)\right) \geq J_i\left(u, w\right)
$$

   holds for each $w \in \mathscr{W}, i \in \{1, \ldots, N\}$.
2. The controls $(\tilde{u}_1, \ldots, \tilde{u}_N) \in \mathscr{U}$ form a *Nash/worst-case equilibrium* if for all $i = 1, \ldots, N$

   (a) There exists a worst-case disturbance from the point of view of the i^{th} player according to all controls $(u_1, \ldots, u_N) \in \mathscr{U}$ and

   (b)
$$
J_i\left(\tilde{u}_i, \bar{\tilde{u}}_i, \hat{w}_i(\tilde{u}_i, \bar{\tilde{u}}_i)\right) \leq J_i\left(u_i, \bar{\tilde{u}}_i, \hat{w}_i(u_i, \bar{\tilde{u}}_i)\right),
$$

   where $\bar{u}_i = (u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_N)$, holds for each worst-case disturbance $\hat{w}_i(u) \in \mathscr{W}$ and admissible control function $u_i \in \mathscr{U}_i$.

We now define some necessary notation.

Let $\mathscr{H}^n_{t_f}$ be the Hilbert space of the set of square integrable $\mathbb{R}^n$–valued functions on $\left[t_0, t_f\right] \subset \mathbb{R}$, with the scalar product:

$$\langle f, g\rangle_{\mathscr{H}^n_{t_f}} = \langle f, g\rangle = f^T(t_f)g(t_f) + \int_{t_0}^{t_f} f^T(t)g(t)\mathrm{d}t, \qquad f, g \in \mathscr{H}^n_{t_f}. \quad (3)$$

Denote by $\Phi(t,\tau)$ the solution of the IVP:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi\left(t,\tau\right) = A(t)\Phi\left(t,\tau\right), \qquad \Phi\left(t,t\right) = I_n, \quad (4)$$

where $I_n \in \mathbb{R}^{n\times n}$ is the identity matrix.

For $i, j = 1, \ldots, N$, we define the following linear operators:

$$\begin{aligned}
\Phi : \mathbb{R}^n &\longrightarrow \mathscr{H}^n_{t_f}\\
x_0 &\rightsquigarrow \Phi\left(\cdot, t_0\right) x_0,\\[1ex]
\mathscr{B}_i : \mathscr{U}_i &\longrightarrow \mathscr{H}^n_{t_f}\\
u_i &\rightsquigarrow \int_{t_0}^{\cdot} \Phi\left(\cdot,\tau\right) B_i\left(\tau\right) u_i\left(\tau\right)\mathrm{d}\tau,\\[1ex]
\mathscr{C} : \mathscr{W} &\longrightarrow \mathscr{H}^n_{t_f}\\
w &\rightsquigarrow \int_{t_0}^{\cdot} \Phi\left(\cdot,\tau\right) C\left(\tau\right) w\left(\tau\right)\mathrm{d}\tau,\\[1ex]
\mathscr{F} : \mathscr{L}_2\left[t_0, t_f\right] &\longrightarrow \mathscr{H}^n_{t_f}\\
F &\rightsquigarrow \int_{t_0}^{\cdot} \Phi\left(\cdot,\tau\right) F\left(\tau\right)\mathrm{d}\tau.
\end{aligned}$$

And we define finally the operators:

$$\begin{aligned}
\bar{Q}_i : \mathscr{H}^n_{t_f} &\longrightarrow \mathscr{H}^n_{t_f}\\
x(\cdot) &\rightsquigarrow \left(t \longmapsto \begin{cases} Q_i(t)x(t), & t \neq t_f\\ K_{if}x(t), & t = t_f\end{cases}\right)\\
\bar{R}_{ij} : \mathscr{H}^{m_i}_{t_f} &\longrightarrow \mathscr{H}^{m_j}_{t_f}\\
x(\cdot) &\rightsquigarrow \left(t \longmapsto \begin{cases} R_{ij}(t)x(t), & t \neq t_f\\ 0, & t = t_f\end{cases}\right)\\
\bar{P}_i : \mathscr{H}^{m}_{t_f} &\longrightarrow \mathscr{H}^{m}_{t_f}\\
x(\cdot) &\rightsquigarrow \left(t \longmapsto \begin{cases} P_i(t)x(t), & t \neq t_f\\ 0, & t = t_f\end{cases}\right).
\end{aligned}$$

Using the above definitions, we write the solution of the differential equation (1) on $\mathscr{H}^n_{t_f}$ as:

$$x(\cdot) = \Phi x_0 + \sum_{j=1}^{N} \mathscr{B}_i u_i + \mathscr{C} w + \mathscr{F}. \quad (5)$$

Similarly, the cost functionals (2) are rewritten in terms of scalar products:

$$J_i\left(u_1(\cdot),\ldots,u_N(\cdot),w\right) = \left\langle x, \bar{Q}_i x\right\rangle_{\mathcal{H}_{t_f}^n} + \sum_{j=1}^{N}\left\langle u_j, \bar{R}_{ij} u_j\right\rangle_{\mathcal{H}_{t_f}^{m_j}} + \left\langle w, \bar{P}_i w\right\rangle_{\mathcal{H}_{t_f}^m}, i = 1,\ldots,N. \tag{6}$$

3 Sufficient Existence Conditions for OL Equilibrium Controls

In this section, we state the modified theorems for the affine linear-quadratic case, generalizing the work presented in [8].

Theorem 1. *For $i = 1,\ldots,N$, let us define the operators*

$$\begin{aligned} &F_i : \mathcal{U}_i \longrightarrow \mathcal{U}_i, && G_i : \mathcal{W} \longrightarrow \mathcal{W}, && H_i : \mathcal{W} \longrightarrow \mathcal{U}_i, \\ &F_i := \mathcal{B}_i^* \bar{Q}_i \mathcal{B}_i + \bar{R}_{ii}, && G_i := \mathcal{C}^* \bar{Q}_i \mathcal{C} + \bar{P}_i, && H_i := \mathcal{B}_i^* \bar{Q}_i \mathcal{C}, \end{aligned}$$

where $-^*$ denotes the adjoint of an operator.

1. *There exists a unique worst-case disturbance $\hat{w}_i \in \mathcal{W}$ from the point of view of the i^{th} player if and only if $G_i < 0$. This disturbance, is then given by:*

$$\hat{w}_i\left(u_1,\ldots,u_N\right) = -G_i^{-1}\left(H_i^* u_i + \mathcal{C}^* \bar{Q}_i\left(\Phi x_0 + \sum_{j\neq i}\mathcal{B}_j u_j + \mathcal{F}\right)\right) \tag{7}$$

   for all $(u_1,\ldots,u_N) \in \mathcal{U}$.
2. *Moreover, for $i = 1,\ldots,N$, let $G_i < 0$ and $F_i > 0$. Then $(\tilde{u}_1,\ldots,\tilde{u}_N) \in \mathcal{U}$ form an OL Nash/worst case equilibrium if and only if for each $i = 1,\ldots,N$*

$$\tilde{u}_i = \left(F_i - H_i G_i^{-1} H_i^*\right)^{-1}\left(H_i G_i^{-1}\mathcal{C}^* - \mathcal{B}_i^*\right)\bar{Q}_i\left(\Phi x_0 + \sum_{j\neq i}\mathcal{B}_j u_j + \mathcal{F}\right). \tag{8}$$

Proof. The proof follows just as for the linear case, see [8]. □

Remark 1. From the proof, one can see that only the positive definiteness of $\left(F_i - H_i G_i^{-1} H_i^*\right)$ is required to ensure a unique best reply. This seems to be a weaker condition than for the undisturbed case, where $F_i > 0$ is required. Although in this article we use the assumptions in Theorem 1, this problem needs to be further investigated.

The next theorem describes the Nash/worst-case controls in a feedback form for a "virtual" worst-case state trajectory for the i^{th} player. Please note at this point that every player may have a different "virtual" worst case trajectory, which in general is not the state trajectory.

Theorem 2. *Suppose that matrices P_i and R_{ii}, $i = 1, \ldots, N$, are negative and positive definite, respectively. Suppose also that the operators G_i and F_i, $i = 1, \ldots, N$, are negative and positive definite, respectively. Thence, $\tilde{u}_1, \ldots, \tilde{u}_N$ form an OL Nash/worst-case equilibrium if and only if the following equations are fulfilled:*

$$\tilde{u}_i = -R_{ii}^{-1} \mathscr{B}_i^* \bar{Q}_i \hat{x}_i \tag{9}$$

$$\hat{w}_i = -P_i^{-1} \mathscr{C}^* \bar{Q}_i \hat{x}_i, \tag{10}$$

where

$$\hat{x}_i = \Phi x_0 + \sum_{j \neq i} \mathscr{B}_j \tilde{u}_j + \mathscr{C} \hat{w}_i + \mathscr{F}. \tag{11}$$

Proof. The proof follows as for the linear case, see [8]. □

In this theorem, the controls are described in terms of adjoint operators. For convenience of the reader in the following lemma, we present the general construction of such adjoint operators for our particular function spaces.

Lemma 1. *Let $L : [t_0, t_f] \longrightarrow \mathbb{R}^{n \times k}$ be a piecewise continuous and bounded mapping for some $k, n \in \mathbb{N}$. Supposing that $\mathscr{L}$ denotes the linear operator*

$$\begin{aligned} \mathscr{L} : \mathscr{L}_2^k [t_0, t_f] &\longrightarrow \mathscr{H}_{t_f}^n \\ u \qquad &\rightsquigarrow \int_{t_0}^{\cdot} \Phi(\cdot, \tau) L(\tau) u(\tau) \mathrm{d}\tau, \end{aligned} \tag{12}$$

the adjoint operator $\mathscr{L}^* : \mathscr{H}_{t_f}^n \longrightarrow \mathscr{L}_2^k [t_0, t_f]$ is given by

$$\mathscr{L}^* y := L^T(\cdot) \left[\Phi^T (t_f, \cdot)\, y(t_f) + \int_{t_0}^{\cdot} \Phi^T(t, \cdot) y(t) \mathrm{d}t \right]. \tag{13}$$

Proof. Obvious from the definition of adjoint operator (see for instance [8]). □

Using this lemma, we can express the equilibrium controls of Theorem 2, as well as the worst-case disturbance, as:

$$\tilde{u}_i = -R_{ii}^{-1}(t) B_i^T(t) \left(\Phi^T(t_f, t) K_{if} \hat{x}_i(t_f) + \int_{t_0}^{t_f} \Phi^T(\tau, t) Q_i(\tau) \hat{x}_i(\tau) \mathrm{d}\tau \right) \tag{14}$$

$$\hat{w}_i = -P_i^{-1}(t) C^T(t) \left(\Phi^T(t_f, t) K_{if} \hat{x}_i(t_f) + \int_{t_0}^{t_f} \Phi^T(\tau, t) Q_i(\tau) \hat{x}_i(\tau) \mathrm{d}\tau \right). \tag{15}$$

Theorem 3. *Consider the assumptions on the matrices R_{ii} and P_i, as well as on the operators F_i, G_i, of Theorem 2 fulfilled. Assume also that the solution of set of TVPs:*

$$\dot{E}_i(t) = -A^T(t)E_i(t) - E_i(t)A(t) - Q_i(t) + E_i(t)\left(S_i(t) + T_i(t)\right)E_i(t)$$
$$E_i\left(t_f\right) = K_{if}, \tag{16}$$

where $E_i(t) \in \mathbb{R}^{n\times n}, i = 1,\ldots,N$, is a symmetric matrix, and $S_i := B_i R_{ii}^{-1} B_i^T$ and $T_i := C P_i^{-1} C^T$, exist in the interval $\left[t_0, t_f\right]$.

1. *The BVP, for $i = 1,\ldots,N$,*

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}e_i(t) = \left(A^T(t) - E_i(t)\left(S_i(t) + T_i(t)\right)\right)e_i(t) \\ \qquad\qquad +E_i(t)\sum_{j\neq i} S_j(t)\left(E_j(t)\hat{x}_j(t) + e_j(t)\right) - E_i(t)F(t) \\ e_i\left(t_f\right) \;= 0, \end{cases} \tag{17}$$

with $e_i(t) \in \mathbb{R}^n$ and $t \in \left[t_0, t_f\right]$, and

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\hat{x}_i(t) = A(t)\hat{x}_i(t) - \sum_{j=1}^N S_j(t)\left(E_j(t)\hat{x}_j(t) + e_j(t)\right) \\ \qquad\qquad -T_i(t)\left(E_i(t)\hat{x}_i(t) + e_i(t)\right) + F(t) \\ \hat{x}_i\left(t_0\right) \;= x_0, \end{cases} \tag{18}$$

with $\hat{x}_i(t) \in \mathbb{R}^n$, is equivalent to (9), (10) and (11).

2. *The control functions*

$$\tilde{u}_i(t) = -R_{ii}^{-1}(t)B_i^T(t)\left(E_i(t)\hat{x}_i(t) + e_i(t)\right) \tag{19}$$

form an OL Nash/worst-case equilibrium if and only if e_i and $\hat{x}_i$ are solutions of (17) and (18), respectively. Moreover, the corresponding worst-case disturbance of the i^{th} player is given by

$$\hat{w}_i = -P_i^{-1}(t)C^T(t)\left(E_i(t)\hat{x}_i(t) + e_i(t)\right). \tag{20}$$

3. *The Nash/worst-case equilibrium represented by (19) is unique if and only if the BVPs (17) and (18) have a unique solution.*

Proof. The proof follows as for the linear case, see [8]. □

For simplicity of exposition, we are going to consider $N = 2$. We also consider that every coefficient, the solution of the standard Riccati equation, as well as the BVP, depend on t.

Theorem 4. *Suppose that for the 2-player game the assumptions on the matrices R_{ii} and P_i and on the operators $F_i, G_i, i = 1, 2$, in Theorem 2 are fulfilled. Suppose further that the symmetric matrix Riccati differential equations (16), for $i = 1, 2$, and the following nonsymmetric matrix Riccati differential equation*

$$\dot{W} = \begin{pmatrix} 0 & E_1 S_2 E_2 \\ E_2 S_1 E_1 & 0 \end{pmatrix} + \begin{pmatrix} -A^T + E_1(S_1 + T_1) & E_1 S_2 \\ E_2 S_1 & -A^T + E_2(S_2 + T_2) \end{pmatrix} W$$

$$-W \begin{pmatrix} A - (S_1 + T_1)E_1 & -S_2 E_2 \\ -S_1 E_1 & A - (S_2 + T_2)E_2 \end{pmatrix} - W \begin{pmatrix} -S_1 - T_1 & -S_2 \\ -S_1 & -S_2 - T_2 \end{pmatrix} W$$

$$W\left(t_f\right) = 0 \in \mathbb{R}^{2n \times 2n} \tag{21}$$

admit bounded solutions over the interval $\left[t_0, t_f\right]$. *Consider* $W = \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix}$. *Then, the BVP, (17) and (18), is uniquely solvable. Moreover, using these solutions the (unique) optimal Nash/worst-case control function for each player can be obtained in the following form:*

$$\tilde{u}_1 = -R_{11}^{-1} B_1^T \left((E_1 + W_{11})\hat{x}_1 + W_{12}\hat{x}_2 + W_{11} y_1 + W_{12} y_2 - y_3\right)$$
$$\tilde{u}_2 = -R_{22}^{-1} B_2^T \left((E_2 + W_{22})\hat{x}_2 + W_{21}\hat{x}_1 + W_{21} y_1 + W_{22} y_2 - y_4\right), \tag{22}$$

as well as the worst-case disturbance:

$$\hat{w}_i = -P_i^{-1} C^T \left((E_i + W_{ii})\hat{x}_i + W_{ij}\hat{x}_j + W_{ii} y_i + W_{ij} y_j - y_{i+2}\right), i,$$
$$j = 1, 2, \quad j \neq i, \tag{23}$$

where $\hat{x}_i$ *denotes the worst-case trajectory from the point of view of the* i^{th} *player, and* $Y = (y_1, y_2, y_3, y_4)^T$ *is the solution of the following TVP:*

$$\frac{\mathrm{d}}{\mathrm{d}t} Y(t) = \underbrace{\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}}_{M(t)} Y(t) - \begin{pmatrix} F \\ F \\ -E_1 F \\ -E_2 F \end{pmatrix}$$
$$Y(t_f) = 0 \tag{24}$$

Proof. The BVP, (17) and (18), can be written in matrix form as a nonhomogeneous differential equation:

$$\frac{\mathrm{d}}{\mathrm{d}t} X(t) = M(t) X(t) + \begin{pmatrix} F \\ F \\ -E_1 F \\ -E_2 F \end{pmatrix} \tag{25}$$

$$\begin{pmatrix} \hat{x}_1(t_0) \\ \hat{x}_2(t_0) \end{pmatrix} = \begin{pmatrix} x_0 \\ x_0 \end{pmatrix}, \tag{26}$$

$$\begin{pmatrix} e_1(t_f) \\ e_2(t_f) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{27}$$

where $X(t) = \begin{pmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \\ e_1(t) \\ e_2(t) \end{pmatrix}$ and

$$M = \begin{pmatrix} A - (S_1 + T_1)\,E_1 & -S_2E_2 & -\,(S_1 + T_1) & -S_2 \\ -S_1E_1 & A - (S_2 + T_2)\,E_2 & -S_1 & -\,(S_2 + T_2) \\ 0 & E_1S_2E_2 & -A^T + E_1\,(S_1 + T_1) & E_1S_2 \\ E_2S_1E_1 & 0 & E_2S_1 & -A^T + E_2\,(S_2 + T_2) \end{pmatrix}.$$

Also consider the auxiliary TVP (24):

$$\dot{Y}(t) = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} Y(t) - \begin{pmatrix} F \\ F \\ -E_1F \\ -E_2F \end{pmatrix},$$
$$Y(t_f) = 0.$$

Adding (24) and (25) yields a homogeneous differential equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\,(X(t) + Y(t)) = M(t)\,(X(t) + Y(t)). \tag{28}$$

Consider $z(t) = X(t) + Y(t)$, and applying Radon's lemma (see [1, Chap. 3]), we solve:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix} = \begin{pmatrix} A - (S_1 + T_1)(E_1 + W_{11}) - S_2W_{21} & -(S_2(E_2 + W_{22}) - (S_1 + T_1)W_{12}) \\ -S_1(E_1 + W_{11}) - (S_2 + T_2)W_{21} & A - (S_2 + T_2)(E_2 + W_{22}) - S_1W_{12} \end{pmatrix}$$
$$\begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix}\begin{pmatrix} z_1(0) \\ z_2(0) \end{pmatrix} = \begin{pmatrix} x_0 + y_1(0) \\ x_0 + y_2(0) \end{pmatrix} \tag{29}$$

and

$$\begin{pmatrix} z_3(t) \\ z_4(t) \end{pmatrix} = W(t)\begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix}$$
$$\begin{pmatrix} z_3(t_f) \\ z_4(t_f) \end{pmatrix} = 0. \tag{30}$$

Thence, calculate the virtual trajectory:

$$\begin{pmatrix} \hat{x}(t) \\ e(t) \end{pmatrix} = \begin{pmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \\ e_1(t) \\ e_2(t) \end{pmatrix} = z(t) - Y(t) \tag{31}$$

after solving the TVP (24).

Now it is possible to calculate the equilibrium controls (22). □

4 Numerical Example

Consider that the players' cost functionals have the following form:

$$J_1 = x(t_f)^2 + \int_{t_0}^{t_f} \left(3u_1(t)^2 - 2w(t)^2\right) \mathrm{d}t \tag{32}$$

$$J_2 = 3x(t_f)^2 + \int_{t_0}^{t_f} \left(u_1(t)^2 - 4w(t)^2\right) \mathrm{d}t. \tag{33}$$

And the underlying one-dimensional system:

$$\dot{x}(t) = 3u_1(t) - u_2(t) + w(t) + 2 \tag{34}$$

$$x_0 = 30. \tag{35}$$

Similarly as for the linear case in [8], since $R_{ii} > 0$ and $\bar{Q}_i \geq 0$, we immediately obtain the positive definiteness of F_i. Also, by the operators' definition in Sect. 2, we have:

$$\left\langle w, \left(\mathscr{C}^* \bar{Q}_i \mathscr{C} + \bar{P}_i\right) w\right\rangle = \left\langle \mathscr{C} w, \bar{Q}_i \mathscr{C} w\right\rangle + \left\langle w, \bar{P}_i w\right\rangle \tag{36}$$

$$= \int_0^{0.5} P_i w(t)^2 \mathrm{d}t + K_{if} \left(\int_0^{0.5} w(t) \mathrm{d}t\right)^2 \tag{37}$$

$$\leq \int_0^{0.5} \left(K_{if} + P_i\right) w(t)^2 \mathrm{d}t < 0 \tag{38}$$

for $i = 1, 2$, which yields the negative definiteness of G_i.

Figure 1 shows the chosen disturbance.

To evaluate the reliability of our approach, we computed the solution of the BVP (17) and (18) first using the decoupled approach stated in Theorem 4, and the obtained equilibrium control profiles are represented in Fig. 2.

The same equilibrium control profiles obtained when using the Scilab routine `bvode` are described in Fig. 3.

In Fig. 2, e_1, e_2 and $\hat{x}_1$, $\hat{x}_2$ correspond to the unique solution of the BVP (17) and (18) obtained via Theorem 4 and calculated in (31), whereas in Fig. 3 e_1-BVP, e_2-BVP and $\hat{x}_1$-BVP, $\hat{x}_2$-BVP correspond to the solution of the BVP (17) and (18) obtained by direct calculation with the Scilab routine bvode.

From the observation of the equilibrium control profiles obtained with both approaches, one may observe that these profiles are similar. Moreover, the CPU time obtained when using the approach described in Theorem 4 is smaller.

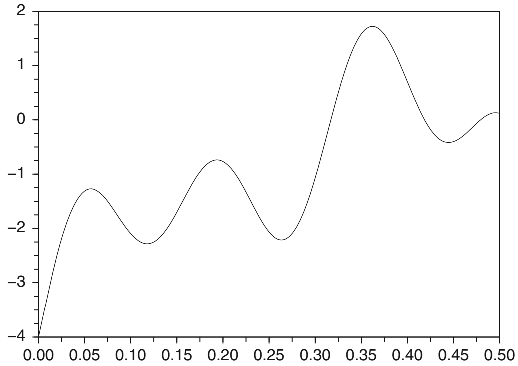


Fig. 1 Chosen disturbance

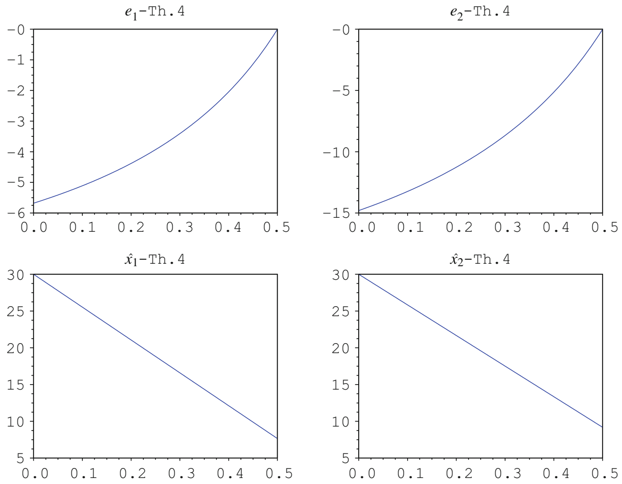


Fig. 2 CPU time = 0.1 s

5 Concluding Remarks

We investigated conditions for the existence/uniqueness of equilibrium controls of an affine disturbed differential game with finite planning horizon. The proposed algorithm to compute these equilibrium controls is based on solving initial and terminal value problems and has been applied to a case with a scalar differential equation. Its numerical performance has been compared to the performance of a Scilab BVP solver used to calculate the solution of the BVP stated in Theorem 3. The reason which makes the proposed algorithm amenable to practical cases is the following: To implement a recursive matricial BVP is not easy for problems whose number of players is $N \geq 2$, and even when $N = 2$ difficulties arise in the nonscalar

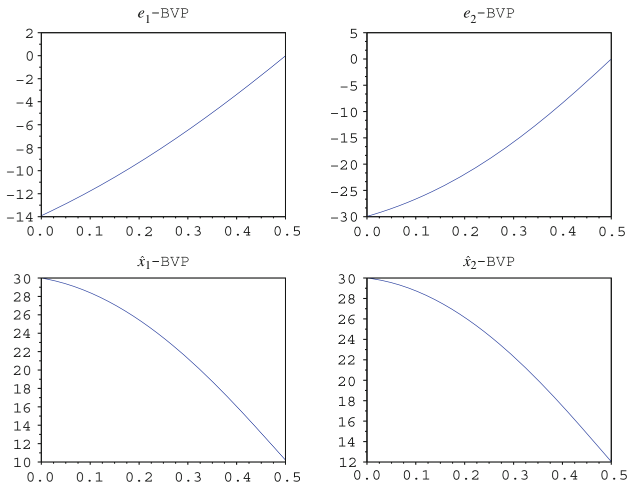


Fig. 3 CPU time = 0.41 s

case; Also the approach avoiding the solution of boundary value problems seems to have a better numerical performance in terms of CPU-times. But, notice that the BVP could still have a solution and hence an equilibrium exists, although condition (21) is violated and, moreover, the BVP need not have a unique solution yielding nonunique Nash equilibriums.

Acknowledgments The authors wish to thank the referees whose comments greatly contributed to the improvement of the work presented here.

Appendix: Algorithm

Assumptions

- All weighting matrices are symmetric
- P_i, R_{ii} are negative and positive definite, respectively
- G_i, F_i are negative and positive definite, respectively

Calculations

We understand that every coefficient varies with t, however we use a simpler notation for clarity sake.

1. Solve backward the Riccati differential equation

$$\dot{E}_i(t) = -A^T E_i - E_i A - Q_i + E_i(S_i + T_i)E_i$$
$$E_i(t_f) = K_{if}, \tag{39}$$

with $S_i = B_i R_{ii}^{-1} B_i^T$, and $T_i = CP_i^{-1}C^T$.

2. Solve backward the TVP

$$\dot{Y}(t) = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} Y(t) - \begin{pmatrix} F \\ F \\ -E_1 F \\ -E_2 F \end{pmatrix}$$
$$Y(t_f) = 0. \tag{40}$$

3. Solve backward the TVP

$$\dot{W}(t) = \underbrace{\begin{pmatrix} 0 & E_1 S_2 E_2 \\ E_2 S_1 E_1 & 0 \end{pmatrix}}_{M_{21}}$$
$$+ \underbrace{\begin{pmatrix} -A^T + E_1(S_1 + T_1) & E_1 S_2 \\ E_2 S_1 & -A^T + E_2(S_2 + T_2) \end{pmatrix}}_{M_{22}} W(t)$$
$$-W(t) \underbrace{\begin{pmatrix} A - (S_1 + T_1)E_1 & -S_2 E_2 \\ -S_1 E_1 & A - (S_2 + T_2)E_2 \end{pmatrix}}_{M_{11}}$$
$$-W(t) \underbrace{\begin{pmatrix} -S_1 - T_1 & -S_2 \\ -S_1 & -S_2 - T_2 \end{pmatrix}}_{M_{12}} W(t), \quad W(t_f) = 0 \in \mathbb{R}^{2n \times 2n} \tag{41}$$

$M_{22} \neq M_{11}^T$.

4. Solve forward the IVP

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix} = \begin{pmatrix} A - (S_1 + T_1)(E_1 + W_{11}) - S_2 W_{21} & -(S_2(E_2 + W_{22}) - (S_1 + T_1)W_{12}) \\ -S_1(E_1 + W_{11}) + (S_2 + T_2)W_{21} & A - (S_2 + T_2)(E_2 + W_{22}) - S_1 W_{12} \end{pmatrix}$$

$$\begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix} \begin{pmatrix} z_1(0) \\ z_2(0) \end{pmatrix} = \begin{pmatrix} x_0 + y_1(0) \\ x_0 + y_2(0) \end{pmatrix}. \tag{42}$$

5. With $W(t) = \begin{pmatrix} W_{11}(t), & W_{12}(t) \\ W_{21}(t), & W_{22}(t) \end{pmatrix}$, $\begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix}$, and $\begin{pmatrix} z_3(t) \\ z_4(t) \end{pmatrix}$ calculate:

$$\begin{aligned} \begin{pmatrix} z_3(t) \\ z_4(t) \end{pmatrix} &= W(t) \begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix} \\ \begin{pmatrix} z_3(t_f) \\ z_4(t_f) \end{pmatrix} &= 0. \end{aligned} \tag{43}$$

6. Calculate the virtual trajectory

$$\begin{pmatrix} \hat{x}(t) \\ e(t) \end{pmatrix} = \begin{pmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \\ e_1(t) \\ e_2(t) \end{pmatrix} = z(t) - Y(t). \tag{44}$$

7. Calculate the controls:

$$\begin{aligned} \tilde{u}_1(t) &= -R_{11}^{-1} B_1^T \left((E_1(t) + W_{11}(t))\, \hat{x}_1(t) + W_{12}(t)\hat{x}_2(t) + W_{11}(t) y_1(t) \right. \\ &\qquad \left. + W_{12}(t) y_2(t) - y_3(t) \right) \\ \tilde{u}_2(t) &= -R_{22}^{-1} B_2^T \left((E_2(t) + W_{22}(t))\, \hat{x}_2(t) + W_{21}(t)\hat{x}_1(t) + W_{21}(t) y_1(t) \right. \\ &\qquad \left. + W_{22}(t) y_2(t) - y_4(t) \right). \end{aligned} \tag{45}$$

8. Calculate the worst-case disturbance:

$$\begin{aligned} \hat{w}_i &= -P_i^{-1} C^T \big((E_i(t) + W_{ii}(t))\, \hat{x}_i(t) + W_{ij}(t)\hat{x}_j(t) + W_{ii}(t) y_i(t) \\ &\qquad + W_{ij}(t) y_j(t) - y_{i+2}(t) \big), i = 1, 2. \end{aligned} \tag{46}$$

9. Calculate the mixed trajectory

$$\dot{x}(t) = Ax(t) + B_1\tilde{u}_1(t) + B_2\tilde{u}_2(t) + C\bar{w}(t) + F(t), \tag{47}$$

where $\bar{w}$ is some chosen disturbance.

10. Calculate the values of the cost functionals at the equilibrium:

$$\begin{aligned} J_i\, (\tilde{u}_1(\cdot), \tilde{u}_2(\cdot), \bar{w}_i(\cdot)) &= x(t_f)^T K_{if} x(t_f) \\ &\quad + \int_0^{t_f} \left(x^T Q_i x + \tilde{u}_i^T R_{ii} \tilde{u}_i + \bar{w}^T P_i \bar{w} \right) \mathrm{d}t, \end{aligned} \tag{48}$$

where matrices K_{if}, Q_i, R_{ij}, P_i are symmetric.

References

1. Abou-Kandil, H., Freiling, G., Ionescu, V., Jank, G.: Matrix Riccati Equations in Control and Systems Theory. Birkhäuser, Basel (2003)
2. Azevedo-Perdicoúlis, T., Jank, G.: A 2–DAE system gas network model in view to optimization as a dynamic game. Proceedings of the 5th International Workshop in Multidimensional (nD) Systems 2007, held at Universidade do Aveiro, Aveiro, Portugal (2007)
3. Azevedo Perdicoúlis, T.-P., Jank, G.: A Disturbed Nash Game Approach for Gas Network Optimization. International Journal of Tomography and Statistics, Special Issue on: Control Applications of Optimisation – optimisation methods, differential games, time delay control systems, economics and management **6**, 43–49 (2007)
4. Basar, T., Olsder, G.: Dynamic Noncooperative Game Theory. Second ed. SIAM, London (1999)
5. Engwerda, J.C.: LQ Dynamic Optimization and Differential Games. Wiley, Chichester (2005)
6. Engwerda, J.C.: Uniqueness conditions for the affine open-loop linear quadratic differential game. Automatica **44**, 504–511 (2008)
7. Engwerda, J.C.: Addendum: Errors in LQ Dynamic Optimization and Differential Games. http://center.uvt.nl/staff/engwerda/books/framefoutjes.pdf
8. Jank, G., Kun, G.: Optimal Control of Disturbed Linear-Quadratic Differential Games. European Journal of Control **8**(2), 152–162 (2002)

Fields of Extremals and Sufficient Conditions for a Class of Open-Loop Variational Games

Dean A. Carlson and George Leitmann

Abstract In a 1967 note, Leitmann observed that coordinate transformations may be used to deduce extrema (minimizers or maximizers) of integrals in the simplest problem of the calculus of variations. This has become known as Leitmann's direct method. Subsequently, in a series of papers, starting in 2001, he revived this approach and extended it in a variety of ways. Shortly thereafter, Carlson presented an important generalization of this approach and connected it to Carathéodory's equivalent problem method. This in turn was followed by a number of joint papers addressing applications to dynamic games, multiple integrals, and other related topics. Recently, for the simplest problem of the calculus of variations in n-variables, making use of the classical notion of fields of extremals we employed Leitmann's direct method, as extended by Carlson, to present an elementary proof of Weierstrass' sufficiency theorem for strong local and global extrema. In this chapter, we extend the notion of a field of extremals to a class of dynamic games and give an extension of the classical Weierstrass sufficiency theorem for open-loop Nash equilibria.

1 Introduction

For the past several years, we have been studying a class of variational games, which may be viewed as an extension of the calculus of variations. In particular, our focus has been on exploiting a direct solution method, originally due to G. Leitmann in [11] to investigate sufficient conditions for open-loop Nash equilibria. The study

D.A. Carlson (✉)
American Mathematical Society, Mathematical Reviews, 416 Fourth Street,
Ann Arbor, MI 48103, USA
e-mail: dac@ams.org

G. Leitmann
Graduate School, University of California at Berkeley, Berkeley, CA 94720, USA
e-mail: gleit@berkeley.edu

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_3,

of such problems pre-dates J. Nash's work in noncooperative games and their study were first found in the 1920s with a series of mathematical papers by C. F. Roos [14–19] exploring the dynamics of competition in economics. The last of these papers provides an extensive investigation into general variational games and provides analogues of the standard first-order necessary conditions, such as the Euler–Lagrange equations, the Weierstrass necessary condition, transversality conditions, Legendre's necessary condition, and the Jacobi necessary condition. Additionally, it also discusses an extension of the classical sufficient condition due to Weierstrass by describing an extremal field, introducing the analogs of the Hilbert invariant integral and presenting several theorems without proof. Recently, using Leitmann's direct method, we were able to give an elementary proof of the Weierstrass sufficient condition for problems in the calculus of variations. A preliminary version of this work appears in Carlson and Leitmann [8] with a more complete version to appear in [9]. Additionally, in Carlson [3] a preliminary extension of the Weierstrass sufficient condition for a class of variational games was presented at the 12th International Symposium on Dynamic Games and Applications in Sophia Antipolis, France (ISDG-2006). The purpose of this paper is to report on an extension of our recent works to N-player variational games. In particular, we extend the classical notion of a field to the case of an N-player game and present an extension of the Weierstrass sufficient condition in this setting. This work extends the earlier work of Carlson [3] by first clearly defining a field and by allowing the states and the strategies of the *j*th player to be n_j-vector valued ($n_j \in \mathbb{N}$). We believe that our presentation is more complete than that presented in Roos [19] in that we provide complete proofs to our results.

The remainder of the paper is organized as follows. In Sect. 2, we define the class of games we consider and provide the definitions of weak and strong local open-loop Nash equilibria. In addition, we give the classical first-order necessary condition, the Euler–Lagrange system of equations, for the class of games considered. Section 3 briefly introduces the Leitmann direct sufficiency method. In Sect. 4, we introduce the notion of a field for the variational game relative to an a priori trajectory. We demonstrate how this gives rise to a family of functions which satisfy some related Euler–Lagrange equations for N single player games. Further, given a family of solutions of the Euler–Lagrange system, we give conditions under which we obtain a field for the variational game. Section 5 starts with a candidate for optimality that solves the system of Euler–Lagrange equations for the game and defines what it means for this trajectory to be embedded in a field. Using this notion and the discussion in Sect. 4, we introduce the Weierstrass E-functions and show how it relates to Leitmann's direct method. Finally, we state and prove the analog of the Weierstrass sufficiency theorem. We end our paper with an example in Sect. 6 and some concluding remarks in Sect. 7.

2 The Class of Games Considered

We consider an N-person game in which the state of player $j = 1, 2, \ldots, N$ is a real-valued function $x_j(\cdot) : [a, b] \to \mathbb{R}^{n_j}$ with fixed initial value $x_j(a) = x_{aj}$ and fixed terminal value $x_j(b) = x_{bj}$. The objective of each player is to minimize a Lagrange-type functional

$$I_j(\mathbf{x}(\cdot)) = \int_a^b L_j(t, \mathbf{x}(t), \dot{x}_j(t))\, \mathrm{d}t, \tag{1}$$

over all of his possible admissible strategies (see below), $\dot{x}_j(\cdot)$. The notation used here is that $\mathbf{x} = (x_1, x_2, \ldots x_N) \in \prod_{j=1}^{N} \mathbb{R}^{n_j} \doteq \mathbb{R}^{\mathbf{n}}$, in which $\mathbf{n} = n_1 + n_2 + \cdots + n_N$. We assume that $L_j : A_j \to \mathbb{R}$ is a continuous function defined on the open set $A_j \subset \mathbb{R} \times \mathbb{R}^{\mathbf{n}} \times \mathbb{R}^{n_j}$ with the additional properties that $L_j(t, \cdot, \cdot)$ is twice continuously differentiable on $A_j(t) \doteq \{(\mathbf{x}, p_j) : \quad (t, \mathbf{x}, p_j) \in A_j\}$.

Clearly, the strategies of the other players influence the decision of the jth player and so each player is unable to minimize independently of the other players. As a consequence, the players seek to play a (open-loop) Nash equilibrium instead. To introduce this concept, we first introduce the following notation. For each fixed $j = 1, 2, \ldots, N$, $\mathbf{x} = (x_1, x_2, \ldots, x_N) \in \mathbb{R}^{\mathbf{n}}$, and $y_j \in \mathbb{R}^{n_j}$ we use the notation

$$[\mathbf{x}^j, y_j] \doteq (x_1, x_2, \ldots, x_{j-1}, y_j, x_{j+1}, \ldots x_N).$$

With this notation, we have the following definitions.

Definition 1. We say a function $\mathbf{x}(\cdot) = (x_1(\cdot), x_2(\cdot), \ldots, x_N(\cdot)) : [a, b] \to \mathbb{R}^{\mathbf{n}}$ is an admissible trajectory for the variational game if it is continuous, has a piecewise continuous first derivative, satisfies the fixed initial and terminal conditions

$$x_j(a) = x_{aj} \quad \text{and} \quad x_j(b) = x_{bj}, \quad j = 1, 2, \ldots, N, \tag{2}$$

and $(t, x_j(t), \dot{x}_j(t)) \in A_j$ for all $t \in [a, b]^{\dagger}$, where the notation $[a, b]^{\dagger}$ consists of those points $t \in [a, b]$ for which the derivative $\dot{x}_j(t)$ exists.

Remark 1. Throughout the chapter, the notation $t \in [a, b]^{\dagger}$ will always be meant to exclude those points where the derivatives of all of the admissible trajectories appearing in the expression exist.

Definition 2. Given an admissible function $\mathbf{x}(\cdot)$ we say a function $y_j(\cdot) : [a, b] \to \mathbb{R}^{n_j}$ is an admissible trajectory for player j relative to $\mathbf{x}(\cdot)$ if the function $[\mathbf{x}^j, y_j](\cdot)$ is an admissible trajectory for the variational game.

With these definitions we can now give the definition of a Nash equilibrium.

Definition 3. An admissible trajectory for the variational game $\mathbf{x}^*(\cdot) : [a,b] \to \mathbb{R}^{\mathbf{n}}$ is called

1. A global Nash equilibrium if and only if for each player $j = 1, 2, \ldots, N$ and each function $y_j(\cdot) : [a,b] \to \mathbb{R}^{n_j}$ that is an admissible trajectory for player j relative to $\mathbf{x}^*(\cdot)$ one has

$$\begin{aligned} I_j(\mathbf{x}^*(\cdot)) &= \int_a^b L_j(t, \mathbf{x}^*(t), \dot{x}_j^*(t))\,\mathrm{d}t \\ &\leq \int_a^b L_j(t, [\mathbf{x}^*(t)^j, y_j(t)], \dot{y}_j(t))\,\mathrm{d}t \\ &= I_j([\mathbf{x}^{*j}, y_j](\cdot)). \end{aligned} \tag{3}$$

2. A strong local Nash equilibrium if and only if for each player $j = 1, 2, \ldots, N$ there exists an $\varepsilon > 0$ such that (3) holds for each function $y_j(\cdot) : [a,b] \to \mathbb{R}^{n_j}$ that is an admissible trajectory for player j relative to $\mathbf{x}^*(\cdot)$ and which satisfies

$$\|x_j^*(t) - y_j(t)\|_{\mathbb{R}^{n_j}} < \varepsilon, \quad t \in [a,b]. \tag{4}$$

3. A weak local Nash equilibrium if and only if for each player $j = 1, 2, \ldots, N$ there exists an $\varepsilon > 0$ such that (3) holds for each function $y_j(\cdot) : [a,b] \to \mathbb{R}^{n_j}$ that is an admissible trajectory for player j relative to $\mathbf{x}^*(\cdot)$ and which satisfies (4) and

$$\|\dot{x}_j^*(t) - \dot{y}_j(t)\|_{\mathbb{R}^{n_j}} < \varepsilon, \quad t \in [a,b]^{\dagger}.$$

Remark 2. From the above definitions, it is clear that when all of the players "play" a Nash equilibrium, then each player's strategy is his/her best response to that of the other players. In other words, if player j applies any other strategy than the equilibrium strategy, his/her cost functional will not decrease. Further, the notions of strong and weak local Nash equilibria are the natural extensions of their classical analogues from the calculus of variations. In addition, it is clear that a global Nash equilibrium is a strong local Nash equilibrium, which in turn is a weak local Nash equilibrium, since each new definition places a restriction on the feasible strategy trajectories considered. Thus, necessary conditions stated for a weak local Nash equilibrium are equally valid for both strong local and global Nash equilibria.

Remark 3. The above dynamic game clearly is not the most general structure one can imagine, even in a variational framework. In particular, the cost functionals are coupled only through their state variables and not through their strategies (i.e., their time derivatives). While not the most general, one can argue that this form is general enough to cover many cases of interest since in a "real-world setting," an individual player will not know the strategies of the other players (see e.g., Dockner and Leitmann [10]).

The similarity of the above dynamic game to a free problem in the calculus of variations begs the question as to how much of the classical variational theory can be extended to this setting. It is indeed one aspect of this question that partly motivates this paper. We first recall the classical first-order necessary condition for this problem, giving the following theorem.

Theorem 1. *If* $\mathbf{x}^*(\cdot)$ *is a weak local Nash equilibrium, then the following system of Euler–Lagrange equations are satisfied:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L_j}{\partial p_j}\left(t,\mathbf{x}^*(t),\dot{x}_j^*(t)\right) = \frac{\partial L_j}{\partial x_j}\left(t,\mathbf{x}^*(t),\dot{x}_j^*(t)\right),\ t \in [a,b]^{\dagger},\quad j = 1,2\ldots N. \tag{5}$$

Proof. The proof follows directly from the classical theory of the calculus of variations upon the recognition that for each $j = 1,2\ldots N$ the trajectory $x_j^*(\cdot)$ minimizes the functional

$$I_j([\mathbf{x}^*(\cdot)^j, y_j(\cdot)]) = \int_a^b L_j(t,[\mathbf{x}^*(t)^j, y_j(t)], \dot{y}_j(t))\,\mathrm{d}t$$

over all continuous functions $y_j(\cdot) : [a,b] \to \mathbb{R}$ with piecewise continuous derivatives, satisfying the fixed end conditions given by (2). □

Clearly, this is a set of standard necessary conditions and as such only provides candidates (i.e., the usual suspects!) for a Nash equilibrium. To insure that the candidate is indeed a Nash equilibrium one must show that it satisfies some additional conditions. In the classical variational theory, these conditions typically mean that one has to show that the candidate can be embedded in a field of extremals and that some additional convexity conditions are also satisfied. As we shall see shortly, an analogous theory can also be developed here.

3 Leitmann's Direct Method

In this section, we briefly outline a coordinate transformation method originally developed for single-player games in [11] and [12] (see also [2]), and further extended to N-player games in [10] and [4], which will enable us to derive our results. In particular, in [13] (see also [4]) we have the following theorem.

Lemma 1. *For* $j = 1,2,\ldots,N$ *let* $x_j = z_j(t,\tilde{x}_j)$ *be a transformation of class* C^1 *having a unique inverse* $\tilde{x}_j = \tilde{z}_j(t,x_j)$ *for all* $t \in [a,b]$ *such that there is a one-to-one correspondence* $\mathbf{x}(t) \Leftrightarrow \tilde{\mathbf{x}}(t)$, *for all admissible trajectories* $\mathbf{x}(\cdot)$ *satisfying the boundary conditions (2) and for all* $\tilde{\mathbf{x}}(\cdot)$ *satisfying*

$$\tilde{x}_j(a) = \tilde{z}_j(a,x_{aj}) \quad \textit{and} \quad \tilde{x}_j(b) = \tilde{z}_j(b,x_{bj})$$

for all $j = 1, 2, \ldots, N$. *Further for each* $j = 1, 2, \ldots, N$ *let* $\tilde{L}_j(\cdot,\cdot,\cdot) : [a,b] \times \mathbb{R}^{\mathbf{n}} \times \mathbb{R}^{n_j} \to \mathbb{R}$ *be a given integrand having the same properties as* $L_j(\cdot,\cdot,\cdot)$. *For a given admissible* $\mathbf{x}^*(\cdot) : [a,b] \to \mathbb{R}^{\mathbf{n}}$ *suppose the transformations* $x_j = z_j(t, \tilde{x}_j)$ *are such that there exist* C^1 *functions* $G_j(\cdot,\cdot) : [a,b] \times \mathbb{R} \to \mathbb{R}$ *so that the functional identities*

$$L_j(t, [\mathbf{x}^*(t)^j, x_j(t)], \dot{x}_j(t)) - \tilde{L}_j(t, [\mathbf{x}^*(t)^j, \tilde{x}_j(t)], \dot{\tilde{x}}_j(t)) = \frac{\mathrm{d}}{\mathrm{d}t} G_j(t, \tilde{x}_j(t)) \quad (6)$$

hold on $[a,b]^\dagger$. *If* $\tilde{x}_j^*(\cdot)$ *yields an extremum of* $\tilde{I}_j([\mathbf{x}^*(\cdot)^j, \cdot])$ *with* $\tilde{x}_j^*(\cdot)$. *satisfying the transformed boundary conditions, then* $x_j^*(\cdot)$ *with* $x_j^*(t) = z_j(t, \tilde{x}^*(t))$ *yields an extremum for* $I_j([\mathbf{x}^{*j}(\cdot), \cdot])$ *with the boundary conditions (2). Moreover, the function* $\mathbf{x}^*(\cdot)$ *is an open-loop Nash equilibrium for the variational game.*

Proof. See [13]. □

Remark 4. In the above, for any piecewise smooth function $\tilde{y}_j(\cdot) : [a,b] \to \mathbb{R}^{n_j}$ the functional $\tilde{I}_j([\mathbf{x}^*(\cdot)^j, \cdot])$ is defined by the formula

$$\tilde{I}_j([\mathbf{x}^*(\cdot)^j, \tilde{y}_j(\cdot)]) = \int_a^b \tilde{L}_j(t, [\mathbf{x}^*(t)^j, \tilde{y}_j(t)], \dot{\tilde{y}}_j(t))\,\mathrm{d}t.$$

Remark 5. This result has been successfully applied to a number of examples having applications in mathematical economics and we refer the reader to the references. Additionally, this lemma has been extended to include various classes of control systems (e.g., affine in the strategies[7]), infinite horizon models [6], as well as multiple integral problems [5].

Two immediate and useful corollaries are the following.

Corollary 1. *The existence of* $G_j(\cdot,\cdot)$, $j = 1, 2 \ldots, N$, *in (6) imply that the following identities hold for* $(t, \tilde{x}_j) \in (a,b) \times \mathbb{R}$ *and* $\tilde{q}_j \in \mathbb{R}^{n_j}$ *for* $j = 1, 2, \ldots, N$:

$$\begin{aligned} L_j\left(t, [\mathbf{x}^{*j}(t), z_j(t,\tilde{x}_j)], \frac{\partial z_j(t,\tilde{x}_j)}{\partial t} + \frac{\partial z_j(t,\tilde{x}_j)}{\partial \tilde{x}_j}\tilde{q}_j\right) - \tilde{L}_j(t, [\mathbf{x}^{*j}(t), \tilde{x}_j], \tilde{q}_j) \\ \equiv \frac{\partial G_j(t,\tilde{x}_j)}{\partial t} + \left(\frac{\partial G_j(t,\tilde{x}_j)}{\partial \tilde{x}_j}\right)^{\mathsf{T}} \tilde{q}_j. \end{aligned} \quad (7)$$

Corollary 2. *For each* $j = 1, 2, \ldots, N$ *the left-hand side of the identity, (7) is linear in* $\tilde{q}_j$, *that is, it is of the form*

$$\Theta_j(t, \tilde{x}_j) + \Psi_j(t, \tilde{x}_j)^{\mathsf{T}} \tilde{q}_j$$

and,

$$\frac{\partial G_j(t,\tilde{x}_j)}{\partial t} = \Theta_j(t,\tilde{x}_j) \quad \text{and} \quad \frac{\partial G_j(t,\tilde{x}_j)}{\partial \tilde{x}_j} = \Psi(t,\tilde{x}_j)$$

on $[a,b] \times \mathbb{R}$.

The utility of the above lemma rests in being able to choose not only the transformation $\mathbf{z}(\cdot,\cdot)$ but also the integrands $\tilde{L}_j(\cdot,\cdot,\cdot)$ and the functions $G_j(\cdot,\cdot)$. It is this flexibility that will enable us to extend the classical calculus of variations theory to the class of dynamic games considered here.

To provide further insight to the direct method of Lemma 1, we wish to suppress the dependence of the other players by introducing the notation $L_j^*(t, x_j, p_j) = L_j(t, [\mathbf{x}^*(t)^j, x_j], p_j)$ and $\tilde{L}_j^*(t, \tilde{x}_j, \tilde{p}_j) = \tilde{L}_j(t, [\mathbf{x}^*(t)^j, \tilde{x}_j], \tilde{p}_j)$. With this notation, the relation (6) becomes

$$L_j^*(t, x_j(t), \dot{x}_j(t)) - \tilde{L}_j^*(t, \tilde{x}_j(t), \dot{\tilde{x}}_j(t)) = \frac{\mathrm{d}}{\mathrm{d}t} G_j(t, \tilde{x}_j(t)).$$

Thus, if $\tilde{x}_j^*(\cdot)$ is an extremum $\tilde{I}_j([\mathbf{x}^*(\cdot)^j, \cdot])$ it necessarily is an extremum of the calculus of variations functional

$$\tilde{y}_j(\cdot) \mapsto \int_a^b \tilde{L}_j^*(t, \tilde{y}_j(t), \dot{\tilde{y}}_j(t))\, \mathrm{d}t.$$

As we shall see in the next sections, this observation allows us to focus our attention on problems of the calculus of variations.

4 Fields of Trajectories

We start by letting $J \subset \mathbb{R}$ be an open interval and letting $D_j \subset \mathbb{R}^{n_j}$, $j = 1, 2, \ldots N$, be such that $\mathscr{D}_j \doteq J \times D_j$ are N open connected regions. Let $\hat{p}_j(\cdot,\cdot) : J \times D \to \mathbb{R}^{n_j}$ be continuously differentiable functions in which $D = D_1 \times D_2 \times \cdots \times D_N \subset \mathbb{R}^{\mathbf{n}}$ and let $\hat{\mathbf{p}}(\cdot,\cdot) = (\hat{p}_1(\cdot,\cdot), \ldots, \hat{p}_N(\cdot,\cdot))$. Further let $\mathbf{x}^*(t) = (x_1^*(t), x_2^*(t), \ldots, x_N^*(t)) : J \to \times_{j=1}^N \mathbb{R}^{n_j}$ be a twice continuously differentiable function. With this notation, we have the following definition.

Definition 4. We say the region $\mathscr{D} \doteq J \times D \subset \mathbb{R} \times \mathbb{R}^{\mathbf{n}}$ is a field for the functionals (1) with slope function $\hat{\mathbf{p}}(\cdot,\cdot) : \mathscr{D} \to \mathbb{R}^{\mathbf{n}}$ relative to the function $\mathbf{x}^*(\cdot)$ if and only if it satisfies the two conditions

1. The components of the vector-valued function $\hat{\mathbf{p}}(\cdot,\cdot)$ have continuous first-order partial derivatives in $\mathscr{D}$.
2. The "Hilbert invariant integrals"

$$I_C^j = \int_C \Bigg[L_j(t, [\mathbf{x}^*(t)^j, x_j], \hat{p}_j(t, [\mathbf{x}^*(t)^j, x_j])) - \left(\left.\frac{\partial L_j}{\partial p_j}\right|_{(t, [\mathbf{x}^*(t)^j, x_j], \hat{p}_j(t, [\mathbf{x}^*(t)^j, x_j]))} \right)^{\mathsf{T}} \hat{p}_j(t, [\mathbf{x}^*(t)^j, x_j]) \Bigg] \mathrm{d}t$$

$$+ \left(\frac{\partial L_j}{\partial p_j} \bigg|_{(t,[\mathbf{x}^*(t)^j, x_j], \hat{p}_j(t,[\mathbf{x}^*(t)^j, x_j]))} \right)^{\mathsf{T}} \mathrm{d}x_j \tag{8}$$

depend only on the end points of the curve along which it is taken (i.e. it is path independent) for any curve C lying completely in $\mathscr{D}_j$.

Remark 6. To compare this with the classical definition of a field, as found in [1], we focus on a fixed index j and let $L_j^*(\cdot,\cdot,\cdot) : [a,b] \times \mathbb{R}^{n_j} \times \mathbb{R}^{n_j} \to \mathbb{R}$ be the integrand defined as $L_j^*(t, x_j, p_j) \doteq L_j(t, [\mathbf{x}^*(t)^j, x_j], p_j)$. With this notation, (8) becomes

$$I_C^j = \int_C \left[L_j^*(t, x_j, \hat{p}_j^*(t, x_j)) - \left(\frac{\partial L_j^*}{\partial p_j} \bigg|_{(t, x_j, \hat{p}_j^*(t, x_j))} \right)^{\mathsf{T}} \hat{p}_j^*(t, x_j) \right] \mathrm{d}t$$
$$+ \left(\frac{\partial L_j^*}{\partial p_j} \bigg|_{(t, x_j, \hat{p}_j^*(t, x_j))} \right)^{\mathsf{T}} \mathrm{d}x_j,$$

in which $\hat{p}_j^*(t, x_j) \doteq \hat{p}_j(t, [\mathbf{x}^*(t)^j, x_j])$. From this it is easy to see that the region $\mathscr{D}_j$ is a field for the functional

$$\hat{I}_j^*(x_j(\cdot)) = \int_a^b L_j^*(t, x_j(t), \dot{x}_j(t))\, \mathrm{d}t \tag{9}$$

with slope function $\hat{p}_j^*(\cdot,\cdot)$ in the classical sense.

Associated with this field for each $j = 1, 2, \ldots N$ we define the *trajectories of the field* relative to $\mathbf{x}^*(\cdot)$ as the solutions of the system of first-order differential equations

$$\dot{x}_j(t) = \hat{p}_j^*(t, x_j(t)) \doteq \hat{p}_j(t, [\mathbf{x}^*(t)^j, x_j(t)]).$$

Through each point $(t, x_j) \in \mathscr{D}_j$, as a consequence of the existence theory for ordinary differential equations, there passes one and only one solution of this differential equation and therefore the collection of all such trajectories is an n_j-parameter family of trajectories which we denote by $x_j(\cdot) = \varphi_j(\cdot, \beta_j)$ in which $\varphi_j(\cdot,\cdot)$ is a twice continuously differentiable function defined on a region $\mathscr{R}_j$ in the space of variables $(t, \beta_j) \subset \mathscr{R}_j \subset J \times \mathbb{R}^{n_j}$. If one follows the presentation in [1], as a consequence of the path independence of the integral (8) it follows easily that $t \to \varphi(t, \beta_j)$ satisfies the Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L_j^*}{\partial p_j} \bigg|_{(t, \varphi_j(t,\beta_j), \dot{\varphi}_j(t,\beta_j))} \right) = \frac{\partial L_j^*}{\partial x_j} \bigg|_{(t, \varphi_j(t,\beta_j), \dot{\varphi}_j(t,\beta_j))}, \tag{10}$$

in which $\dot{\varphi}_j(\cdot,\cdot)$ denotes the partial derivative of $\varphi_j(\cdot,\cdot)$ with respect to t (see [1, Chap. 2]). That is, the functions $x_j(\cdot) = \varphi_j(\cdot,\beta_j) = \varphi_j(\cdot,\beta_{1j},\ldots,\beta_{n_j j})$ form an n_j-parameter family of solutions of the above Euler–Lagrange equations. Following the classical theory a little further, a natural question to ask now is under what conditions does an n_j-parameter family of solutions, say $x_j(\cdot) = \varphi_j(\cdot,\beta_j)$, of (10) correspond to a field $\mathscr{D}_j$ for the functional (9) and what is the corresponding slope function $\hat{p}_j^*(\cdot,\cdot)$. We now direct our attention to this question.

To this end, let $\mathbf{x}^*(\cdot)$ be a twice continuously differentiable function defined on an interval J and let $\mathscr{D}_j$ be as described above in Definition 4. For each fixed $j = 1,2,\ldots,N$ consider an n_j-parameter family of functions $x_j(\cdot) \doteq \varphi_j^*(\cdot,\beta_j)$ defined on J. For each fixed $\beta_j \in \mathbb{R}^{n_j}$, the map $t \to \varphi_j^*(t,\beta_j)$ defines a curve in $\mathbb{R}^{n_j}$. We further suppose that this family of curves "simply covers" the region $\mathscr{D}_j \subset J \times \mathbb{R}^n$ in the sense that through every point $(t,x_j) \in \mathscr{D}_j$ exactly one and only one member of the family passes through this point (i.e., there exists only one choice of β_j such that $x_j = \varphi_j^*(t,\beta_j)$). In this way, this family of curves defines a point set $\mathscr{R}_j$ in the space of variables $(t,\beta_j) \in \mathbb{R} \times \mathbb{R}^{n_j}$, which we assume is simply connected. Further, we assume that the functions $\varphi_j^*(\cdot,\cdot)$ are twice continuously differentiable in $\mathscr{R}_j$, that the Jacobian

$$\det\left(\frac{\partial \varphi_j^*}{\partial \beta_j}\right) \neq 0, \tag{11}$$

everywhere in $\mathscr{R}_j$ and that for each β_j the vector valued function $t \to \varphi_j^*(t,\beta_j)$ satisfies the Euler–Lagrange equations (10). A consequence of (11) is that we can uniquely solve the equation $x_j = \varphi_j^*(t,\beta_j)$ for β_j to obtain a twice continuously differentiable function which we denote by $\beta_j = \psi_j^*(t,x_j)$, for $(t,x_j) \in \mathscr{D}_j$. Now, through each given point $(t,x_j) \in \mathscr{D}_j$ there exists one and only one curve of the family which we denote by the function $s \to y_j(s) = \varphi_j^*(s,\psi_j^*(t,x_j))$, $s \in J$ (i.e., we select the unique function from the family $\varphi_j^*(\cdot,\beta_j)$ corresponding to the parameter $\beta_j = \psi_j^*(t,x_j)$). The slope of the curve at any point $(s,y_j(s))$ on the curve is given by the formula

$$\dot{y}_j(s) = \dot{\varphi}_j^*(s,\beta_j) = \dot{\varphi}_j^*(s,\psi_j^*(t,x_j)).$$

In particular, we notice that when $s = t$ this becomes

$$\dot{y}_j(t) = \dot{\varphi}_j^*(t,\psi_j^*(t,x_j)).$$

As $(t,x_j) \in \mathscr{D}_j$ was chosen arbitrarily the mapping $(t,x_j) \to \dot{\varphi}_j^*(t,\psi_j^*(t,x_j))$ is well defined on $\mathscr{D}_j$. With this observation, we select as a candidate for the "slope function" $\hat{p}_j^*(\cdot,\cdot) : \mathscr{D}_j \to \mathbb{R}^n$ defined by the formula

$$\hat{p}_j^*(t,x_j) = \dot{\varphi}_j^*(t,\psi_j^*(t,x_j)). \tag{12}$$

This function has continuous partial derivatives of the first order in $\mathscr{D}_j$.

Remark 7. We emphasize that the functions $\hat{p}_j^*(\cdot,\cdot)$ implicitly depend on the fixed function $\mathbf{x}^*(\cdot)$ since it implicitly appears in equation (10), since $L_j^*(t,x_j,p_j) = L_j(t,[\mathbf{x}^*(t)_j,x_j],p_j)$. As this can be done for each $j = 1,2,\ldots N$ we obtain as our candidate for the slope function $\hat{\mathbf{p}}(\cdot,\cdot) = (\hat{p}_1^*(\cdot,\cdot),\ldots,\hat{p}_N^*(\cdot,\cdot))$ defined on $\mathscr{D}$ into $\mathbb{R}^{\mathbf{n}}$.

Our goal now is to find conditions which allow the region $\mathscr{D}$ to be a field for the functionals (1) relative to $\mathbf{x}(\cdot)$ with slope function $\hat{\mathbf{p}}(\cdot,\cdot)$. Observe that, as a consequence of our construction, if $\mathscr{D}$ is a field relative to $\mathbf{x}(\cdot)$ with slope function $\hat{\mathbf{p}}(\cdot,\cdot)$ the n_j-parameter families we started will be the trajectories of the field, since any curve $y_j(t) = \varphi_j^*(t,\beta_j)$ (for some parameter $\beta_j \in \mathbb{R}^{n_j}$) necessarily satisfies the differential equation

$$\dot{y}_j(t) = \dot{\varphi}_j^*(t,\psi_j^*(t,y_j(t))) = \hat{p}_j^*(t,y_j(t)),$$

since by definition $\beta_j = \psi_j^*(t,y_j(t))$ for all $t \in J$. To find these additional conditions, we rewrite the jth Hilbert invariant integral in terms of the variables (t,β_j) as

$$\begin{aligned} I_C^j = \int_{\mathscr{C}} & L_j^*(t,\varphi_j^*(t,\beta_j),\dot{\varphi}_j^*(t,\beta_j))\,\mathrm{d}t \\ & + \left(\left.\frac{\partial L_j^*}{\partial p}\right|_{(t,\varphi_j^*(t,\beta_j),\dot{\varphi}_j^*(t,\beta_j))} \right)^{\mathsf{T}} \left.\frac{\partial \varphi_j^*}{\partial \beta_j}\right|_{(t,\beta_j)} \mathrm{d}\beta_j, \end{aligned}$$

in which $L_j^*(t,x_j,p_j) = L_j(t,[x^*(t)^j,x_j],p_j)$, $\mathscr{C}$ is the path in $\mathscr{R}_j$ corresponding to the path C in $\mathscr{D}_j$, $\partial\varphi_j^*/\partial\beta_j$ denotes the Jacobian matrix of $\varphi_j^*(t,\cdot)$ (considered as a function of β_j alone for each fixed t), and $\dot{\varphi}_j^*(t,\beta_j)$ denotes the partial derivative of $\varphi_j^*(\cdot,\cdot)$ with respect to t. Now for this integral to be path independent it is necessary (and sufficient since $\mathscr{R}_j$ is simply connected) that the integrand be an exact differential. That is, there exists a function $G_j^*(\cdot,\cdot)$ such that

$$\begin{aligned} &\frac{\partial G_j^*}{\partial t}(t,\beta_j) = L_j^*(t,\varphi_j^*(t,\beta_j),\dot{\varphi}_j^*(t,\beta_j)) \\ &\frac{\partial G_j^*}{\partial \beta_{kj}}(t,\beta_j) = \sum_{i=1}^{n_j} \frac{\partial \varphi_{ij}^*}{\partial \beta_{kj}}(t,\beta_j)\frac{\partial L_j^*}{\partial p_{ij}}(t,\varphi_j^*(t,\beta_j),\dot{\varphi}_j^*(t,\beta_j)), \quad k = 1,2,\ldots,n_j, \end{aligned} \tag{13}$$

where $\beta_j = (\beta_{1j},\beta_{2j},\ldots\beta_{n_{jj}})$. The conditions for such a $G_j^*(\cdot,\cdot)$ to exist is that all of the second-order mixed partial derivatives of $G_j^*(\cdot,\cdot)$ be equal. This means that we first have the conditions

$$\frac{\partial L_j^*}{\partial \beta_{kj}} = \frac{\partial}{\partial t}\left(\sum_{i=1}^{n_j} \frac{\partial \varphi_{ij}^*}{\partial \beta_{kj}} \frac{\partial L_j^*}{\partial p_{ij}} \right), \tag{14}$$

for $k = 1, 2, \ldots, n_j$, and the conditions

$$\frac{\partial}{\partial \beta_{rj}} \left(\sum_{i=1}^{n_j} \frac{\partial \varphi_{ij}^*}{\partial \beta_{sj}} \frac{\partial L_j^*}{\partial p_{ij}} \right) = \frac{\partial}{\partial \beta_{sj}} \left(\sum_{i=1}^{n_j} \frac{\partial \varphi_{ij}^*}{\partial \beta_{rj}} \frac{\partial L_j^*}{\partial p_{ij}} \right), \tag{15}$$

for $r, s = 1, 2, \ldots, n_j$, $r \neq s$. In the above two equations, the partial derivatives of $L_j^*(\cdot,\cdot,\cdot)$ are evaluated at $(t, \varphi_j^*(t, \beta_j), \dot{\varphi}_j^*(t, \beta_j))$ and the partial derivatives of $\varphi_{ij}^*(\cdot,\cdot)$ are evaluated at (t, β_j). Carrying out the differentiation in (14) we have

$$\sum_{i=1}^{n_j} \frac{\partial L_j^*}{\partial x_i} \frac{\partial \varphi_{ij}^*}{\partial \beta_{kj}} + \sum_{i=1}^{n_j} \frac{\partial L_j^*}{\partial p_i} \frac{\partial \dot{\varphi}_{ij}^*}{\partial \beta_{kj}} = \sum_{i=1}^{n_j} \frac{\partial L_j^*}{\partial p_i} \frac{\partial^2 \varphi_{ij}^*}{\partial t \partial \beta_{kj}} + \sum_{i=1}^{n_j} \frac{\partial^2 L_j^*}{\partial t \partial p_i} \frac{\partial \varphi_{ij}^*}{\partial \beta_{kj}},$$

for $k = 1, 2, \ldots, n_j$, where again we have suppressed the arguments of the functions for the sake of brevity. Using the fact that $\varphi_{kj}^*(\cdot,\cdot)$ is twice continuously differentiable we have that

$$\frac{\partial \dot{\varphi}_{ij}^*}{\partial \beta_{kj}} = \frac{\partial^2 \varphi_{ij}^*}{\partial \beta_{kj} \partial t} = \frac{\partial^2 \varphi_{ij}^*}{\partial t \partial \beta_{kj}},$$

for $i, k = 1, 2, \ldots, n_j$ so that the above expansion reduces to

$$\sum_{i=1}^{n_j} \left\{ \frac{\partial}{\partial t} \left(\frac{\partial L_j^*}{\partial p_i} \right) - \frac{\partial L_j^*}{\partial x_i} \right\} \frac{\partial \varphi_{ij}^*}{\partial \beta_{kj}} = 0, \quad k = 1, 2, \ldots, n_j.$$

This last set of equalities is automatically satisfied since we have assumed that the n_j-parameter family $\varphi_j^*(\cdot,\cdot)$ satisfies the Euler–Lagrange equations (10). Thus, we immediately have that (14) is satisfied. Expanding (15) we have

$$\begin{aligned} &\sum_{i=1}^{n_j} \left\{ \frac{\partial^2 \varphi_{ij}^*}{\partial \beta_{rj} \partial \beta_{sj}} \left(\frac{\partial L_j^*}{\partial p_i} \right) + \frac{\partial \varphi_{ij}^*}{\partial \beta_{sj}} \frac{\partial}{\partial \beta_{rj}} \left(\frac{\partial L_j^*}{\partial p_i} \right) \right\} \\ &\quad = \sum_{i=1}^{n_j} \left\{ \frac{\partial^2 \varphi_{ij}^*}{\partial \beta_{sj} \partial \beta_{rj}} \left(\frac{\partial L_j^*}{\partial p_i} \right) + \frac{\partial \varphi_{ij}^*}{\partial \beta_{rj}} \frac{\partial}{\partial \beta_{sj}} \left(\frac{\partial L_j^*}{\partial p_i} \right) \right\}, \end{aligned}$$

for $r, s = 1, 2, \ldots, n_j$, $r \neq s$. Again, since $\varphi_j^*(\cdot,\cdot)$ is twice continuously differentiable this last expression reduces to

$$\sum_{i=1}^{n_j} \left\{ \frac{\partial \varphi_{ij}^*}{\partial \beta_{sj}} \frac{\partial}{\partial \beta_{rj}} \left(\frac{\partial L_j^*}{\partial p_i} \right) - \frac{\partial \varphi_{ij}^*}{\partial \beta_{rj}} \frac{\partial}{\partial \beta_{sj}} \left(\frac{\partial L_j^*}{\partial p_i} \right) \right\} = 0, \tag{16}$$

for $r, s = 1, 2, \ldots n_j$ (observe that the summand corresponding to $r = s$ is automatically zero). The left-hand side of (16) is referred to as a Lagrange bracket. Thus, to summarize what we have done we state the following.

Theorem 2. *For the n_j-parameter families of functions $\varphi_j^*(\cdot,\cdot)$, $j = 1, 2, \dots N$, satisfying the conditions outlined at the beginning of this section to be the trajectories of a field $\mathscr{D}$ relative to $\mathbf{x}^*(\cdot)$ for the functional (1) with slope functions $\hat{p}_j^*(\cdot,\cdot)$ given by (12) it is necessary and sufficient that all of the Lagrange brackets (i.e., the left-hand side of (16) for all $r, s = 1, 2, \dots, n$) vanish identically on $\mathscr{R}_j$ for each $j = 1, 2, \dots N$.*

Remark 8. We notice that since (16) is automatically satisfied when $r = s$, it follows that whenever $n_j = 1$ for each $j = 1, 2, \dots N$ it suffices to show that the families $\varphi_j^*(\cdot,\cdot)$ satisfy the Euler–Lagrange equations (10).

We conclude our discussion of fields by making some connections to Leitmann's direct sufficiency method. Let $\mathscr{D} \doteq J \times D$ be a field of for the objective functional (1) relative to the function $\mathbf{x}^*(\cdot)$ with slope function $\hat{\mathbf{p}}(\cdot,\cdot)$ and let $x_j(\cdot) = \varphi_j^*(\cdot, \beta_j)$ $(j = 1, 2, \dots, N)$ denote the trajectories for the field relative to $\mathbf{x}^*(\cdot)$. For each $j = 1, 2, \dots, N$ and each piecewise smooth trajectory $y_j(\cdot) : [a, b] \to \mathbb{R}^{n_j}$ whose graph lies in $\mathscr{D}$ (i.e., $\{(t, y_j(t)) : t \in J\} \subset \mathscr{D}_j$) we can define, the unique piecewise smooth function $\tilde{y}_j(\cdot) : J \to \mathbb{R}^{n_j}$ by means of the equation

$$y_j(t) = \varphi_j^*(t, \tilde{y}_j(t)), \quad t \in J, \tag{17}$$

or equivalently by the inverse relation

$$\tilde{y}_j(t) = \psi_j^*(t, y_j(t)), \quad t \in J. \tag{18}$$

Moreover, we observe that the graph of $\tilde{y}_j(\cdot)$ is contained in $\mathscr{R}_j$ (i.e., $\{(t, \tilde{y}_j(t)) : t \in J\} \subset \mathscr{R}_j$). Conversely, if $\tilde{y}_j(\cdot)$ is a piecewise smooth trajectory whose graph lies in $\mathscr{R}_j$, then one can uniquely define a trajectory $y_j(\cdot) : J \to \mathbb{R}^{n_j}$ whose graph lies in $\mathscr{D}_j$ by means of (17). In this way, we see that the n_j-parameter family of trajectories can be used to establish a one-to-one correspondence between the piecewise smooth trajectories with graphs in $\mathscr{D}_j$ and the piecewise smooth trajectories with graphs in $\mathscr{R}_j$. Further, if we restrict $y_j(\cdot)$ to satisfy the fixed end conditions (2), then $\tilde{y}_j(\cdot) = \psi_j^*(\cdot, y_j(\cdot))$ satisfies the end conditions

$$\tilde{y}_j(a) = \tilde{x}_{aj} = \psi_j^*(a, x_{aj}) \quad \text{and} \quad \tilde{y}_j(b) = \tilde{x}_{bj} = \psi_j^*(b, x_{bj}).$$

From the above, we see that we have a transformation of coordinates which we can exploit in Leitmann's direct sufficiency method. More specifically, we can define $z_j(t, \tilde{x}_j) = \varphi_j^*(t, \tilde{x}_j)$ for each $(t, \tilde{x}_j) \in \mathscr{R}_j$. Furthermore, the inverse transformation $\tilde{z}_j(\cdot,\cdot) : \mathscr{D}_j \to \mathbb{R}^{n_j}$ is defined by $\tilde{z}_j(t, x_j) = \psi_j^*(t, x_j)$ for each $(t, x_j) \in \mathscr{D}_j$. In addition, we also have that because the Hilbert invariant integral (8) is path independent, there exists a function $G_j^*(\cdot,\cdot)$ for which (13) holds on $\mathscr{R}_j$. Thus, we see that two of the three components required to apply the direct sufficiency method can be obtained by this extension of the classical field theory.

5 The Weierstrass Sufficiency Theorem

We now return to our original problem beginning with the following definition.

Definition 5. Let $\mathbf{x}^*(\cdot) : [a,b] \times \mathbb{R}^{\mathbf{n}}$ be a twice continuously differentiable solution of the Euler–Lagrange equations (5) satisfying the fixed end conditions (2) and let $\mathscr{D} \doteq [a,b] \times D \subset \mathbb{R} \times \mathbb{R}^{\mathbf{n}}$ be a field for the functionals (1) relative to $\mathbf{x}^*(\cdot)$ with slope function $\hat{\mathbf{p}}(\cdot,\cdot)$. We say $\mathbf{x}^*(\cdot)$ is embedded in the field $\mathscr{D}$ for the functional (1) if and only if it corresponds to one of the trajectories of the field, that is, if there exists a (unique) parameter $\mathbf{fi}^* = (\beta_1^*, \beta_2^*, \ldots, \beta_N^*)$ such that $x_j^*(\cdot) = \varphi_j^*(\cdot, \beta_j^*)$.

Observe that if $\mathbf{x}^*(\cdot)$ is embedded in a field $\mathscr{D}$ for the functional (1) relative to itself, then, for each $j = 1,2,\ldots,N$ there exists a parameter β_j^* such that $x_j^*(t) = \varphi_j^*(t, \beta_j^*)$ for all $t \in [a,b]$. This means that we have $\beta_j^* = \psi_j^*(t, x_j^*(t))$ for all $t \in [a,b]$ and, in particular, we have $\beta_j^* = \psi_j^*(a, x_{aj}) = \psi_j^*(b, x_{bj})$. Thus for every piecewise smooth trajectory $y_j(\cdot)$ that satisfies the fixed end conditions (2), its related trajectory $\tilde{y}_j(\cdot) = \psi_j^*(t, y_j(\cdot))$ defined by (18) satisfies the periodic end conditions $\tilde{y}_j(a) = \beta_j^*$ and $\tilde{y}_j(b) = \beta_j^*$. In particular, this means that the related trajectory $\tilde{x}_j^*(\cdot)$ given by $\tilde{x}_j^*(t) = \psi_j^*(t, x_j^*(t))$ is a constant function, namely $\tilde{x}_j^*(t) \equiv \beta_j^*$. That is, the trajectory $\mathbf{x}^*(\cdot)$ (which is our candidate for a Nash equilibrium) is transformed into the constant trajectory $\tilde{\mathbf{x}}^*(t) \equiv \boldsymbol{\beta}^*$.

At present, we only have two thirds of what is required to apply Leitmann's direct sufficiency method, namely the transformation $\mathbf{z}(\cdot,\cdot)$ and the functions $G_j^*(\cdot,\cdot)$. We still need to select the appropriate integrands $\tilde{L}_j(\cdot,\cdot,\cdot) : [a,b] \times \mathbb{R}^{\mathbf{n}} \times \mathbb{R}^{nj} \to \mathbb{R}$. To do this we suppose, as above, that $\mathbf{x}^*(\cdot)$ is embedded in a field $\mathscr{D}$ for the functional (1) relative to itself with slope function $\hat{p}_j(\cdot,\cdot) = \dot{\varphi}_j^*(\cdot, \psi_j^*(\cdot,\cdot))$ and define the Weierstrass E-functions, $E_j^* : [a,b] \times \mathbb{R}^{nj} \times \mathbb{R}_j^n \times \mathbb{R}^{nj}, \to \mathbb{R}$ by the formula

$$E_j^*(t, x_j, p_j, q_j) = L_j^*(t, x_j, q_j) - L_j^*(t, x_j, p_j) - \left(\left. \frac{\partial L_j^*}{\partial p_j} \right|_{(t, x_j, p_j)} \right)^{\mathsf{T}} (q_j - p_j), \tag{19}$$

in which $L_j^*(t, x_j, p_j) = L_j(t, [\mathbf{x}^*(t)^j, x_j], p_j)$ (as in Remark 6). Observe that if we consider the function $p_j \mapsto L_j^*(t, x_j, p_j)$, with (t, x_j) fixed, the Weierstrass E-function, $E_j^*(\cdot,\cdot,\cdot,\cdot)$ represents the difference between $L_j^*(t, x, \cdot)$ evaluated at q_j and the linear part of its Taylor expansion about the point p_j. This means that for each fixed (t, x_j) there exists $\theta_j^*(t, x_j) \in (0,1)$ such that

$$E_j^*(t, x_j, p_j, q_j) = \frac{1}{2}(q_j - p_l)^{\mathsf{T}} \left. \frac{\partial^2 L_j^*}{\partial p_j^2} \right|_{(t, x_j, p_j + \theta_j^*(t, x_j)(q_j - p_j))} (q_j - p_j),$$

in which $\partial^2 L_j^*/\partial p_j^2$ denotes the Hessian matrix of the function $z_j \to L_j^*(t, x_j, z_j)$. In particular, when we have a field $\mathscr{D}$ for the functionals (1) with slope functions $\hat{p}_j^*(t, x_j) = \hat{p}_j(t, [\mathbf{x}^*(t)^j, x_j])$, we have the formula,

$$\begin{aligned}
E_j^*(t, x_j, \hat{p}_j^*(t, x), q_j) &= \frac{1}{2}(q_j - \hat{p}_j^*(t, x_j))^{\mathsf{T}} Q_j^*(t, x_j, q_j)(q_j - \hat{p}_j^*(t, x_j)) \\
&= \frac{1}{2}(q_j - \dot{\varphi}_j^*(t, \psi_j^*(t, x_j)))^{\mathsf{T}} Q_j^*(t, x_j, q_j) \\
&\quad \times (q_j - \dot{\varphi}_j^*(t, \psi_j^*(t, x_j))),
\end{aligned}$$

where

$$\begin{aligned}
Q_j^*(t, x_j, q_j) &= \left. \frac{\partial^2 L_j^*}{\partial p_j^2} \right|_{(t, x_j, \hat{p}_j^*(t, x_j) + \theta_j^*(t, x_j)(q_j - \hat{p}_j^*(t, x_j)))} \\
&= \left. \frac{\partial^2 L_j^*}{\partial p_j^2} \right|_{(t, x_j, \dot{\varphi}_j^*(t, \psi_j^*(t, x_j)) + \theta_j^*(t, x_j)(q_j - \dot{\varphi}_j^*(t, \psi_j^*(t, x_j))))} .
\end{aligned}$$

Remark 9. As a consequence of the definition of $L_j^*(\cdot,\cdot,\cdot)$, one can easily see that in terms of the original integrands $L_j(\cdot,\cdot,\cdot)$ we can write

$$\begin{aligned}
Q_j(t, [\mathbf{x}^*(t)^j, x_j], q_j) &\doteq Q_j^*(t, x_j, q_j) \\
&= \left. \frac{\partial^2 L_j}{\partial p_j^2} \right|_{(*)},
\end{aligned}$$

in which $(*) = (t, [\mathbf{x}^*(t)^j, x_j], \dot{\varphi}_j^*(t, \psi_j^*(t, x_j)) + \theta_j^*(t, x_j)(q_j - \dot{\varphi}_j^*(t, \psi_j^*(t, x_j))))$.

With this notation, we define our integrand $\tilde{L}_j(\cdot,\cdot,\cdot) : [a, b] \times \mathbb{R}^{\mathbf{n}} \times \mathbb{R}^{n_j} \to \mathbb{R}$ by the formula,

$$\begin{aligned}
\tilde{L}_j^*(t, \tilde{x}_j, q_j) &= \tilde{L}_j(t, [\mathbf{x}^*(t)^j, \tilde{x}_j], q_j) \\
&= E_j^*\left(t, \varphi_j^*(t, \tilde{x}_j), \dot{\varphi}_j^*(t, \tilde{x}_j), \dot{\varphi}_j^*(t, \tilde{x}_j) + \frac{\partial \varphi_j^*}{\partial \beta_j}(t, \tilde{x}_j) q_j\right) \\
&= \frac{1}{2} q_j^{\mathsf{T}} \left. \left(\frac{\partial \varphi_j^*}{\partial \beta_j} \right)^{\mathsf{T}} \right|_{(t, \tilde{x}_j)} Q_j^*(t, \varphi_j^*(t, \tilde{x}_j), q_j) \left. \frac{\partial \varphi_j^*}{\partial \beta_j} \right|_{(t, \tilde{x}_j)} q_j,
\end{aligned}$$

for $(t, \tilde{x}_j, q_j) \in [a, b] \times \mathbb{R}^{n_j} \times \mathbb{R}^{n_j}$.

With these notations, we let $\tilde{x}_j \in \mathbb{R}^{n_j}$ be fixed, but otherwise arbitrarily chosen. By Taylor's theorem with remainder we have

$$\begin{aligned}
&L_j^*\left(t, \varphi_j^*(t,\tilde{x}_j), \dot{\varphi}_j^*(t,\tilde{x}_j) + \left.\frac{\partial \varphi_j^*}{\partial \tilde{x}_j}\right|_{(t,\tilde{x}_j)} q_j\right)\\
&\quad = L_j^*(t, \varphi_j^*(t,\tilde{x}_j), \dot{\varphi}_j^*(t,\tilde{x}_j)) + \left(\left.\frac{\partial L_j^*}{\partial p_j}\right|_{(*)}\right)\left(\left.\frac{\partial \varphi_j^*}{\partial \tilde{x}_j}\right|_{(t,\tilde{x}_j)}\right) q_j\\
&\qquad + \frac{1}{2} q_j^{\mathsf{T}} \left(\left.\frac{\partial \varphi_j^*}{\partial \tilde{x}_j}\right|_{(t,\tilde{x}_j)}\right)^{\mathsf{T}} \left.\frac{\partial^2 L_j^*}{\partial p_j^2}\right|_{(**)} \left(\left.\frac{\partial \varphi_j^*}{\partial \tilde{x}_j}\right|_{(t,\tilde{x}_j)}\right) q_j,
\end{aligned}$$

in which $(*) = (t, \varphi_j^*(t,\tilde{x}_j), \dot{\varphi}_j^*(t,\tilde{x}_j))$ and $(**) = \big(t, \varphi_j^*(t,\tilde{x}_j), \dot{\varphi}_j^*(t,\tilde{x}_j) + \theta_j^*(t,\tilde{x}_j) \left.\frac{\partial \varphi^*}{\partial \tilde{x}_j}\right|_{(t,\tilde{x}_j)} q_j\big)$. In particular, when we have an admissible trajectory for the variational game, say $\mathbf{x}(\cdot)$ and the transformed trajectory $\tilde{\mathbf{x}}(\cdot)$, which for each player j are related through the equation $x_j(t) = \varphi_j^*(t, \tilde{x}_j(t))$, $j = 1, 2, \ldots, N$, this last equation becomes,

$$\begin{aligned}
L_j^*(t, x_j(t), \dot{x}_j(t)) = L_j^*(t, x_j(t), \dot{\varphi}_j^*(t, \tilde{x}_j(t))) + \left(\left.\frac{\partial L_j^*}{\partial p_j}\right|_{(*)}\right)\left(\left.\frac{\partial \varphi_j^*}{\partial \tilde{x}_j}\right|_{(t,\tilde{x}(t))}\right)\dot{\tilde{x}}(t)\\
+ \tilde{L}_j^*(t, \tilde{x}(t), \dot{\tilde{x}}(t)),
\end{aligned}$$

or equivalently

$$\begin{aligned}
L_j^*(t, x_j(t), \dot{x}_j(t)) - \tilde{L}_j^*(t, \tilde{x}(t), \dot{\tilde{x}}(t)) &= \frac{\partial G_j^*}{\partial t}(t, \tilde{x}(t)) + \left(\frac{\partial G_j^*}{\partial \tilde{x}_j}(t, \tilde{x}_j(t))\right)\dot{\tilde{x}}(t)\\
&= \frac{\mathrm{d}}{\mathrm{d}t} G_j^*(t, \tilde{x}_j(t)),
\end{aligned}$$

in which this last equality is a consequence that $\mathbf{x}^*(\cdot)$ is assumed to be embedded in a field $\mathscr{D}$ relative to itself. Notice that this is the fundamental identity (6). With these observations, we are now ready present our extension of the Weierstrass sufficiency theorem for variational games.

Theorem 3. *Let $\mathbf{x}^*(\cdot) : [a,b] \to \mathbb{R}^{\mathbf{n}}$ be a twice continuously differentiable admissible trajectory for the N-player variational game with objectives (1) and end conditions (2) and suppose that it satisfies the system of Euler–Lagrange equations (5). Further suppose that there exists a field $\mathscr{D}$ for the functionals (1) relative to $\mathbf{x}^*(\cdot)$ with slope functions $\hat{p}_j(\cdot,\cdot)$ in which $\mathbf{x}^*(\cdot)$ can be embedded. If for each $j = 1, 2, \ldots, N$, one has that the Weierstrass E-functions defined by (19) satisfy*

$$E_j^*(t, x_j, \hat{p}_j^*(t, x_j), q_j) \geq 0 \tag{20}$$

for all $(t, x_j, q_j) \in \mathscr{D}_j \times \mathbb{R}^{n_j}$, *in which* $\hat{p}_j^*(t, x_j) = \hat{p}_j(t, [\mathbf{x}^*(t)^j, x_j])$, *then* $x^*(\cdot)$ *is a strong local open-loop Nash equilibrium for the variational game.*

Remark 10. We note due to the fact that the notion of embedding an admissible trajectory into a field $\mathscr{D}$ can only be ensured in a strong neighborhood of the trajectory $\mathbf{x}^*(\cdot)$, it is clear that we only have a strong local Nash equilibrium. On the other hand, whenever one can show that the field $\mathscr{D}$ is the entire region $[a, b] \times \mathbb{R}^{\mathbf{n}}$ we obtain a global Nash equilibrium.

Proof. We begin by recalling that $\hat{p}_j^*(t, x_j) = \dot{\varphi}_j^*(t, \tilde{x}_j)$, in which x_j and $\tilde{x}_j$ are related through the transformation $x_j = \varphi_j^*(t, \tilde{x}_j)$. This means that the condition (20) may also be written as

$$E_j^*(t, \varphi_j^*(t, \tilde{x}_j), \dot{\varphi}_j^*(t, \tilde{x}_j), q_j) \geq 0,$$

for all $(t, \tilde{x}_j, q_j) \in \mathscr{R}_j \times \mathbb{R}^{n_j}$, in which $\mathscr{R}_j$ is the image of $\mathscr{D}_j$ via the mapping $(t, \tilde{x}_j) \mapsto (t, \psi_j^*(t, x_j))$ In particular, as a result of the definition of $\tilde{L}_j^*(\cdot, \cdot, \cdot)$ we have from (20) that

$$\tilde{L}_j^*(t, \tilde{x}_j, q_j) \doteq E_j^*(t, \varphi_j^*(t, \tilde{x}_j), \dot{\varphi}_j^*(t, \tilde{x}_j), \dot{\varphi}_j^*(t, \tilde{x}_j) + \frac{\partial \varphi_j^*}{\partial \beta_j}(t, \tilde{x}_j) q_j) \geq 0,$$

for all $(t, \tilde{x}_j, q_j) \in \mathscr{R}_j \times \mathbb{R}^{n_j}$. Now, since $\mathbf{x}^*(\cdot)$ is embedded in the field, there exists a unique set of parameters, $\mathbf{fi}^* = (\beta_1^*, \beta_2^*, \ldots, \beta_N^*)$ so that $x_j^*(t) = \varphi_j^*(t, \beta_j^*)$ on $[a, b]$ and moreover by using the transformation $\tilde{x}_j = \tilde{z}_j(t, x_j) \doteq \psi_j^*(t, x_j)$ to define $\tilde{x}_j^*(\cdot)$ we immediately obtain that $\tilde{x}_j^*(t) \equiv \beta_j^*$. Also, as remarked earlier, for any other admissible trajectory for the variational game, say $\mathbf{x}(\cdot)$, we know that its image, obtained via the transformation $\tilde{x}_j(t) = \psi_j^*(t, x_j(t))$, satisfies the end conditions $\tilde{x}_j(a) = \tilde{x}_j(b) = \beta_j^*$ and moreover, the process is reversible. That is, if $\tilde{\mathbf{x}}(\cdot)$ is a piecewise smooth function such that $(t, \tilde{x}_j(t)) \in \mathscr{R}_j$ for all $t \in [a, b]$ and for which $\tilde{x}_j(a) = \tilde{x}_j(b) = \beta_j^*$ for all $j = 1, 2, \ldots, N$, then the function $\mathbf{x}(\cdot) = (x_1(\cdot), \ldots, x_N(\cdot))$ defined through the transformations $x_j(t) = \varphi_j^*(t, \tilde{x}_j(t))$ is an admissible trajectory for the variational game. In addition, we know from above that for each j there exists smooth functions $G_j^*(\cdot, \cdot)$ so that (6) holds for $G_j(\cdot, \cdot) \doteq G_j^*(\cdot, \cdot)$ and $\tilde{L}_j(\cdot, \cdot, \cdot)$ as described above. To complete the proof, we now observe that the functional

$$\tilde{x}_j(\cdot) \mapsto \int_a^b \tilde{L}_j^*(t, \tilde{x}_j(t), \dot{\tilde{x}}_j(t))\, \mathrm{d}t$$

is nonnegative and, since $\tilde{L}_j^*(t, \tilde{x}_j, 0) = 0$ for all $(t, \tilde{x}_j) \in [a, b] \times \mathbb{R}^{n_j}$, it attains it minimum at $\tilde{x}_j^*(t) \equiv \beta_j^*$ over the set of all piecewise smooth trajectories $\tilde{x}_j(\cdot)$ satisfying the end conditions $\tilde{x}_j(a) = \tilde{x}_j(b) = \beta_j^*$. The desired conclusion now follows by appealing to the direct method (Lemma 1) and the observation that $x_j^*(t) = \varphi_j^*(t, \tilde{x}_j^*(t))$. □

6 An Example

This example is taken from the earlier paper [3, Example 6.2] and is presented here to illustrate the above result.We consider a two-player game in which the objective of player $j = 1, 2$ is to maximize the objective functional

$$I_j(x_1(\cdot), x_2(\cdot)) = \int_0^1 \sqrt{\dot{x}_j(t) - x_1(t) - x_2(t)}\,\mathrm{d}t$$

with the fixed end conditions

$$x_j(0) = x_{j0} \quad \text{and} \quad x_j(1) = x_{j1}.$$

We first observe that for each $j = 1, 2$ we have $p_j \to L_j(x_1, x_2, p_j) = \sqrt{p_j - x_1 - x_2}$ is a concave function for fixed (x_1, x_2) since we have

$$\frac{\partial^2 L_j}{\partial p_j^2} = \frac{-1}{4(p_j - x_1 - x_2)^{\frac{3}{2}}} < 0,$$

so that our hypotheses apply. In addition, the Euler–Lagrange equations take the form

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{1}{\sqrt{\dot{x}_j(t) - x_1(t) - x_2(t)}} = -\frac{1}{\sqrt{\dot{x}_j(t) - x_1(t) - x_2(t)}}, \quad j = 1, 2.$$

To solve this pair of coupled equations we observe that if we let $v_j(t) = (\dot{x}_j(t) - x_1(t) - x_2(t))^{-1/2}$ we see that $\dot{v}_j(t) = -v_j(t)$ which gives us that $v_j(t) = A_j \mathrm{e}^{-t}$ for each $j = 1, 2$. Thus, we have

$$\dot{x}_j(t) - x_1(t) - x_2(t) = \alpha_j \mathrm{e}^{2t}, \quad j = 1, 2,$$

where $\alpha_j \geq 0$ is otherwise an arbitrary constant. Adding the two equation together gives us

$$\dot{x}_1(t) + \dot{x}_2(t) - 2(x_1(t) + x_2(t)) = (\alpha_1 + \alpha_2)\mathrm{e}^{2t},$$

which has the general solution

$$x_1(t) + x_2(t) = (\alpha_1 + \alpha_2)t\mathrm{e}^{2t} + C\mathrm{e}^{2t},$$

where C is a constant of integration. From this, it follows that for each $j = 1, 2$ we have

$$\dot{x}_j(t) = (\alpha_j + C)\mathrm{e}^{2t} + (\alpha_1 + \alpha_2)t\mathrm{e}^{2t},$$

giving us

$$x_j(t) = \left[\frac{1}{2}(\alpha_j + C) + \frac{\alpha_1 + \alpha_2}{2}t - \frac{\alpha_1 + \alpha_2}{4}\right] \mathrm{e}^{2t} + \beta_j,$$

in which β_j is an arbitrary constant of integration. Choosing $C = \frac{\alpha_1+\alpha_2}{2}$ we obtain

$$x_j(t) = \frac{1}{2}\left[\alpha_j + (\alpha_1 + \alpha_2)t\right] \mathrm{e}^{2t} + \beta_j.$$

To get a candidate for optimality we need to choose α_j and β_j so that the fixed end conditions are satisfied. It is easy to see that this gives a linear system of equations for the four unknown constants $\alpha_1,\ \alpha_2,\ \beta_1,\ \beta_2$ which is uniquely solvable for any set of boundary conditions. This means we can find $\alpha_1^*,\ \alpha_2^*,\ \beta_1^*, \beta_2^*$ to give us a candidate for the optimal solution given by

$$x_j^*(t) = \frac{1}{2}\left[\alpha_j^* + (\alpha_1^* + \alpha_2^*)t\right] \mathrm{e}^{2t} + \beta_j^*.$$

From this it is easy to see that if we choose

$$\xi_j(t, \beta_j) = \frac{1}{2}\left[\alpha_j^* + (\alpha_1^* + \alpha_2^*)t\right] \mathrm{e}^{2t} + \beta_j$$

we obtain a one-parameter family of extremals for the dynamic game which allows us to conclude that $\mathbf{x}^*(\cdot) = (x_1^*(\cdot), x_2^*(\cdot))$ is an open loop Nash equilibrium for this dynamic game.

7 Conclusion

In this paper, we presented a proof of the Weierstrass sufficiency theorem for a strong local open-loop Nash equilibrium for a class of variational games. We extended the notion of a field from the classical calculus of variations to the class of dynamic games considered here. Our approach is to view an n-parameter family of solutions of the Euler–Lagrange equations (i.e., the trajectories of the field) as a coordinate transformation which defines a one-to-one mapping between the piecewise smooth trajectories satisfying the requisite end conditions (2) and the piecewise smooth trajectories which satisfy a fixed set of periodic end conditions. In this way, we could exploit Leitmann's direct sufficiency method to present our proof. From this, we see that the direct sufficiency method can be viewed as a generalization of Weierstrass's classical result since in general the endpoints of the transformed trajectories need not satisfy constant end point conditions in the direct sufficiency method. The results presented here provide valuable insight into the early work on competitive systems, viewed as an extension of the calculus of variations, and the more familiar presentations of differential games.

Acknowledgements The authors would like to express their sincere thanks and appreciation to the anonymous referee for the careful reading of our paper. The care to detail given by this person should serve as a model to all referees.

References

1. Akhiezer, N.I.: The Calculus of Variations. Translated from the Russian by Aline H. Frink. A Blaisdell Book in the pure and Applied Sciences. Blaisdell Publishing Co. (A Division of Random House, Inc.), New York-London (1962)
2. Carlson, D.A.: An observation on two methods of obtaining solutions to variational problems. Journal of Optimization Theory and Applications **114**, 345–362 (2002)
3. Carlson, D.A.: Fields of extremals and sufficient conditions for a class of variational games. In: P. Bernhard, V. Gaitsgory, O. Pourtalier (eds.) Advances in Dynamic Games and Applications: Anayliatical and Numerical Developments, *Annals of the International Society of Dynamic Games*, vol. 10, pp. 19–36. Birkhäuser Boston (2009)
4. Carlson, D.A., Leitmann, G.: An extension of the coordinate transformation method for open-loop Nash equilibria. Journal of Optimization Theory and its Applications **123**(1), 27–47 (2004)
5. Carlson, D.A., Leitmann., G.: A coordinate transformation method for the extremization of multiple integrals. Journal Optimization Theory and Applications **127**, 523–533 (2005)
6. Carlson, D.A., Leitmann, G.: The direct method for a class of infinite horizon dynamic games. In: C. Deissenberg, R.F. Hartl (eds.) Optimal Control and Dynamic Games, Applications in Finance, Management Science and Economics. Springer, New York, NY (2005)
7. Carlson, D.A., Leitmann, G.: A direct method for open-loop dynamic games for affine control systems. In: A. Haurie, G. Zaccour (eds.) Dynamic Games: Theory and Applications, pp. 37–55. Springer, New York, NY (2005)
8. Carlson, D.A., Leitmann, G.: Fields of extremals and sufficient conditions for the simplest problem of the calculus of variations. Journal Global Optimization **40**, 41–50 (2008)
9. Carlson, D.A., Leitmann, G.: Fields of extremals and sufficient conditions for the simplest problem of the calculus of variations in n-variables. In: G. Buttazzo, A. Frediani (eds.) Variational Analysis and Aerospace Engineering, *Springer Optimization and it Applications*, vol. 33, pp. 75–90. Springer (2009)
10. Dockner, E.J., Leitmann, G.: Coordinate transformation and derivation of open-loop Nash equilibrium. Journal of Optimization Theory and Applications **110**(1), 1–16 (2001)
11. Leitmann, G.: A note on absolute extrema of certain integrals. International Journal of Non-Linear Mechanics **2**, 55–59 (1967)
12. Leitmann, G.: On a class of direct optimization problems. Journal Optimization Theory and Applications **108**(3), 467–481 (2001)
13. Leitmann, G.: A direct method of optimization and its applications to a class of differential games. Dynamics of Continuous, Discrete and Impulsive Systems, Series A **11**, 191–204 (2004)
14. Roos, C.F.: A mathematical theory of competition. American Journal of Mathematics **46**, 163–175 (1925)
15. Roos, C.F.: Dynamical economics. Proceedings of the National Academy of Sciences **13**, 145–150 (1927)
16. Roos, C.F.: A dynamical theory of economic equilibrium. Proceedings of the National Academy of Sciences **13**, 280–285 (1927)
17. Roos, C.F.: A dynamical theory of economics. Journal Political Economy **35**(5), 632–656 (1927)
18. Roos, C.F.: Generalized Lagrange problems in the calculus of variations. Transactions of the American Mathematical Society **30**(2), 360–384 (1928)
19. Roos, C.F.: Generalized Lagrange problems in the calculus of variations. Transactions of the American Mathematical Society **31**(1), 58–70 (1929)

Riemann–Liouville, Caputo, and Sequential Fractional Derivatives in Differential Games

Arkadii Chikrii and Ivan Matychyn

Abstract There exists a number of definitions of the fractional order derivative. The classical one is the definition by Riemann–Liouville [19, 21, 23, 26]. The Riemann–Liouville fractional derivatives have a singularity at zero. That is why differential equations involving these derivatives require initial conditions of special form lacking clear physical interpretation. These shortcomings do not occur with the regularized fractional derivative in the sense of Caputo. Both the Riemann–Liouville and Caputo derivatives possess neither semigroup nor commutative property. That is why so-called sequential derivatives were introduced by Miller and Ross [19]. In this chapter, we treat sequential derivatives of special form [7–9]. Their relation to the Riemann–Liouville and Caputo fractional derivatives and to each other is established. Differential games for the systems with the fractional derivatives of Riemann–Liouville, Caputo, as well as with the sequential derivatives are studied. Representations of such systems' solutions involving the Mittag–Leffler generalized matrix functions [6] are given. The use of asymptotic representations of these functions in the framework of the Method of Resolving Functions [2–5,10,11,17] allows to derive sufficient conditions for solvability of corresponding game problems. These conditions are based on the modified Pontryagin's condition [2]. The results are illustrated on a model example where a dynamic system of order π pursues another system of order e.

1 Introduction

Operations of fractional differentiation and integration go back a long way. It seems likely that the most straightforward way to their definitions is associated with the Abel integral equation and the Cauchy formula for multiple integration of functions. These issues are widely covered in the monograph [26], see also the books [19, 21, 23]. Various kinds of fractional derivatives were generated by the needs of practice.

A. Chikrii (✉) and I. Matychyn
Glushkov Institute of Cybernetics NASU, 40 Glushkov Prsp 03680, MSP Kiev 187, Ukraine
e-mail: chikrii@gmail.com; atychyn@gmail.com

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_4,

Detailed historical review of this subject matter is contained in [26]. The fractional derivatives under study, namely, Riemann–Liouville, Caputo, and Miller–Ross also have their own specifics. The Riemann–Liouville fractional derivative has singularity at the origin. Therefore, the trajectories of corresponding differential system start at infinity, which seems not always justified from the physical point of view. That is why Caputo regularized derivatives appeared, in which this defect was eliminated. This means that the trajectories of corresponding systems do not arrive at infinity at any finite moment of time. However, both the Riemann–Liouville and the Caputo fractional operators do not possess neither semigroup nor commutative properties, which are inherent to the derivatives of integer order. This gap is filled by the Miller–Ross fractional derivatives, which, in particular, make it feasible to reduce the order of a system of differential equations by increasing their number.

Representations of solutions of the abovementioned fractional order linear nonhomogeneous systems play an important role in the mathematical control theory and the theory of dynamic games [2–11, 13, 17, 18, 20, 22, 25]. Some formulas can be found in [19, 23]. In [6], by introducing of the Mittag–Leffler generalized matrix functions, an analog of the Cauchy formula was derived in the case of fractional derivatives Riemann–Liouville and Caputo for order $0 < \alpha < 1$ (without the help of the Laplace transform). Corresponding formulas for derivatives of arbitrary order can be found in [7–9].

It seems likely that research into the game problems for the fractional-order systems go back to the paper [6]. The basic method in these studies is the method of resolving functions [2–11, 17]. This method is sometimes called the method of the inverse Minkowski functionals. Originally, the method was developed to solve the group pursuit problems (see [2, 13, 25]). In [13], it is referred to as the method of the guaranteed position nondeterioration. The method of resolving functions is based on the Pontryagin condition or its modifications. This method is sufficiently universal: it allows exploring the conflict-controlled processes for objects of different inertia as well as of oscillatory and rotary dynamics, the game problems under state and integral constraints and game problems with impulse controls [2, 10, 13, 17, 25]. The method of resolving functions fully substantiates the classical rule of parallel pursuit [2]. The gist of the method consists in constructing special measurable set-valued mappings with closed images and their support functions. These functions integrally characterize the course of the conflict-controlled process [2, 11]. They are referred to as the resolving functions. Instead of the Filippov–Castaing lemma, $\mathscr{L} \times \mathscr{B}$-measurability [11, 12] and resulting superpositional measurability of certain set-valued mappings and their selections is employed to substantiate a measurable selection of the pursuer's control.

2 Fractional Derivatives of Riemann–Liouville and Caputo

Let $\mathbb{R}^m$ be the m-dimensional Euclidean space, $\mathbb{R}_+$ the positive semi-axis. Suppose $n - 1 < \alpha < n$, $n \in \mathbb{N}$, and let the fractional part of α be denoted by $\{\alpha\}$ and its integer part by $[\alpha]$. Thus, $[\alpha] = n - 1$, $\{\alpha\} = \alpha - n + 1$. The Riemann–Liouville fractional derivative of arbitrary order α ($n-1<\alpha<n$, $n\in\mathbb{N}$) is defined as follows [26]:

$$D^{\alpha} f(t)=\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{[\alpha]} D^{\{\alpha\}} f(t)=\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{[\alpha]} \frac{1}{\Gamma(1-\{\alpha\})} \frac{\mathrm{d}}{\mathrm{d}t} \int\limits_0^t \frac{f(\tau)}{(t-\tau)^{\{\alpha\}}} \mathrm{d}\tau$$

$$=\frac{1}{\Gamma(n-\alpha)}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{n} \int\limits_0^t \frac{f(\tau)}{(t-\tau)^{\alpha-n+1}} \mathrm{d}\tau, \tag{1}$$

where the function $f(t)$ has absolutely continuous derivatives up to order $(n-1)$. The following formula is true

$$D^{\alpha} f(t)=\sum_{k=0}^{n-1} \frac{t^{k-\alpha}}{\Gamma(k-\alpha+1)} f^{(k)}(0)+\frac{1}{\Gamma(n-\alpha)} \int\limits_0^t \frac{f^{(n)}(\tau)}{(t-\tau)^{\alpha-n+1}} \mathrm{d}\tau. \tag{2}$$

The Riemann–Liouville fractional derivatives have singularity at zero due to (2). That is why fractional order differential equations (FDEs) in the sense of Riemann–Liouville require initial conditions of special form lacking clear physical interpretation. These shortcomings do not occur with the regularized Caputo derivative. The Caputo derivative of order α ($n-1<\alpha<n$, $n\in\mathbb{N}$) is defined as the integral part of (2):

$$D^{(\alpha)} f(t)={}_0^C\mathbf{D}_t^{\alpha} f(t)=\frac{1}{\Gamma(n-\alpha)} \int\limits_0^t \frac{f^{(n)}(\tau)}{(t-\tau)^{\alpha+1-n}} \mathrm{d}\tau$$

$$=D^{\alpha} f(t)-\sum_{k=0}^{n-1} \frac{t^{k-\alpha}}{\Gamma(k-\alpha+1)} f^{(k)}(0). \tag{3}$$

3 Miller–Ross Sequential Derivatives

Both the Riemann–Liouville and Caputo derivatives possess neither semigroup nor commutative property, i.e. in general,

$$\mathbf{D}^{\alpha+\beta} f(t) \neq \mathbf{D}^{\alpha} \mathbf{D}^{\beta} f(t),\ \mathbf{D}^{\alpha} \mathbf{D}^{\beta} f(t) \neq \mathbf{D}^{\beta} \mathbf{D}^{\alpha} f(t),$$

where $\mathbf{D}^{\alpha}$ stands for the Riemann–Liouville or Caputo fractional differentiation operator of order α. This renders reduction of the order of FDEs impossible.

To reduce the order of FDEs, sequential derivatives were introduced by Miller and Ross [19]. The Miller–Ross sequential derivatives are defined as follows:

$$\mathscr{D}^{\alpha} f(t) = \mathbf{D}^{\alpha_1}\,\mathbf{D}^{\alpha_2}\ldots\mathbf{D}^{\alpha_k}\; f(t),$$

where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ is a multiindex and function $f(t)$ is sufficiently smooth. In general, the operator $\mathbf{D}^{\alpha}$ underlying the sequential Miller–Ross derivative can be either the Riemann–Liouville or Caputo or any other kind of integro-differentiation operator. In particular, in the case of integer α_i it is conventional differentiation operator $\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{\alpha_i}$.

It should be noted that the Riemann–Liouville and Caputo fractional derivatives can be considered as particular cases of sequential derivatives. Indeed, suppose $n - 1 < \alpha < n$ and denote $p = n - \alpha$, then, by definition (1),

$$D^{\alpha} f(t) = \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{n-1} D^{\alpha-n+1} f(t) = \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{n} D^{-p} f(t),$$

$$D^{(\alpha)} f(t) = D^{(\alpha-n+1)} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{n-1} f(t) = D^{-p} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{n} f(t),$$

where D^{-p} is the fractional integration operator of order $p = n - \alpha$:

$$D^{-p} f(t) = \frac{1}{\Gamma(p)} \int\limits_0^t \frac{f(\tau)}{(t-\tau)^{1-p}} \mathrm{d}\tau.$$

Here, we suggest an example of sequential derivatives to establish their relation to the Riemann–Liouville and Caputo derivatives. Let us choose some ν, $n - 1 < \nu < n$, $n \in \mathbb{N}$, and let us study the case when

$$\alpha = (j, \nu - n + 1, n - 1 - j) = (j, \{\nu\}, [\nu] - j), \;\; j = 0, \ldots, n - 1.$$

We introduce the following notation

$$\mathscr{D}_j^{\nu} f(t) = \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j} D^{\{\nu\}} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{[\nu]-j} f(t) = \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j} D^{\nu-n+1} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{n-1-j} f(t),$$

$$\mathscr{D}_j^{(\nu)} f(t) = \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j} D^{(\{\nu\})} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{[\nu]-j} f(t) = \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j} D^{(\nu-n+1)} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{n-1-j} f(t),$$

where $j = 0, 1, \ldots, n - 1$.

The following lemma shows a relationship between the sequential derivatives $\mathscr{D}_j^{\nu}$, $\mathscr{D}_j^{(\nu)}$, and classical derivatives of Riemann–Liouville and Caputo.

Lemma 1. *Let $n-1<\nu<n$, $n\in\mathbb{N}$, and the function $f(t)$ have absolutely continuous derivatives up to the order $(n-1)$. Then the following equalities hold true*

$$\mathscr{D}_0^{(\nu)} f(t) = D^{(\nu)} f(t) = D^\nu f(t) - \sum_{k=0}^{n-1} \frac{t^{k-\nu}}{\Gamma(k-\nu+1)} f^{(k)}(0),$$

$$\mathscr{D}_1^{(\nu)} f(t) = \mathscr{D}_0^\nu f(t) = D^{(\nu)} f(t) + \frac{t^{n-1-\nu}}{\Gamma(n-\nu)} f^{(n-1)}(0)$$

$$= D^\nu f(t) - \sum_{k=0}^{n-2} \frac{t^{k-\nu}}{\Gamma(k-\nu+1)} f^{(k)}(0),$$

$$\cdots\cdots\cdots$$

$$\mathscr{D}_j^{(\nu)} f(t) = \mathscr{D}_{j-1}^\nu f(t) = D^{(\nu)} f(t) + \sum_{k=n-j}^{n-1} \frac{t^{k-\nu}}{\Gamma(k-\nu+1)} f^{(k)}(0)$$

$$= D^\nu f(t) - \sum_{k=0}^{n-1-j} \frac{t^{k-\nu}}{\Gamma(k-\nu+1)} f^{(k)}(0),$$

$$\cdots\cdots\cdots$$

$$\mathscr{D}_{n-1}^{(\nu)} f(t) = \mathscr{D}_{n-2}^\nu f(t) = D^{(\nu)} f(t) + \sum_{k=1}^{n-1} \frac{t^{k-\nu}}{\Gamma(k-\nu+1)} f^{(k)}(0)$$

$$= D^\nu f(t) - \frac{t^{-\nu}}{\Gamma(1-\nu)} f(0),$$

$$\mathscr{D}_{n-1}^\nu f(t) = D^{(\nu)} f(t) + \sum_{k=0}^{n-1} \frac{t^{k-\nu}}{\Gamma(k-\nu+1)} f^{(k)}(0) = D^\nu f(t).$$

Proof. It is evident that

$$\mathscr{D}_0^{(\nu)} f(t) = D^{(\nu-n+1)} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{n-1} f(t) = \frac{1}{\Gamma(n-\nu)} \int_0^t \frac{f^{(n)}(\tau)}{(t-\tau)^{\nu+1-n}} \mathrm{d}\tau,$$

whence, by virtue of (3),

$$\mathscr{D}_0^{(\nu)} f(t) = D^{(\nu)} f(t) = D^\nu f(t) - \sum_{k=0}^{n-1} \frac{t^{k-\nu}}{\Gamma(k-\nu+1)} f^{(k)}(0).$$

Let $j = 1, \ldots, n-1$, then

$$\begin{aligned}
\mathscr{D}_j^{(\nu)} f(t) &= \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^j D^{(\nu-n+1)} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{n-1-j} f(t) = \\
&= \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^j \frac{1}{\Gamma(n-\nu)} \int_0^t \frac{f^{(n-j)}(\tau)}{(t-\tau)^{\nu+1-n}} \mathrm{d}\tau = \\
&= \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j-1} \frac{1}{\Gamma(n-\nu)} \frac{\mathrm{d}}{\mathrm{d}t} \int_0^t \frac{f^{(n-j)}(\tau)}{(t-\tau)^{\nu+1-n}} \mathrm{d}\tau = \\
&= \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j-1} D^{\nu-n+1} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{n-j} f(t) = \mathscr{D}_{j-1}^{\nu} f(t).
\end{aligned}$$

On the other hand, in virtue of (3)

$$\begin{aligned}
\mathscr{D}_j^{(\nu)} f(t) = \mathscr{D}_{j-1}^{\nu} f(t) &= \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^j \frac{1}{\Gamma(n-\nu)} \int_0^t \frac{f^{(n-j)}(\tau)}{(t-\tau)^{\nu+1-n}} \mathrm{d}\tau \\
&= \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^j D^{(\nu-j)} f(t) = \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^j \Bigg[D^{\nu-j} f(t) \\
&\quad - \sum_{k=0}^{n-1-j} \frac{t^{k-\nu+j}}{\Gamma(k-\nu+j+1)} f^{(k)}(0) \Bigg] \\
&= D^{\nu} f(t) - \sum_{k=0}^{n-1-j} \frac{t^{k-\nu}}{\Gamma(k-\nu+1)} f^{(k)}(0).
\end{aligned}$$

The equality

$$\mathscr{D}_{n-1}^{\nu} f(t) = D^{(\nu)} f(t) + \sum_{k=0}^{n-1} \frac{t^{k-\nu}}{\Gamma(k-\nu+1)} f^{(k)}(0) = D^{\nu} f(t)$$

is the direct consequence of (1) and (2). □

Taking into account the equalities $\mathscr{D}_j^{(\nu)} f(t) = \mathscr{D}_{j-1}^{\nu} f(t)$ and setting $\mathscr{D}_n^{(\nu)} f(t) = D^{\nu} f(t)$, one can introduce a common notation

$$\mathfrak{D}_j^{\nu} f(t) = \mathscr{D}_j^{(\nu)} f(t) = \mathscr{D}_{j-1}^{\nu} f(t),$$

where $n-1 < \nu < n$, $j = 0, \ldots, n$. It goes without saying that $\mathfrak{D}_0^{\nu} f(t) = D^{(\nu)} f(t)$ and $\mathfrak{D}_n^{\nu} f(t) = D^{\nu} f(t)$. In the sequel the formula

$$\mathfrak{D}_j^{\nu} f(t) = \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^j D^{(\nu-j)} f(t) \tag{4}$$

will be of use.

The following formulas for the Laplace transforms of the Riemann–Liouville and Caputo fractional derivatives hold true [23]:

$$L\{D^{\alpha} f(t); s\} = s^{\alpha} F(s) - \sum_{k=0}^{n-1} s^k D^{\alpha-k-1} f(t)\big|_{t=0}, \tag{5}$$

$$L\{D^{(\alpha)} f(t); s\} = s^{\alpha} F(s) - \sum_{k=0}^{n-1} s^{\alpha-k-1} f^{(k)}(0). \tag{6}$$

For integer n, the following formula is true:

$$L\{f^{(n)}(t); s\} = s^n F(s) - \sum_{k=0}^{n-1} s^k f^{(n-k-1)}(0). \tag{7}$$

Using (4), (6), (7), one can derive Laplace transform of the derivative $\mathfrak{D}_j^{\nu} f(t)$:

$$\begin{aligned} L\{\mathfrak{D}_j^{\nu} f(t); s\} &= L\left\{\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^j D^{(\nu-j)} f(t); s\right\} \\ &= s^j L\left\{D^{(\nu-j)} f(t); s\right\} - \sum_{k=0}^{j-1} s^k \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j-k-1} D^{(\nu-j)} f(t)\bigg|_{t=0} \\ &= s^j L\left\{D^{(\nu-j)} f(t); s\right\} - \sum_{k=0}^{j-1} s^k \mathfrak{D}_{j-k-1}^{\nu-k-1} f(t)\big|_{t=0} \\ &= s^{\nu} F(s) - \sum_{l=0}^{n-j-1} s^{\nu-l-1} f^{(l)}(0) - \sum_{k=0}^{j-1} s^k \mathfrak{D}_{j-k-1}^{\nu-k-1} f(t)\big|_{t=0}, \end{aligned} \tag{8}$$

where $\sum_{l=0}^{-1} s^{\nu-l-1} f^{(l)}(0) = 0$, $\sum_{k=0}^{-1} s^k \mathfrak{D}_{j-k-1}^{\nu-k-1} f(t)\big|_{t=0} = 0$. Taking into account that $\mathfrak{D}_{n-k-1}^{\nu-k-1} f(t) = D^{\nu-k-1} f(t)$, one can see that (5), (6) are particular cases of (8):

$$L\{\mathfrak{D}_0^{\nu} f(t); s\} = L\{D^{(\nu)} f(t); s\},\ L\{\mathfrak{D}_n^{\nu} f(t); s\} = L\{D^{\nu} f(t); s\}.$$

4 The Mittag–Leffler Generalized Matrix Function

In [6], the Mittag–Leffler generalized matrix function was introduced:

$$E_{\rho}(B; \mu) = \sum_{k=0}^{\infty} \frac{B^k}{\Gamma(k\rho^{-1} + \mu)},$$

where $\rho > 0$, $\mu \in \mathbb{C}$, and B is an arbitrary square matrix of order m.

The Mittag–Leffler generalized matrix function plays important role in studying the linear systems of fractional order. Denote by I the identity matrix of order m. The following lemma allows to find the Laplace transforms of expressions involving the Mittag–Leffler matrix function.

Lemma 2. *Let $\alpha > 0$, $\beta > 0$, and let A be an arbitrary square matrix of order m. Then the following formula is true:*

$$L\left\{t^{\beta-1}E_{\frac{1}{\alpha}}(At^{\alpha};\beta);s\right\} = s^{\alpha-\beta}(s^{\alpha}I - A)^{-1}.$$

Proof. Taking into account definitions of the Mittag–Leffler generalized matrix function, Gamma-function, and substituting $\tau = st$, we obtain:

$$\begin{aligned}
L\left\{t^{\beta-1}E_{\frac{1}{\alpha}}(At^{\alpha};\beta);s\right\} &= \int\limits_0^{\infty} \mathrm{e}^{-st}t^{\beta-1}E_{\frac{1}{\alpha}}(At^{\alpha};\beta)\mathrm{d}t \\
&= \int\limits_0^{\infty} \mathrm{e}^{-st}t^{\beta-1}\sum_{k=0}^{\infty}\frac{A^k t^{\alpha k}}{\Gamma(\alpha k+\beta)}\mathrm{d}t = \sum_{k=0}^{\infty}\frac{A^k}{\Gamma(\alpha k+\beta)}\int\limits_0^{\infty}\mathrm{e}^{-st}t^{\alpha k+\beta-1}\mathrm{d}t \\
&= \sum_{k=0}^{\infty}\frac{A^k}{\Gamma(\alpha k+\beta)s^{\alpha k+\beta}}\int\limits_0^{\infty}\mathrm{e}^{-\tau}\tau^{\alpha k+\beta-1}\mathrm{d}\tau = \sum_{k=0}^{\infty}A^k s^{-(\alpha k+\beta)}.
\end{aligned}$$

Now let us show that $\sum\limits_{k=0}^{\infty}A^k s^{-(\alpha k+\beta)} = s^{\alpha-\beta}(s^{\alpha}I - A)^{-1}$. The last formula is equivalent to the equation

$$\sum_{k=0}^{\infty}A^k s^{-(k+1)\alpha} = (s^{\alpha}I - A)^{-1}. \tag{9}$$

Let us multiply both sides of (9) by $(s^{\alpha}I - A)$ (either on the left or on the right, it makes no difference as these matrices commute). We obtain

$$\sum_{k=0}^{\infty}A^k s^{-(k+1)\alpha}(s^{\alpha}I - A) = \sum_{k=0}^{\infty}A^k s^{-k\alpha} - \sum_{k=0}^{\infty}A^{(k+1)}s^{-(k+1)\alpha} = I.$$

Since the inverse matrix is unique, this completes the proof. □

5 Cauchy Problem for Fractional Order Systems

Suppose $g(t)$ is a measurable and bounded function, then $g(t)\mathrm{e}^{-st} \in L_1(\mathbb{R}_+)$, $s \in \mathbb{C}$. Thus, $g(t)$ is Laplace transformable.

Consider a dynamic system whose evolution is described by the equation:

$$D^{\alpha} z = Az + g, \quad n-1 < \alpha < n, \tag{10}$$

under the initial conditions

$$D^{\alpha-k} z(t)\big|_{t=0} = z_{k1}^{0}, \quad k = 1, \dots, n. \tag{11}$$

Lemma 3. *The trajectory of the system (10), (11) has the form:*

$$z(t) = \sum_{k=1}^{n} t^{\alpha-k} E_{\frac{1}{\alpha}}(At^{\alpha}; \alpha - k + 1) z_{k1}^{0} + \int_{0}^{t} (t-\tau)^{\alpha-1} E_{\frac{1}{\alpha}}(A(t-\tau)^{\alpha}; \alpha) g(\tau) \mathrm{d}\tau. \tag{12}$$

Now consider a dynamic system of fractional order in the sense of Caputo described by the equation:

$$D^{(\alpha)} z = Az + g, \quad n-1 < \alpha < n, \tag{13}$$

under the initial conditions

$$z^{(k)}(0) = z_{k2}^{0}, \quad k = 0, \dots, n-1. \tag{14}$$

Lemma 4. *The trajectory of the system (13), (14) has the form:*

$$z(t) = \sum_{k=0}^{n-1} t^{k} E_{\frac{1}{\alpha}}(At^{\alpha}; k+1) z_{k2}^{0} + \int_{0}^{t} (t-\tau)^{\alpha-1} E_{\frac{1}{\alpha}}(A(t-\tau)^{\alpha}; \alpha) g(\tau) \mathrm{d}\tau. \tag{15}$$

It should be noted that the formulas (12), (15) in some specific cases were obtained in [6, 26], by a different method.

Finally, let us study systems involving sequential derivatives of special form $\mathfrak{D}_j^{\alpha}$. Consider a dynamic system whose evolution is described by the equation:

$$\mathfrak{D}_j^{\alpha} z = Az + g, \; n-1 < \alpha < n, \; j \in \{0, 1, \dots, n\} \tag{16}$$

under the initial conditions

$$\begin{aligned} \mathfrak{D}_{j-k-1}^{\alpha-k-1} z(t)\big|_{t=0} &= \tilde{z}_k^{0}, \quad k = 0, \dots, j-1, \\ z^{(l)}(0) &= z_l^{0}, \quad l = 0, \dots, n-j-1. \end{aligned} \tag{17}$$

Lemma 5. *The trajectory of the system (16), (17) has the form:*

$$z(t) = \sum_{l=0}^{n-j-1} t^l E_{\frac{1}{\alpha}}(At^\alpha; l+1) z_l^0 + \sum_{k=0}^{j-1} t^{\alpha-k-1} E_{\frac{1}{\alpha}}(At^\alpha; \alpha - k)\tilde{z}_k^0$$
$$+ \int_0^t (t-\tau)^{\alpha-1} E_{\frac{1}{\alpha}}(A(t-\tau)^\alpha; \alpha) g(\tau) \mathrm{d}\tau.$$

The proofs of Lemmas 3–5 can be derived by straightforward application of direct and inverse Laplace transforms as well as of Lemma 2.

6 Problem Statement

In this section, a statement for the problem of approaching the terminal set will be given for conflict-controlled processes, the dynamics of which is described using fractional derivatives of Riemann–Liouville, Caputo, and Miller–Ross.

Consider conflict-controlled process whose evolution is defined by the fractional order system

$$\mathbf{D}^\alpha z = Az + \varphi(u, v), \quad n - 1 < \alpha < n. \tag{18}$$

Here $\mathbf{D}^\alpha$, as before, stands for the operator of fractional differentiation in the sense of Riemann–Liouville, Caputo or Miller–Ross. It will be clear from the context which type of the fractional differentiation operator is meant. The state vector z belongs to the m-dimensional real Euclidean space $\mathbb{R}^m$, A is a square matrix of order m, the control block is defined by the jointly continuous function $\varphi(u, v)$, $\varphi : U \times V \to \mathbb{R}^m$, where u and v, $u \in U$, $v \in V$ are control parameters of the first and second players respectively, and the control sets U and V are from the set $K(\mathbb{R}^m)$ of all nonempty compact subsets of $\mathbb{R}^m$.

When $\mathbf{D}^\alpha$ is the operator of fractional differentiation in the sense of Riemann–Liouville, i.e. $\mathbf{D}^\alpha = D^\alpha$, the initial conditions for the process (18) are given in the form (11). In this case denote : $z^0 = (z_{11}^0, \dots, z_{n1}^0)$. When the derivative in (18) is understood in Caputo's sense, $\mathbf{D}^\alpha = D^{(\alpha)}$, the initial conditions are of the form (14) and $z^0 = (z_{02}^0, \dots, z_{n-1\,2}^0)$. For sequential derivatives of special form $\mathbf{D}^\alpha = \mathfrak{D}_j^\alpha$, the initial conditions are given by (17) and $z^0 = (\tilde{z}_0^0, \dots, \tilde{z}_{j-1}^0, z_0^0, \dots, z_{n-j-1}^0)$.

Along with the process dynamics (18) and the initial conditions, a terminal set of cylindrical form is given

$$M^* = M_0 + M, \tag{19}$$

where M_0 is a linear subspace of $\mathbb{R}^m$, $M \in K(L)$, and $L = M_0^\perp$ is the orthogonal complement of the subspace M_0 in $\mathbb{R}^m$.

When the controls of the both players are chosen in the form of Lebesgue measurable functions $u(t)$ and $v(t)$ taking values from U and V, respectively, the Cauchy problem for the process (18) with corresponding initial values has a unique absolutely continuous solution [19, 26].

Consider the following dynamic game. The first player aims to bring a trajectory of the process (18) to the set (19), while the other player strives to delay the moment of hitting the terminal set as much as possible. We assume that the second player's control is an arbitrary measurable function $v(t)$ taking values from V, and the first player at each time instant t, $t \geq 0$, forms his control on the basis of information about z^0 and $v(t)$:

$$u(t) = u(z^0, v(t)), \quad u(t) \in U. \tag{20}$$

Therefore, $u(t)$ is Krasovskii's counter-control [16] prescribed by the O. Hajek stroboscopic strategies [14].

By solving the problem stated above, we employ the Method of Resolving Functions [2, 6]. Usually this method implements the pursuit process in the class of quasistrategies. However in this chapter, we use the results from [11] providing sufficient conditions for the termination of the pursuit in the aforementioned method with the help of counter-controls.

7 Method of Resolving Functions

Denote Π the orthoprojector from $\mathbb{R}^m$ onto L. Set $\varphi(U, v) = \{\varphi(u, v) : u \in U\}$ and consider set-valued mappings

$$W(t, v) = \Pi t^{\alpha-1} E_{\frac{1}{\alpha}}(At^{\alpha}; \alpha)\varphi(U, v), \quad W(t) = \bigcap_{v \in V} W(t, v)$$

defined on the sets $\mathbb{R}_+ \times V$ and $\mathbb{R}_+$, respectively. The condition

$$W(t) \neq \varnothing, \quad t \in \mathbb{R}_+, \tag{21}$$

is usually referred to as Pontryagin's condition. This condition reflects some kind of first player's advantage in resources over the second player. In the case when the condition (21) fails, i.e. for some $t \in \mathbb{R}_+$, $W(t) = \varnothing$, we will use the modified Pontryagin's condition. It consists in rearranging resources in favor of the first player. Namely, at the moments when $W(t) = \varnothing$ the players' control resources are equalized and the resource consumed for this purpose is then subtracted from the terminal set. Formally the procedure is arranged as follows. A measurable bounded with respect to t matrix function $C(t)$ is introduced. Consider set-valued mappings

$$W^*(t, v) = \Pi t^{\alpha-1} E_{\frac{1}{\alpha}}(At^{\alpha}; \alpha)\varphi(U, C(t)v), \quad W^*(t) = \bigcap_{v \in V} W^*(t, v),$$

$$M(t) = M \stackrel{*}{-} \int_0^t \tau^{\alpha-1} \Pi E_{\frac{1}{\alpha}}(A\tau^{\alpha}; \alpha)\varphi^*(\tau, U, V) \mathrm{d}\tau,$$

where $\varphi^*(t,u,v) = \varphi(u,v) - \varphi(u,C(t)v)$ and $X \overset{*}{-} Y = \{z : z + Y \subset X\} = \bigcap_{y \in Y} (X - y)$ is the Minkowski (geometrical) subtraction [24]. By the integral of the set-valued mapping we mean the Aumann integral, i.e. a union of integrals of all possible measurable selections of the given set-valued mapping [15]. Hereafter, we will say that the modified Pontryagin condition is fulfilled whenever a measurable bounded matrix function $C(t)$ exists such that

$$W^*(t) \neq \varnothing \ \forall t \in \mathbb{R}_+, \tag{22}$$

$$M(t) \neq \varnothing \ \forall t \in \mathbb{R}_+. \tag{23}$$

Thus, Pontryagin's condition (21) is replaced with the conditions (22), (23). Obviously, as $C(t) = I$, the condition (23) is fulfilled by default and the condition (22) coincides with the condition (21), since in this case $W^*(t) \equiv W(t)$. It follows that the modified Pontryagin's condition (22), (23) is, generally speaking, a less restrictive assumption than Pontryagin's condition (21).

By virtue of the properties of the process (18) parameters, the set-valued mapping $\varphi(U, C(t)v)$, $v \in V$, is continuous in the Hausdorff metric. Therefore, taking into account the analytical properties of the Mittag–Leffler generalized matrix function, the set-valued mapping $W^*(t,v)$ is measurable with respect to t, $t \in \mathbb{R}_+$, and closed with respect to v, $v \in V$. Then [1] the set-valued mapping $W^*(t)$ is measurable with respect to t and closed-valued. Let $\mathscr{P}(\mathbb{R}^m)$ be the collection of all nonempty closed subsets of space $\mathbb{R}^m$. Then it is evident that

$$W^*(t,v) : \mathbb{R}_+ \times V \to \mathscr{P}(\mathbb{R}^m), \ \ W^*(t) : \mathbb{R}_+ \to \mathscr{P}(\mathbb{R}^m).$$

In this case the measurable with respect to t set-valued mappings $W^*(t,v)$, $W^*(t)$ are said to be normal [15].

It follows from the condition (22) and from the measurable choice theorem [1] that there exists at least one measurable selection $\gamma(\cdot)$ such that $\gamma(t) \in W^*(t)$, $t \in \mathbb{R}_+$. Denote by $\varGamma$ the set of all such selections.

Let us also denote by $g(t,z^0)$ the solution of homogeneous system (18) for $\varphi(u,v) \equiv 0$. Thus, when $\mathbf{D}^\alpha$ is the Riemann–Liouville fractional differentiation operator ($\mathbf{D}^\alpha = D^\alpha$)

$$g(t,z^0) = \sum_{k=1}^{n} t^{\alpha-k} E_{\frac{1}{\alpha}}(At^\alpha; \alpha - k + 1) z_k^0,$$

where

$$z_k^0 = D^{\alpha-k} z(t)\big|_{t=0}, \quad k = 1, \dots, n.$$

For the Caputo regularized fractional derivative ($\mathbf{D}^\alpha = D^{(\alpha)}$), we have

$$g(t,z^0) = \sum_{k=0}^{n-1} t^k E_{\frac{1}{\alpha}}(At^\alpha; k+1) z_k^0,$$

where

$$z_k^0 = z^{(k)}(0), \quad k = 0, \ldots, n-1.$$

And for the sequential derivatives $\mathbf{D}^\alpha = \mathfrak{D}_j^\alpha$, we obtain

$$g(t, z^0) = \sum_{l=0}^{n-j-1} t^l E_{\frac{1}{\alpha}}(At^\alpha; l+1) z_l^0 + \sum_{k=0}^{j-1} t^{\alpha-k-1} E_{\frac{1}{\alpha}}(At^\alpha; \alpha - k) \tilde{z}_k^0,$$

where

$$\tilde{z}_k^0 = \mathfrak{D}_{j-k-1}^{\alpha-k-1} z(t)\big|_{t=0}, \quad k = 0, \ldots, j-1,$$

$$z_l^0 = z^{(l)}(0), \quad l = 0, \ldots, n-j-1.$$

Let us introduce the function

$$\xi(t) = \Pi g(t, z^0) + \int_0^t \gamma(\tau) \mathrm{d}\tau, \quad t \in \mathbb{R}_+,$$

where $\gamma(\cdot) \in \Gamma$ is a certain fixed selection. By virtue of the assumptions made, the selection $\gamma(\cdot)$ is summable.

Consider the set-valued mapping

$$\mathfrak{A}(t, \tau, v) = \{\alpha \geq 0 : [W^*(t-\tau, v) - \gamma(t-\tau)] \cap \alpha[M(t) - \xi(t)] \neq \varnothing\},$$

defined on $\Delta \times V$, where $\Delta = \{(t, \tau) : \ 0 \leq \tau \leq t < \infty\}$. Let us study its support function in the direction of $+1$

$$\alpha(t, \tau, v) = \sup\{\alpha : \ \alpha \in \mathfrak{A}(t, \tau, v)\}, \quad (t, \tau) \in \Delta, \quad v \in V.$$

This function is called the resolving function [2].

Taking into account the modified Pontryagin condition (22), (23), the properties of the conflict-controlled process (18) parameters, as well as characterization and inverse image theorems, one can show that the set-valued mapping $\mathfrak{A}(t, \tau, v)$ is $\mathscr{L} \times \mathscr{B}$-measurable [1] with respect to τ, v, $\tau \in [0, t]$, $v \in V$, and the resolving function $\alpha(t, \tau, v)$ is $\mathscr{L} \times \mathscr{B}$-measurable in τ, v by virtue of the support function theorem [1] when $\xi(t) \notin M(t)$.

It should be noted that for $\xi(t) \in M(t)$ we have $\mathfrak{A}(t, \tau, v) = [0, \infty)$ and hence $\alpha(t, \tau, v) = +\infty$ for any $\tau \in [0, t]$, $v \in V$.

Denote

$$\mathfrak{T} = \left\{ t \in \mathbb{R}_+ : \int_0^t \inf_{v \in V} \alpha(t, \tau, v) \mathrm{d}\tau \geq 1 \right\}. \tag{24}$$

If for some $t > 0$ $\xi(t) \notin M(t)$, we assume the function $\inf\limits_{v \in V} \alpha(t, \tau, v)$ to be measurable with respect to τ, $\tau \in [0, t]$. If it is not the case, then let us define the set $\mathfrak{T}$ as follows

$$\mathfrak{T} = \left\{ t \in \mathbb{R}_+ : \inf_{v(\cdot)} \int_0^t \alpha(t, \tau, v(\tau)) \mathrm{d}\tau \geq 1 \right\}.$$

Since the function $\alpha(t, \tau, v)$ is $\mathscr{L} \times \mathscr{B}$-measurable with respect to τ, v, it is superpositionally measurable [12]. If $\xi(t) \in M(t)$, then $\alpha(t, \tau, v) = +\infty$ for $\tau \in [0, t]$ and in this case it is natural to set the value of the integral in (24) to be equal $+\infty$. Then the inequality in (24) is fulfilled by default. In the case when the inequality in braces in (24) fails for all $t > 0$, we set $\mathfrak{T} = \varnothing$. Let $T \in \mathfrak{T} \neq \varnothing$.

Condition 1. *[11] The set $\mathfrak{A}(T, \tau, v)$ is convex-valued for all $\tau \in [0, T]$, $v \in V$.*

Theorem 1. *Assume for the game problem (18), (19) that there exists a bounded measurable matrix function $C(t)$ such that the conditions (22), (23) hold true and the set M is convex. If there exists a finite number T, $T \in \mathfrak{T} \neq \varnothing$, such that Condition 1 is fulfilled, then the trajectory of the process (18) can be brought to the set (19) from the initial position z^0 at the time instant T using the control of the form (20).*

The proof of this theorem is essentially based on the measurable selection theorem [1] and is similar to that given in [11].

8 Separate Motions

Consider the case when each player's motion is governed by an independent system of fractional differential equations. Let the dynamics of the first player whom we will refer to as *Pursuer* be described by the equation:

$$\mathbf{D}^{\alpha} x = Ax + u, \quad x \in \mathbb{R}^{m_1}, \quad n_1 - 1 < \alpha < n_1. \tag{25}$$

The dynamics of the second player whom we will refer to as *Evader* is governed by the equation:

$$\mathbf{D}^{\beta} y = By + v, \quad y \in \mathbb{R}^{m_2}, \quad n_2 - 1 < \beta < n_2. \tag{26}$$

Here, A and B are square matrices of order m_1 and m_2, respectively, $U \in K(\mathbb{R}^{m_1})$, $V \in K(\mathbb{R}^{m_2})$. It should be noted that (25), (26) are not a particular case of (18) if α and β are distinct. It is supposed that the initial conditions for the systems (25), (26) are given in an appropriate form depending on the operators $\mathbf{D}^{\alpha}$, $\mathbf{D}^{\beta}$, and they are defined by the initial position vectors x^0, y^0 of the Pursuer and Evader, respectively.

The terminal set is defined by ε-distance with respect to the first s ($s \leq \min(m_1, m_2)$) components of the vectors x and y, i.e. the game is terminated whenever

$$\|x - y\|_s \leq \varepsilon. \tag{27}$$

Here, ε is a fixed number, $0 \leq \varepsilon < \infty$.

Let us introduce orthoprojectors $\Pi_1 : \mathbb{R}^{m_1} \to \mathbb{R}^s$, $\Pi_2 : \mathbb{R}^{m_2} \to \mathbb{R}^s$, "cutting" the vectors x, y, leaving only the first s coordinates. Using this notation, the inequality (27) can be rewritten in the form

$$\|\Pi_1 x - \Pi_2 y\| \leq \varepsilon. \tag{28}$$

The situation when the inequality (27)–(28) holds is called capture.

In virtue of Lemmas 3–5, the trajectories of the systems (25), (26) are of the form:

$$x(t) = g_x(x^0, t) + \int_0^t (t-\tau)^{\alpha-1} E_{\frac{1}{\alpha}}(A(t-\tau)^\alpha; \alpha) u(\tau) \mathrm{d}\tau,$$

$$y(t) = g_y(y^0, t) + \int_0^t (t-\tau)^{\beta-1} E_{\frac{1}{\beta}}(B(t-\tau)^\beta; \beta) v(\tau) \mathrm{d}\tau,$$

where $g_x(x^0, t)$ and $g_y(y^0, t)$ are the general solutions to the homogeneous systems $\mathbf{D}^\alpha x = Ax$, $\mathbf{D}^\beta y = By$, respectively, with initial positions x^0, y^0.

In accordance with the method of resolving functions general scheme, let us consider the set-valued mappings

$$W_1(t, v) = t^{\alpha-1} \Pi_1 E_{\frac{1}{\alpha}}(At^\alpha; \alpha) U - t^{\beta-1} \Pi_2 E_{\frac{1}{\beta}}(Bt^\beta; \beta) C_1(t) v;$$

$$W_1(t) = t^{\alpha-1} \Pi_1 E_{\frac{1}{\alpha}}(At^\alpha; \alpha) U \overset{*}{-} t^{\beta-1} \Pi_2 E_{\frac{1}{\beta}}(Bt^\beta; \beta) C_1(t) V. \tag{29}$$

Here, $C_1(t)$ is a matrix function that enables it to equalize the control resources. Along with the set-valued mapping $W_1(t)$, the modified Pontryagin condition involves the mapping:

$$M_1(t) = \varepsilon S \overset{*}{-} \int_0^t \tau^{\beta-1} \Pi_2 E_{\frac{1}{\beta}}(B\tau^\beta; \beta)(C_1(\tau) - I) V \mathrm{d}\tau, \tag{30}$$

where S stands for the unit ball in $\mathbb{R}^s$ centered at the origin.

We say that the modified Pontryagin condition holds if for some measurable bounded matrix function $C_1(t)$ the set-valued mappings defined by (29), (30) are nonempty for all $t \geq 0$.

Let us fix a measurable selection $\gamma_1(t)$ of $W_1(t)$ ($\gamma_1(t) \in W_1(t)\ \forall t \geq 0$) and set

$$\xi_1(t) = \Pi_1 g_x(x^0, t) - \Pi_2 g_y(y^0, t) + \int_0^t \gamma_1(\tau) \mathrm{d}\tau.$$

As before, we define the set-valued mapping

$$\mathfrak{A}_1(t, \tau, v) = \{\alpha \geq 0 : [W_1(t-\tau) - \gamma_1(t-\tau)] \cap \alpha[M_1(t) - \xi(t)] \neq \varnothing\},$$
$$\mathfrak{A}_1 : \Delta \times V \to 2^{\mathbb{R}_+},$$

and its support function in the direction of $+1$ (resolving function)

$$\alpha_1(t, \tau, v) = \sup\{\alpha : \ \alpha \in \mathfrak{A}_1\}, \quad \alpha_1 : \Delta \times V \to \mathbb{R}_+.$$

Now, using the resolving function we define the set

$$\mathfrak{T}_1 = \left\{ t \geq 0 : \int_0^t \inf_{v \in V} \alpha_1(t, \tau, v) \mathrm{d}\tau \geq 1 \right\}.$$

Let $T \in \mathfrak{T}_1 \neq \varnothing$.

Condition 2. *The set $\mathfrak{A}_1(T, \tau, v)$ is convex-valued for all $\tau \in [0, T]$, $v \in V$.*

Theorem 2. *Suppose that for the game problem (25)–(28) with separate dynamics of the players there exists a bounded measurable matrix function $C_1(t)$, $t \geq 0$, such that the set-valued mappings $W_1(t)$ and $M_1(t)$ are nonempty-valued for all $t \geq 0$. If there exists a finite number T, $T \in \mathfrak{T}_1 \neq \varnothing$, such that Condition 2 holds true, then the capture in the game (25)–(28) occurs at the time instant T.*

The proof of this theorem is very similar to that of Theorem 1.

9 Example

Here, we consider an example of fractional order pursuit-evasion dynamic game. Let the dynamics of the Pursuer be described by the equation:

$$\mathbf{D}^{\pi}\, x = u, \quad \|u\| \leq 1, \tag{31}$$

where $\pi = 3,14159\ldots$ is the ratio of a circle's circumference to its diameter.

The dynamics of the Evader is governed by the equation:

$$\mathbf{D}^{e}\, y = v, \quad \|v\| \leq 1, \tag{32}$$

where $e = 2,71828\ldots$ is the base of the natural logarithm.

Here as before $\mathbf{D}^\alpha$ is the operator of fractional differentiation of order α in the sense of Riemann–Liouville, Caputo or Miller–Ross. The phase vectors x and y define the current position in $\mathbb{R}^m$ of the pursuer and the evader, respectively. We suppose that $x = x(t)$ is triple and $y = y(t)$ is twice absolutely continuously differentiable on $\mathbb{R}_+$ functions of time t, i.e. $x(t) \in AC^3(\mathbb{R}_+)$, $y(t) \in AC^2(\mathbb{R}_+)$. The control vectors $u = u(t)$, $v = v(t)$, $u, v \in \mathbb{R}^m$ are measurable functions of time t.

Since A and B are $m \times m$ zero matrices, $E_{\frac{1}{\pi}}(At^\pi; \pi) = \frac{1}{\Gamma(\pi)} I$ and $E_{\frac{1}{e}}(Bt^e; e) = \frac{1}{\Gamma(e)} I$.

If the differentiation is taken in the sense of Riemann–Liouville, i.e. $\mathbf{D}^\alpha = D^\alpha$, the initial conditions for (31), (32) are of the form

$$D^{\pi-1}x(t)\big|_{t=0} = x_{11}^0,\ D^{\pi-2}x(t)\big|_{t=0} = x_{21}^0,\ D^{\pi-3}x(t)\big|_{t=0} = x_{31}^0,\ D^{\pi-4}x(t)\big|_{t=0} = x_{41}^0$$

and

$$D^{e-1}y(t)\big|_{t=0} = y_{11}^0,\ D^{e-2}y(t)\big|_{t=0} = y_{21}^0,\ D^{e-3}y(t)\big|_{t=0} = y_{31}^0,$$

respectively. In this case, we denote

$$\begin{aligned} x^0 &= (x_{11}^0, x_{21}^0, x_{31}^0, x_{41}^0),\ y^0 = (y_{11}^0, y_{21}^0, y_{31}^0), \\ g_x(x^0, t) &= \frac{t^{\pi-1}}{\Gamma(\pi)} x_{11}^0 + \frac{t^{\pi-2}}{\Gamma(\pi-1)} x_{21}^0 + \frac{t^{\pi-3}}{\Gamma(\pi-2)} x_{31}^0 + \frac{t^{\pi-4}}{\Gamma(\pi-3)} x_{41}^0, \\ g_y(y^0, t) &= \frac{t^{e-1}}{\Gamma(e)} y_{11}^0 + \frac{t^{e-2}}{\Gamma(e-1)} y_{21}^0 + \frac{t^{e-3}}{\Gamma(e-2)} y_{31}^0. \end{aligned}$$

Now suppose that $\mathbf{D}^\alpha$ stands for the operator of fractional differentiation in the sense of Caputo, i.e. $\mathbf{D}^\alpha = D^{(\alpha)}$. Then the initial conditions for (31), (32) can be written down the form

$$x(0) = x_{02}^0,\ \dot{x}(0) = x_{12}^0,\ \ddot{x}(0) = x_{22}^0,\ \dddot{x}(0) = x_{32}^0$$

and

$$y(0) = y_{02}^0,\ \dot{y}(0) = y_{12}^0,\ \ddot{y}(0) = y_{22}^0,$$

respectively. We denote

$$\begin{aligned} x^0 &= (x_{02}^0, x_{12}^0, x_{22}^0, x_{32}^0),\ y^0 = (y_{02}^0, y_{12}^0, y_{22}^0), \\ g_x(x^0, t) &= x_{02}^0 + t x_{12}^0 + \frac{t^2}{2} x_{22}^0 + \frac{t^3}{6} x_{32}^0, \\ g_y(y^0, t) &= y_{02}^0 + t y_{12}^0 + \frac{t^2}{2} y_{22}^0. \end{aligned}$$

Finally, if $\mathbf{D}^\pi = \mathfrak{D}_i^\pi$ for some i, $i = 1, 2, 3$, and $\mathbf{D}^e = \mathfrak{D}_j^e$, $j = 1, 2$, the initial conditions for (31), (32) take on the form

$$x^0 = (\tilde{x}_0^0, \ldots, \tilde{x}_{i-1}^0, x_0^0, \ldots, x_{3-i}^0),\ y^0 = (\tilde{y}_0^0, \ldots, \tilde{y}_{j-1}^0, y_0^0, \ldots, y_{2-j}^0),$$

$$g_x(x^0, t) = \sum_{l=0}^{3-i} \frac{t^l}{l!} x_l^0 + \sum_{k=0}^{i-1} \frac{t^{\pi-k-1}}{\Gamma(\pi-k)} \tilde{x}_k^0,$$
$$g_y(y^0, t) = \sum_{s=0}^{2-j} \frac{t^s}{s!} y_s^0 + \sum_{r=0}^{j-1} \frac{t^{e-r-1}}{\Gamma(e-r)} \tilde{y}_r^0.$$

The goal of the pursuer is to achieve the fulfillment of the inequality

$$\|x(T) - y(T)\| \leq \varepsilon, \quad \varepsilon > 0, \tag{33}$$

for some finite time instant T. The goal of the evader is to prevent the fulfillment of the inequality (33) or, provided it is impossible, to maximally postpone the time instant T.

Here $\Pi_1 = \Pi_2 = I$. The set-valued mappings (29) for this example take on the form

$$W_1(t, v) = \frac{t^{\pi-1}}{\Gamma(\pi)} U - \frac{t^{e-1}}{\Gamma(e)} C_1(t) v = \frac{t^{\pi-1}}{\Gamma(\pi)} S - \frac{t^{e-1}}{\Gamma(e)} C_1(t) v,$$

$$W_1(t) = \bigcap_{v \in V} W_1(t, v) = \frac{t^{\pi-1}}{\Gamma(\pi)} S \overset{*}{-} \frac{t^{e-1}}{\Gamma(e)} C_1(t) S,$$

where the matrix function $C_1(t)$ equalizing control resources can be chosen in the form

$$C_1(t) = c_1(t) I, \quad c_1(t) = \begin{cases} \frac{\Gamma(e)}{\Gamma(\pi)} t^{\pi-e} & \text{if } 0 \leq t < \left(\frac{\Gamma(\pi)}{\Gamma(e)}\right)^{\frac{1}{\pi-e}} \\ 1 & \text{if } t \geq \left(\frac{\Gamma(\pi)}{\Gamma(e)}\right)^{\frac{1}{\pi-e}}. \end{cases} \tag{34}$$

Then

$$W_1(t) = \begin{cases} \{0\} & \text{if } 0 \leq t < \left(\frac{\Gamma(\pi)}{\Gamma(e)}\right)^{\frac{1}{\pi-e}} \\ \left(\frac{t^{\pi-1}}{\Gamma(\pi)} - \frac{t^{e-1}}{\Gamma(e)}\right) S & \text{if } t \geq \left(\frac{\Gamma(\pi)}{\Gamma(e)}\right)^{\frac{1}{\pi-e}}. \end{cases}$$

So, $W_1(t) \neq \varnothing$ for all $t \geq 0$. Further,

$$M_1(t) = \varepsilon S \overset{*}{-} \int_0^t \frac{\tau^{e-1}}{\Gamma(e)}(c_1(\tau) - 1)V\mathrm{d}\tau$$

$$= \begin{cases} \left[\varepsilon - \left|\frac{t^\pi}{\Gamma(\pi+1)} - \frac{t^e}{\Gamma(e+1)}\right|\right] S & \text{if } 0 \leq t < \left(\frac{\Gamma(\pi)}{\Gamma(e)}\right)^{\frac{1}{\pi-e}} \\ \left[\varepsilon - \left|\frac{(\Gamma(\pi)/\Gamma(e))^{\frac{\pi}{\pi-e}}}{\Gamma(\pi+1)} - \frac{(\Gamma(\pi)/\Gamma(e))^{\frac{e}{\pi-e}}}{\Gamma(e+1)}\right|\right] S & \text{if } t \geq \left(\frac{\Gamma(\pi)}{\Gamma(e)}\right)^{\frac{1}{\pi-e}}. \end{cases}$$

Thus, the modified Pontryagin condition holds if

$$\varepsilon \geq \left|\frac{(\Gamma(\pi)/\Gamma(e))^{\frac{\pi}{\pi-e}}}{\Gamma(\pi+1)} - \frac{(\Gamma(\pi)/\Gamma(e))^{\frac{e}{\pi-e}}}{\Gamma(e+1)}\right|$$

$$= \frac{(\Gamma(\pi)/\Gamma(e))^{\frac{e}{\pi-e}}}{\Gamma(e+1)} - \frac{(\Gamma(\pi)/\Gamma(e))^{\frac{\pi}{\pi-e}}}{\Gamma(\pi+1)}.$$

Let us choose $\gamma_1(t) \equiv 0$ as a measurable selection of $W_1(t)$. Then

$$\xi_1(t) = \Pi_1 g_x(x^0, t) - \Pi_2 g_y(y^0, t).$$

Denote

$$m(t) = \begin{cases} \varepsilon - \left|\frac{t^\pi}{\Gamma(\pi+1)} - \frac{t^e}{\Gamma(e+1)}\right| & \text{if } 0 \leq t < \left(\frac{\Gamma(\pi)}{\Gamma(e)}\right)^{\frac{1}{\pi-e}} \\ \varepsilon - \left|\frac{(\Gamma(\pi)/\Gamma(e))^{\frac{\pi}{\pi-e}}}{\Gamma(\pi+1)} - \frac{(\Gamma(\pi)/\Gamma(e))^{\frac{e}{\pi-e}}}{\Gamma(e+1)}\right| & \text{if } t \geq \left(\frac{\Gamma(\pi)}{\Gamma(e)}\right)^{\frac{1}{\pi-e}} \end{cases}$$

then

$$M_1(t) = m(t)S.$$

Consider resolving function

$$\alpha_1(t, \tau, v) = \sup\{\alpha \geq 0 : W_1(t - \tau, v) \cap \alpha[M_1(t) - \xi_1(t)] \neq \varnothing\}.$$

This function can be evaluated as the greater root of the quadratic equation:

$$\left\|\alpha\xi_1(t) - \frac{(t-\tau)^{e-1}c_1(t-\tau)}{\Gamma(e)}v\right\| = \frac{(t-\tau)^{\pi-1}}{\Gamma(\pi)} + \alpha m(t).$$

Solving this equation, we find that at $v = -\frac{\xi_1(t)}{\|\xi_1(t)\|}$ the minimum value is attained:

$$\min_{\|v\|\le 1} \alpha_1(t,\tau,v) = \frac{\frac{(t-\tau)^{\pi-1}}{\Gamma(\pi)} - \frac{(t-\tau)^{e-1}c_1(t-\tau)}{\Gamma(e)}}{\|\xi_1(t)\| - m(t)}. \tag{35}$$

In virtue of continuity of the numerator and denominator in (35), time of (31)–(33) game termination can be found as the least positive root of the equation

$$\int_0^t \left[\frac{\tau^{\pi-1}}{\Gamma(\pi)} - \frac{\tau^{e-1}}{\Gamma(e)}c_1(\tau)\right] \mathrm{d}\tau = \|\xi_1(t)\| - m(t). \tag{36}$$

Equation (36) can be simplified taking into account form of the functions $c_1(t)$ and $m(t)$. Indeed, it follows from (34) that for $0 \le t < \left(\frac{\Gamma(\pi)}{\Gamma(e)}\right)^{\frac{1}{\pi-e}}$ the integrand is equal to zero and the game cannot be terminated on this interval.

Suppose $t \ge \left(\frac{\Gamma(\pi)}{\Gamma(e)}\right)^{\frac{1}{\pi-e}}$. Then

$$\int_0^t \left[\frac{\tau^{\pi-1}}{\Gamma(\pi)} - \frac{\tau^{e-1}}{\Gamma(e)}c_1(\tau)\right] \mathrm{d}\tau = \int_{\left(\frac{\Gamma(\pi)}{\Gamma(e)}\right)^{\frac{1}{\pi-e}}}^{t} \left[\frac{\tau^{\pi-1}}{\Gamma(\pi)} - \frac{\tau^{e-1}}{\Gamma(e)}\right] \mathrm{d}\tau$$

and (36) for obtaining the time of game termination takes on its final form:

$$\frac{t^{\pi}}{\Gamma(\pi+1)} - \frac{t^{e}}{\Gamma(e+1)} + \varepsilon = \|\xi_1(t)\|.$$

References

1. Aubin, J.-P., Frankowska, H.: Set-valued analysis. Birkhäuser, Boston (1990)
2. Chikrii, A.A.: Conflict-Controlled Processes. Kluwer, Boston (1997)
3. Chikrii, A.A.: Game problems for systems with fractional derivatives of arbitrary order. In: Haurie, A., Muto, S., Petrosyan, L., Raghavan, T. (eds.) Advances in Dynamic Games, **8**, II, pp. 105–113. Birkhäuser, Boston (2006)
4. Chikrii, A.A.: Game dynamic problems for systems with fractional derivatives. In: Chinchuluun, A., Pardalos, P.M., Migdalas, A., Pitsoulis, L. (eds.) Pareto Optimality, Game Theory and Equilibria, **17**, Springer, New York (2008a)
5. Chikrii, A.A.: Optimization of game interaction of fractional-order controlled systems. J. Optim. Methods Softw. **23**(1), 39–72 (2008b)
6. Chikrii, A.A., Eidelman, S.D.: Generalized Mittag–Leffler matrix functions in game problems for evolutionary equations of fractional order. Cybern. Syst. Anal. **36**(3), 315–338 (2000)

7. Chikrii, A.A., Matychyn, I.I.: On an analogue of the Cauchy formula for linear systems of arbitrary fractional order. Dopovidi Natsional'noi Akademii Nauk Ukrainy. **1**, 50–55 (2007) (in Russian)
8. Chikriy, A.A., Matichin, I.I.: Presentation of solutions of linear systems with fractional derivatives in the sense of Riemann–Liouville, Caputo and Miller–Ross. J. Autom. Inform. Sci. **40**(6), 1–11 (2008)
9. Chikrii, A.A., Matychyn, I.I.: Dynamic games of fractional order in the sense of Riemann–Liouville, Caputo, and Miller–Ross. Proc. Int. Conf. Devoted to the 90-th Anniversary of L.S. Pontryagin. 369-370, Moscow (2008)
10. Chikrii, A.A., Matychyn, I.I., Chikrii, K.A.: Differential games with impulse control. In Jørgensen, S. J., Quincampoix, M., and Vincent, T. L. (eds.) Advances in Dynamic Games, **9**, part I, pp. 37-57. Birkhäuser, Boston (2007)
11. Chikrii, A.A., Rappoport, I.S., Chikrii, K.A.: Multivalued mappings and their selectors in the theory of conflict-controlled processes. Cybern. Syst. Anal. **43**(5), 719–730 (2007)
12. Clarke F.: Optimization and Nonsmooth Analysis. Wiley-Interscience, New York (1983)
13. Grigorenko, N.A.: Mathematical Methods of Control of Multiple Dynamic Processes. MSU Publ., Moscow (1990) (in Russian)
14. Hajek, O.: Pursuit Games. Mathematics in Science and Engineering, **120**. Academic, New York (1975)
15. Ioffe, A.D., Tikhomirov, V.M.: Theory of Extremal Problems. North-Holland, Amsterdam (1979)
16. Krasovskii, N.N., Subbotin, A.I.: Game-Theoretical Control Problems. Springer, New York (1988)
17. Krivonos, Yu.G., Matychyn, I.I., Chikrii, A.A.: Dynamic Games with Discontinuous Trajectories. Naukova Dumka, Kiev (2005) (in Russian)
18. Melikyan, A.A.: Generalized characteristics of first order: applications in optimal control and differential games. Birkhäuser, Boston (1998)
19. Miller, K.S., Ross, B.: An Introduction to the Fractional Calculus and Fractional Differential Equations. Wiley, New York (1993)
20. Mordukhovich, B.S.: Variational analysis and generalized differentiation. Springer, Berlin (2006)
21. Nakhushev, A.M.: Fractional calculus and its applications. Fizmatlit, Moscow (2003) (in Russian)
22. Osipov, Yu.S., Kryazhimskii, A.V.: Inverse problems for ordinary differential equations: dynamical solutions. Gordon & Breach, Basel (1995)
23. Podlubny, I.: Fractional Differential Equations. Academic, San Diego (1999)
24. Pontryagin, L.S.: Selected works, volume 2. Gordon & Breach, Amsterdam (1986)
25. Pshenichnyi, B.N.: Simple pursuit by several objects. Kibernetika, **3**, 145–146 (1976) (in Russian)
26. Samko, S.G., Kilbas, A.A., Marichev, O.I.: Fractional Integrals and Derivatives. Gordon & Breach, Amsterdam (1993)

The Mixed Zero-Sum Stochastic Differential Game in the Model with Jumps

Saïd Hamadène and Hao Wang

Abstract In this chapter, we show that the mixed zero-sum differential-integral game has a value. The main tool is the notion of Backward Stochastic Differential Equations (BSDE for short) with two reflecting right continuous with left limits obstacles (or barriers) when the noise is given by a Brownian motion and a Poisson random measure mutually independent.

1 Introduction

In this chapter, we are concerned with the existence of a value for a mixed zero-sum integral-differential game which is related to the value of callable options and convertible bonds in a financial market. First, let us describe it briefly. Assume that two agents (or players) π_1 and π_2 who intervene in a system, which could be a stock in the market, have antagonistic advantages. The intervention of the agents can take two forms, control and stopping. The dynamics of the system when controlled is given by:

$$x_t = x_0 + \int_0^t f(s, x_s, u_s, v_s)\mathrm{d}s + \int_0^t \int_E \gamma(s, e, x_{s-})\beta(s, e, x_{s-}, u_s, v_s)\lambda(\mathrm{d}e)\mathrm{d}s + \int_0^t \sigma(s, x_s)\mathrm{d}B_s + \int_0^t \int_E \gamma(s, e, x_{s-})\widetilde{\mu}(\mathrm{d}s, \mathrm{d}e), \ \forall\, t \in [0, T],$$

S. Hamadène (✉)
Laboratoire Manceau de Mathématiques, Université du Maine, 72085 Le Mans Cedex 9, France
e-mail: hamadene@univ-lemans.fr

H. Wang
Laboratoire Manceau de Mathématiques, Université du Maine, 72085 Le Mans Cedex 9, France
and
School of Mathematics, Shandong University, Jinan 250100, P. R. China
e-mail: hao.wang.etu@univ-lemans.fr

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_5,

where B is a Brownian motion and $\tilde{\mu}$ is a compensated Poisson measure. The agent π_1 (resp. π_2) controls the system with the help of the process $u = (u_t)_{t\leq T}$ (resp. $v = (v_t)_{t\leq T}$) up to the time when he/she decides to stop controlling at τ (resp. σ) a stopping time. The control of the system is stopped at $\tau \wedge \sigma$, i.e., when one of the agents decides first to stop controlling. As previously said, the advantages of the agents are antagonistic, i.e., there is a payoff $J(u, \tau; v, \sigma)$ which is a cost (resp. a reward) for π_1 (resp. π_2) who aims at minimizing (resp. maximizing) $J(u, \tau; v, \sigma)$. In the particular case of agents who do not have control actions, this mixed game is just the well-known Dynkin game which is widely studied in the literature (see [10] and the references therein). The main objective of this paper is to show that the mixed zero-sum differential-integral game has a value, i.e., the following equality holds true:

$$\inf_{(u,\tau)} \sup_{(v,\sigma)} J(u, \tau; v, \sigma) = \sup_{(v,\sigma)} \inf_{(u,\tau)} J(u, \tau; v, \sigma).$$

When the filtration is Brownian and the process $(x_t)_{t\leq T}$ has no jumps, this mixed zero-sum game is completely solved in [9] in its general setting. However according to our knowledge, the problem of zero-sum mixed differential-integral game is still open. Therefore, our work completes and closes this problem of zero-sum stochastic games of diffusion processes with jumps.

This paper is organized as follows. In Sect. 2, we set up the problem of mixed zero-sum differential-integral game. In Sect. 3, we introduce the notion of BSDE with two reflecting right continuous with left limits (*rcll* for short) barriers which enables us to tackle the problem. The proof of the main theorem of this section, which is challenging and quite long, is postponed to the appendix. In Sect. 4, we give the main result of this paper, i.e., the game has actually a value. This value is expressed by means of a solution of BSDE with two *rcll* reflecting barriers associated with the game. Section 5 is devoted to the proof of the existence and uniqueness of the solutions of the BSDEs involved in the mixed game problem.

2 Setting of the Problem

We begin by fixing the general framework of the game problem and the notation. Throughout this paper, $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\leq T}, P)$ is a stochastic basis such that $\mathscr{F}_0$ contains all P-null sets of $\mathscr{F}$ and $\mathscr{F}_{t+} := \bigcap_{\epsilon>0} \mathscr{F}_{t+\epsilon} = \mathscr{F}_t$, $\forall t < T$. Moreover, we assume that the filtration is generated by the following two mutually independent processes:

- A d-dimensional Brownian motion $(B_t)_{t\leq T}$,
- A Poisson random measure μ on $R^+ \times E$, where $E := R^l \backslash \{0\}$ $(l \geq 1)$ is equipped with its Borel σ-algebra $\mathscr{E}$, with compensator $\nu(\mathrm{d}t, \mathrm{d}e) = \mathrm{d}t\lambda(\mathrm{d}e)$, such that $\widetilde{\mu}([0,t] \times A) = (\mu - \nu)([0,t] \times A)_{t\leq T}$ is a martingale for every $A \in \mathscr{E}$ satisfying $\lambda(A) < \infty$. The measure λ is assumed to be σ-finite on $(E, \mathscr{E})$ and integrates the function $(1 \wedge |e|^2)_{e\in E}$.

Define:

- $\mathscr{P}$ (resp. $\mathscr{P}^d$) the σ-algebra of $\mathscr{F}_t$-progressively measurable (resp. predictable) sets on $[0, T] \times \Omega$.
- $\mathscr{H}^k$ $(k \geq 1)$ the set of $\mathscr{P}$-measurable processes $Z = (Z_t)_{t \leq T}$ with values in R^k s.t. P-$a.s.$, $\int_0^T |Z_s(\omega)|^2 \mathrm{d}s < \infty$; $\mathscr{H}^{2,k}$ is the subset of $\mathscr{H}^k$ of processes Z s.t. $E[\int_0^T |Z_s(\omega)|^2 \mathrm{d}s] < \infty$.
- $\mathscr{S}^2$ the set of $\mathscr{F}_t$-adapted *rcll* processes $Y = (Y_t)_{t \leq T}$ s.t. $E[\sup_{t \leq T} |Y_t|^2] < \infty$.
- $\mathscr{L}$ the set of $\mathscr{P}^d \otimes \mathscr{E}$-measurable mappings $V : \Omega \times [0, T] \times E \to R$ s.t. $P - a.s.$, $\int_0^T \mathrm{d}s \int_E (V_s(\omega, e))^2 \lambda(\mathrm{d}e) < \infty$; $\mathscr{L}^2$ is the subset of $\mathscr{L}$ of processes V s.t. $E[\int_0^T \mathrm{d}s \int_E |V_s(e)|^2 \lambda(\mathrm{d}e)] < \infty$.
- $\mathscr{A}$ the set of $\mathscr{P}^d$-measurable, *rcll* nondecreasing processes $K = (K_t)_{t \leq T}$ s.t. $K_0 = 0$ and P-$a.s.$, $K_T < \infty$; $\mathscr{A}^2$ is the subset of $\mathscr{A}$ of processes K s.t. $E[K_T^2] < \infty$.
- For $\pi = (\pi_t)_{t \leq T} \in \mathscr{S}^2$, $\pi_- := (\pi_{t-})_{t \leq T}$ is the process of its left limits, i.e., $\forall t > 0, \pi_{t-} = lim_{s \nearrow t} \pi_s$ $(\pi_{0-} = \pi_0)$. On the other hand, we denote by $\Delta\pi_t = \pi_t - \pi_{t-}$ the size of the jump of π at t.
- For any stopping time τ, $\mathscr{F}_\tau$ is the σ-algebra of events prior to τ (see [4] for more details).

We now describe the mixed zero-sum stochastic differential-integral game we deal with. Let $x_0 \in R^d$ and let $x = (x_t)_{t \leq T}$ be the solution of the following standard differential equation:

$$x_t = x_0 + \int_0^t \sigma(s, x_s) \mathrm{d}B_s + \int_0^t \int_E \gamma(s, e, x_{s-}) \widetilde{\mu}(\mathrm{d}s, \mathrm{d}e), \forall t \leq T, \tag{1}$$

where the mapping σ: $(t, x) \in [0, T] \times R^d \mapsto \sigma(t, x) \in R^d$ and γ: $(t, e, x) \in [0, T] \times E \times R^d \mapsto \gamma(t, e, x) \in R^d$ satisfy the following assumptions:

(i): There exists a constant C s.t. for any t, x, y we have:

- $$tr[(\sigma(t, x) - \sigma(t, y))(\sigma^*(t, x) - \sigma^*(t, y))] + \int_E |\gamma(t, e, x) - \gamma(t, e, y)|^2 \lambda(\mathrm{d}e) \leq C|x - y|^2$$
- $$tr(\sigma\sigma^*(t, x)) + \int_E |\gamma(t, e, x)|^2 \lambda(\mathrm{d}e) \leq C(1 + |x|^2).$$

(ii): $\forall (t, x) \in [0, T] \times R^d$, the matrix $\sigma(t, x)$ is invertible and $\sigma^{-1}(t, x)$ is bounded.

Therefore, according to Theorem 1.19 in [15], the process $(x_t)_{t \leq T}$ exists and is unique. It stands for the dynamics of a system when noncontrolled.

Let A (*resp.* B) be a compact metric space and $\mathscr{U}$ (*resp.* $\mathscr{V}$) be the space of $\mathscr{P}$-measurable processes $u = (u_t)_{t \leq T}$ (*resp.* $v = (v_t)_{t \leq T}$) with values in A (*resp.* B); $\mathscr{U}$ (*resp.* $\mathscr{V}$) is called the set of admissible controls for π_1 (*resp.* π_2).

Now let f be a function from $([0,T]\times R^d)\times A\times B$ into R^d which is $\mathscr{B}([0,T]\times R^d)\otimes B(A\times B)$-measurable and which satisfies:

(iii): f is bounded, i.e., there exists a constant M s.t. $|f(t,x,u,v)|\le M$ for any (t,x,u,v), and for any (t,x), the mapping $(u,v)\mapsto f(t,x,u,v)$ is continuous.

Let β be a function from $[0,T]\times E\times R^d\times A\times B$ into R which is $\mathscr{B}([0,T]\times E\times R^d)\otimes\mathscr{B}(A\times B)$-measurable and which satisfies:

(iv): For any $(t,e,x)\in[0,T]\times E\times R^d$, the mapping $(u,v)\mapsto\beta(t,e,x,u,v)$ is continuous and for any t,x,e,u,v we have $-1<\beta(t,x,e,u,v)$ and $|\beta(t,x,e,u,v)|\le c_0(1\wedge|e|)$ with c_0 a real constant.

For $(u,v)=(u_t,v_t)_{t\le T}\in\mathscr{U}\times\mathscr{V}$, let $L^{u,v}:=(L_t^{u,v})_{t\le T}$ be the positive martingale solution of:

$$\frac{\mathrm{d}L_t^{u,v}}{L_{t-}^{u,v}}=\sigma^{-1}(t,x_t)f(t,x_t,u_t,v_t)\mathrm{d}B_t+\int_E\beta(t,e,x_{t-},u_t,v_t)\widetilde{\mu}(\mathrm{d}t,\mathrm{d}e)\text{ and }L_0^{u,v}=1.$$

Then the measure $P^{u,v}$ defined by: $\mathrm{d}P^{u,v}=L_T^{u,v}\mathrm{d}P$ is actually a probability ([3], Corollary 5.1, pp. 244) equivalent to P. Moreover, under the new probability $P^{u,v}$, $\mu(\mathrm{d}t,\mathrm{d}e)$ remains a random measure, whose compensator is $\bar{\nu}(\mathrm{d}t,\mathrm{d}e)=(1+\beta(t,e,x_{t-},u_t,v_t))\lambda(\mathrm{d}e)\mathrm{d}t$, i.e. $\widetilde{\mu}^{u,v}([0,t]\times A):=(\mu-\bar{\nu})([0,t]\times A)_{t\le T}$ is a martingale for any $A\in\mathscr{E}$ satisfying $\lambda(A)<\infty$. Additionally, the process $B_t^{u,v}=B_t-\int_0^t\sigma^{-1}(s,x_s)f(s,x_s,u_s,v_s)\mathrm{d}s$ is a Brownian motion and $(x_t)_{t\le T}$ satisfies:

$$\begin{aligned}x_t&=x_0+\int_0^t f(s,x_s,u_s,v_s)\mathrm{d}s+\int_0^t\sigma(s,x_s)\mathrm{d}B_s^{u,v}\\&\quad+\int_0^t\int_E\gamma(s,e,x_{s-})\widetilde{\mu}^{u,v}(\mathrm{d}s,\mathrm{d}e)\\&\quad+\int_0^t\int_E\gamma(s,e,x_{s-})\beta(s,e,x_{s-},u_s,v_s)\lambda(\mathrm{d}e)\mathrm{d}s,\end{aligned}\tag{2}$$

therefore the process $(x_t)_{t\le T}$ is a weak solution of the SDE (2). When the system is controlled, its dynamics has the same law as the one of $(x_t)_{t\le T}$ under the probability $P^{u,v}$. Consequently, comparing (1) and (2), it turns out that the actions of the controllers π_1 and π_2, by choosing the probability $P^{u,v}$, which fixes the law of the controlled system, raise a drift in the dynamics of the system. This formulation of the differential-integral game problem is usually termed as of weak type.

In mixed zero-sum game problems (see e.g. [9, 12]), two agents π_1, π_2 intervene in a system and act with admissible controls u,v respectively. Moreover, they can make the decision to stop controlling at τ for π_1 and σ for π_2, where τ and σ are two stopping times. Therefore, a strategy for π_1 (resp. π_2) is a pair (u, τ) (resp. (v, σ)) and the system is actually stopped at $\tau \wedge \sigma$. Meanwhile, the interventions of the agents will generate a payoff which is a cost for π_1 and a reward for π_2 whose expression is given by:

$$J(u, \tau; v, \sigma) = E^{u,v}\left[\int_0^{\tau\wedge\sigma} h(s, x_s, u_s, v_s)\mathrm{d}s + U_\tau 1_{[\tau<\sigma]} \right.$$
$$\left. + L_\sigma 1_{[\sigma\le\tau<T]} + \xi 1_{[\tau=\sigma=T]}\right],$$

where:

(v): $h : [0, T] \times R^d \times A \times B \mapsto R^+$ is $\mathscr{P} \otimes \mathscr{B}(A \times B)$-measurable function which stands for the instantaneous payoff between the two agents. In addition, the mapping is continuous *w.r.t.* (u, v) and there exists a constant C_h s.t. for any (t, x, u, v), $|h(t, x, u, v)| \le C_h(1 + |x|)$;

(vi): The stopping payoffs $U = (U_t)_{t\le T}$ and $L = (L_t)_{t\le T}$ are processes of $\mathscr{S}^2$ and satisfy $L_t < U_t$ and $L_{t-} < U_{t-}$ $\forall\ t \le T$, i.e., they are *completely separated*;

(vii): ξ is a $\mathscr{F}_T$-measurable random variable s.t. $E[\xi^2] < \infty$ and $L_T \le \xi \le U_T$.

The payoff $J(u, \tau; v, \sigma)$ means that if π_1 (resp. π_2) makes use of control (u, τ) (resp. (v, σ)) and makes the decision to stop the control first before the maturity T then he pays to π_2 (resp. he obtains from π_1) an amount equal to $\int_0^\tau h(s, x_s, u_s, v_s)\mathrm{d}s + U_\tau$ (resp. $\int_0^\sigma h(s, x_s, u_s, v_s)\mathrm{d}s + L_\sigma$). Now if they agree not to stop the control of the system before the maturity T, then π_2 obtains from π_1 the amount $\int_0^T h(s, x_s, u_s, v_s)\mathrm{d}s + \xi$. Note that the difference $U - L$ is a penalty that pays π_1 to π_2 for her decision to stop first the control.

The main objective of this paper is to show that the game has a value, i.e., it holds true that:

$$\operatorname{essinf}_{(u,\tau)}\operatorname{esssup}_{(v,\sigma)} J(u, \tau; v, \sigma) = \operatorname{esssup}_{(v,\sigma)} \operatorname{essinf}_{(u,\tau)} J(u, \tau; v, \sigma).$$

3 Connection with BSDEs with two Reflecting *rcll* Barriers

To tackle this mixed zero-sum stochastic differential game, we are going to use BSDEs with two reflecting barriers which we introduce now.

Actually, besides ξ, L and U, let us consider a function $\Phi : (t, \omega, z, v) \in [0, T] \times \Omega \times R^d \times L^2(E, \mathscr{E}, \lambda; R) \longmapsto \Phi(t, \omega, z, v) \in R$ s.t. $(\Phi(t, \omega, z, v))_{t\le T}$ is $\mathscr{P}$-measurable for any $(z, v) \in R^d \times L^2(E, \mathscr{E}, \lambda; R))$ and $E[\int_0^T |\Phi(t, \omega, 0, 0)|^2\mathrm{d}t] < \infty$. Moreover, we assume that Φ is uniformly Lipschitz with respect to (z, v),

$$\mathrm{P} - a.s., \ |\Phi(t, z, v) - \Phi(t, z', v')| \le C_\Phi(|z - z'| + ||v - v'||), \ \forall\ t, z, z', v \text{ and } v'.$$

A solution for the BSDE, driven by the Brownian motion B and the independent Poisson random measure μ, with two reflecting *rcll* barriers associated with (Φ, ξ, L, U) is a quintuple $(Y, Z, V, K^+, K^-) := (Y_t, Z_t, V_t, K_t^+, K_t^-)_{t \le T}$ of processes with values in $R^{1+d} \times L^2_R(E, \mathscr{E}, \lambda) \times R^{1+1}$ s.t.: $\forall t \le T$,

$$\begin{cases} (i) \;\; Y \in \mathscr{S}^2, K^{\pm} \in \mathscr{A}, Z \in \mathscr{H}^d \text{ and } V \in \mathscr{L}; \\ (ii) \;\; Y_t = \xi + \int_t^T \Phi(s, \omega, Z_s, V_s)\mathrm{d}s + (K_T^+ - K_t^+) - (K_T^- - K_t^-) \\ \qquad\quad - \int_t^T Z_s \mathrm{d}B_s - \int_t^T \int_E V_s(e)\widetilde{\mu}(\mathrm{d}s, \mathrm{d}e); \\ (iii) \;\; L \le Y \le U \text{ and if } K^{c,\pm} \text{ is the continuous part of } K^{\pm} \text{ then} \\ \qquad\quad \int_0^T (Y_t - L_t)\mathrm{d}K_t^{c,+} = \int_0^T (U_t - Y_t)\mathrm{d}K_t^{c,-} = 0; \\ (iv) \;\; \text{if } K^{d,\pm} \text{ is the purely discontinuous part of } K^{\pm} \\ \qquad\quad \text{then } K^{d,\pm} \text{ is predictable and} \\ \qquad\quad K_t^{d,-} = \sum_{0<s\le t} (Y_s - U_{s-})^+ 1_{[\Delta U_s > 0]} \\ \qquad\quad \text{and } K_t^{d,+} = \sum_{0<s\le t} (Y_s - L_{s-})^- 1_{[\Delta L_s < 0]} \; ; \end{cases} \tag{3}$$

here $x^+ = max\{x, 0\}$ and $x^- = -min\{x, 0\}$ for any $x \in R$.

First notice that obviously for arbitrary barriers L and U this equation does not have a solution. However under Mokobodski's condition **[Mk]** which reads as:

$$\textbf{[Mk]:} \begin{cases} \text{there exist two supermartingales of } \mathscr{S}^2, (h_t)_{t \le T} \text{ and } (\theta_t)_{t \le T} \text{ s.t.} \\ \qquad P - a.s., \;\; \forall t \le T, \;\; h_t \ge 0, \;\; \theta_t \ge 0 \text{ and } L_t \le h_t - \theta_t \le U_t, \end{cases} \tag{4}$$

there are several works which establish existence/uniqueness of a solution to (3) (see e.g. [10]). Additionally, it is shown that $E[(K_T^{\pm})^2] < \infty$. Now without assuming **[Mk]** we have the following result related to existence and uniqueness of the solution of (3) whose proof is given in the appendix (Sect. 5, Theorem 7).

Theorem 1. *Under the following assumptions on L, U, ξ and Φ:*

(i) *L and U belong to $\mathscr{S}^2$ and completely separated, i.e., for any $t \le T$, $L_t < U_t$ and $L_{t-} < U_{t-}$,*
(ii) *ξ is $\mathscr{F}_T$-measurable and $L_T \le \xi \le U_T$,*
(iii) *Φ is uniformly Lipschitz in (z, v) uniformly in (t, ω),*

the BSDE with two reflecting barriers (3) associated with (Φ, ξ, L, U) has a unique solution.

4 The Main Result

Let us go back now to the mixed zero-sum differential-integral game problem. First define the Hamilton function associated with this game problem as following: for any $(t, x, z, r, u, v) \in [0, T] \times R^d \times R^d \times L^2_R(E, \mathrm{d}\lambda) \times A \times B$,

$$H(t, x, z, r, u, v) := z\sigma^{-1}(t, x) f(t, x, u, v) + h(t, x, u, v) + \int_E r(e)\beta(t, e, x, u, v)\lambda(\mathrm{d}e).$$

Next assume that Isaacs' condition is fulfilled, i.e., for any (t, x, z, r),

$$\textbf{[IC]} : \inf_{u\in A}\sup_{v\in B} H(t, x, z, r, u, v) = \sup_{v\in B}\inf_{u\in A} H(t, x, z, r, u, v).$$

Under **[IC]**, through the assumptions above on f, β and h and Benes' selection theorem, the following result holds true (see e.g. [6]).

Proposition 1. *There exist two measurable functions $u^*(t, x, z, r)$ and $v^*(t, x, z, r)$ from $[0, T] \times R^d \times R^d \times L^2_R(E, \mathrm{d}\lambda)$ into A and B, respectively, s.t.:*

(i) The pair $(u^*, v^*)(t, x, z, r)$ is a saddle-point for the function H, i.e., for any u, v we have: $H(t, x, z, r, u^*, v) \le H(t, x, z, r, (u^*, v^*)(t, x, z, r)) \le H(t, x, z, r, u, v^*)$.
(ii) The function $(z, r) \mapsto H(t, x, z, r, (u^*, v^*)(t, x, z, r))$ is uniformly Lipschitz.

Next set $H^*(t, x_t(\omega), z, r) = H(t, x_t(\omega), z, r, (u^*, v^*)(t, x_t(\omega), z, r))$ and let $(Y_t, Z_t, R_t, K^\pm_t)$ be the solution associated with (H^*, ξ, L, U), which exists according to Theorem 1. Therefore, we have: $\forall t \in [0, T]$,

$$\begin{cases} (i)\ \ Y_t = \xi + \int_t^T H^*(s, x_s, Z_s, R_s)\mathrm{d}s + (K^+_T - K^+_t) - (K^-_T - K^-_t) \\ \qquad\quad - \int_t^T Z_s \mathrm{d}B_s - \int_t^T\int_E R_s(e)\widetilde{\mu}(\mathrm{d}s, \mathrm{d}e) \\ (ii)\ \ L_t \le Y_t \le U_t,\ \int_0^T (U_s - Y_s)\mathrm{d}K^{c,-}_s = \int_0^T (Y_s - L_s)\mathrm{d}K^{c,+}_s = 0 \\ \qquad\quad \text{where } K^{c,\pm} \text{ is the continuous part of } K^\pm\ (K^{c,\pm}_0 = 0); \\ (iii)\ K^{d,\pm},\ \text{the purely discontinuous part of } K^\pm \text{ is predictable and verifies} \\ \qquad\quad K^{d,+}_t = \sum_{0<s\le t}(L_{s-} - Y_s)^+ \text{ and } K^{d,-}_t = \sum_{0<s\le t}(Y_s - U_{s-})^+; \\ (iv)\ \int_0^T |Z_s|^2\mathrm{d}s + \int_0^T\int_E |R_s(e)|^2\lambda(\mathrm{d}e)\mathrm{d}s < \infty,\ P - a.s.. \end{cases} \tag{5}$$

The following is the main result of this part:

Theorem 2. *Assume that:*

(i) The processes L and U are completely separated, i.e, for any $t \le T$, $L_t < U_t$ and $L_{t-} < U_{t-}$.
*(ii) Isaac's condition **[IC]** is fulfilled.*

Then the zero-sum differential-integral game has a value which is equal to Y_0, i.e., we have:

$$\begin{aligned}&\mathrm{esssup}_{\sigma\in\mathscr{T}_0,v\in\mathscr{V}}\mathrm{essinf}_{\tau\in\mathscr{T}_0,u\in\mathscr{U}}\, J(u,\tau;v,\sigma)\\ &=\mathrm{essinf}_{\tau\in\mathscr{T}_0,u\in\mathscr{U}}\mathrm{esssup}_{\sigma\in\mathscr{T}_0,v\in\mathscr{V}}\, J(u,\tau;v,\sigma)=Y_0.\end{aligned}$$

Proof. First note that Y_0 is a constant since $\mathscr{F}_0$ contains only the P-null sets of $\mathscr{F}$. Now, for any fixed $(u,v)\in\mathscr{U}\times\mathscr{V}$, let $(Y^{u,v},\bar{Z},\bar{R},\bar{K}^{\pm})$ be the solution of the following reflected BSDE:

$$\begin{cases}(i)\;\; Y_t^{u,v}=\xi+\int_t^T H(s,x_s,\bar{Z}_s,\bar{R}_s,u_s,v_s)\mathrm{d}s+(\bar{K}_T^+-\bar{K}_t^+)-(\bar{K}_T^--\bar{K}_t^-)\\ \qquad\qquad -\int_t^T \bar{Z}_s\mathrm{d}B_s-\int_t^T\int_E \bar{R}_s(e)\widetilde{\mu}(\mathrm{d}s,\mathrm{d}e);\\ (ii)\;\; L_t\le Y_t^{u,v}\le U_t \text{ and } \int_0^T(U_s-Y_s^{u,v})\mathrm{d}\bar{K}_s^{c,-}=\int_0^T(Y_s^{u,v}-L_s)\mathrm{d}\bar{K}_s^{c,+}=0\\ \qquad\qquad \text{where } \bar{K}^{c,\pm} \text{ is the continuous part of } \bar{K}^{\pm}\ (\bar{K}_0^{c,\pm}=0);\\ (iii)\; \bar{K}_t^{d,+}=\sum_{0<s\le t}(L_{s-}-Y_s^{u,v})^+ \text{ and } \bar{K}_t^{d,-}=\sum_{0<s\le t}(Y_s^{u,v}-U_{s-})^+;\\ (iv)\;\; \int_0^T|\bar{Z}_s|^2\mathrm{d}s+\int_0^T\int_E|\bar{R}_s(e)|^2\lambda(\mathrm{d}e)\mathrm{d}s<\infty,\ P-a.s..\end{cases}\tag{6}$$

Although there are general jumps in (6), making a change of probability and arguing as in [17], we obtain that $Y_t^{u,v}$ is the value function of the Dynkin game, i.e.,

$$Y_t^{u,v}=\mathrm{esssup}_{\sigma\in\mathscr{T}_t}\mathrm{essinf}_{\tau\in\mathscr{T}_t}\, J_t(u,\tau;v,\sigma)=\mathrm{essinf}_{\tau\in\mathscr{T}_t}\mathrm{esssup}_{\sigma\in\mathscr{T}_t}\, J_t(u,\tau;v,\sigma),$$

where

$$\begin{aligned}J_t(u,\tau;v,\sigma)=E^{u,v}\Bigg[&\int_t^{\tau\wedge\sigma} h(s,x_s,u_s,v_s)\mathrm{d}s+U_\tau 1_{[\tau<\sigma]}\\ &+L_\sigma 1_{[\sigma\le\tau<T]}+\xi 1_{[\tau=\sigma=T]}|\mathscr{F}_t\Bigg].\end{aligned}$$

We now prove that:

$$Y_t=\mathrm{esssup}_{v\in\mathscr{V}}\mathrm{essinf}_{u\in\mathscr{U}}\, Y_t^{u,v}=\mathrm{essinf}_{u\in\mathscr{U}}\mathrm{esssup}_{v\in\mathscr{V}}\, Y_t^{u,v}.\tag{7}$$

However since $\mathrm{esssup}_{v\in\mathscr{V}}\mathrm{essinf}_{u\in\mathscr{U}}\, Y_t^{u,v}\le\mathrm{essinf}_{u\in\mathscr{U}}\mathrm{esssup}_{v\in\mathscr{V}}\, Y_t^{u,v}$, we just need to prove that: $\mathrm{essinf}_{u\in\mathscr{U}}\mathrm{esssup}_{v\in\mathscr{V}}\, Y_t^{u,v}\le Y_t\le\mathrm{esssup}_{v\in\mathscr{V}}\mathrm{essinf}_{u\in\mathscr{U}}\, Y_t^{u,v}$, where Y_t is the solution of (5).

First note that the processes $(u_t^*,v_t^*)=((u^*,v^*)(t,x_t,Z_t,R_t))_{t\le T}$ are admissible controls for π_1 and π_2, respectively, and for any admissible control $(u_t)_{t\le T}$, the generator $H(t,x_t,z,r,u_t,v^*(t,x_t,Z_t,R_t))$ is uniformly Lipschitz *w.r.t.* (z,r). Therefore thanks to Theorem 1 there exists a process Y^{u,v^*} s.t. for any $t\le T$:

$$\begin{aligned}Y_t^{u,v*}=\xi+&\int_t^T H(s,x_s,\tilde{Z}_s,\tilde{R}_s,u_s,v_s^*)\mathrm{d}s+\big(\tilde{K}_T^+-\tilde{K}_t^+\big)-\big(\tilde{K}_T^--\tilde{K}_t^-\big)\\ &-\int_t^T\tilde{Z}_s\mathrm{d}B_s-\int_t^T\int_E\tilde{R}_s(e)\widetilde{\mu}(\mathrm{d}s,\mathrm{d}e);\ \tilde{Z},\tilde{R},\tilde{K}^{\pm} \text{ are as in Theorem 1.}\end{aligned}$$

Next let us define a new probability P^{u,v^*} by $\mathrm{d}P^{u,v^*} = L_T^{u,v^*}\mathrm{d}P$. Using Itô–Meyer's formula ([18], pp. 221) for $(Y - Y^{u,v^*})^{+2}$ and taking into account that

$$\begin{aligned} H^*(s,x_s, Z_s, R_s) - H\left(s, x_s, \tilde{Z}_s, \tilde{R}_s, u_s, v_s^*\right) \\ = H^*(s, x_s, Z_s, R_s) - H(s, x_s, Z_s, R_s, u_s, v_s^*) \\ +(Z_s - \tilde{Z}_s)\sigma^{-1}(s, x_s) f(s, x_s, u_s, v_s^*) \\ + \int_E (R_{s-}(e) - \tilde{R}_{s-}(e))\beta(s, e, x_s, u_s, v_s^*)\lambda(\mathrm{d}e), \end{aligned}$$

we obtain for any $t \in [0, T]$ and stopping time θ,

$$\begin{aligned} \left(Y_{t\wedge\theta} - Y_{t\wedge\theta}^{u,v^*}\right)^{+2} &\le (Y_\theta - Y_\theta^{u,v^*})^{+2} + 2\int_{t\wedge\theta}^{\theta} (Y_s - Y_s^{u,v^*})^+(H^*(s, x_s, Z_s, R_s) \\ &-H(s, x_s, Z_s, R_s, u_s, v_s^*))\mathrm{d}s + 2\int_{t\wedge\theta}^{\theta} (Z_s - \tilde{Z}_s)\mathrm{d}B_s^{u,v^*} \\ &+2\int_t^{\theta}\int_E (R_s(e) - \tilde{R}_s(e))\widetilde{\mu}^{u,v^*}(\mathrm{d}s, \mathrm{d}e), \end{aligned}$$

where under the new probability P^{u,v^*}, the process B^{u,v^*} is a Brownian motion and $\tilde{\mu}^{u,v^*}(\mathrm{d}s, \mathrm{d}e)$ is a martingale measure. Since $H^*(s, x_s, Z_s, R_s) \le H(s, x_s, Z_s, R_s, u_s, v_s^*)$, using localization (see the proof of Proposition 2 below), taking expectation under P^{u,v^*} and then the limit, we obtain P^{u,v^*}-$a.s.$, $Y_t \le Y_t^{u,v^*}$. Therefore, P-$a.s.$ for any $t \le T$, $Y_t \le Y_t^{u,v^*}$ since the two probabilities are equivalent. In the same way, we can show that $Y_t^{u^*,v} \le Y_t$, P-$a.s.$ for any $t \le T$ and any admissible control $(v_t)_{t\le T}$. Therefore, for any $t \le T$ we have $Y_t^{u^*,v} \le Y_t \le Y_t^{u,v^*}$ and then $\operatorname{essinf}_{u\in\mathscr{U}} \operatorname{esssup}_{v\in\mathscr{V}} Y_t^{u,v} \le Y_t \le \operatorname{esssup}_{v\in\mathscr{V}} \operatorname{essinf}_{u\in\mathscr{U}} Y_t^{u,v}$,

which ends the proof of (7).

We now focus on the main claim, that is:

$$\operatorname{essinf}_{u\in\mathscr{U}} Y_t^{u,v} = \operatorname{esssup}_{\sigma\in\mathscr{T}_t} \operatorname{essinf}_{\tau\in\mathscr{T}_t} \operatorname{essinf}_{u\in\mathscr{U}} J_t(u, \tau; v, \sigma),$$

i.e. we can commute the control and the stopping times. So for any u, v and σ, τ let $(J_t, z_t, r_t)_{t\le\tau\wedge\sigma}$ be the solution of the following standard BSDE:

$$J_t = \bar{\xi} + \int_t^{\tau\wedge\sigma} H(s, x_s, z_s, r_s, u_s, v_s)\mathrm{d}s - \int_t^{\tau\wedge\sigma} z_s\mathrm{d}B_s - \int_t^{\tau\wedge\sigma}\int_E r_s(e)\widetilde{\mu}(\mathrm{d}s, \mathrm{d}e),$$

where $\bar{\xi} = U_\tau 1_{[\tau<\sigma]} + L_\sigma 1_{[\sigma\le\tau<T]} + \xi 1_{[\tau=\sigma=T]}$. This solution exists thanks to a result by Li-Tang [14]. Therefore $P - a.s.$, for any $t \le \tau \wedge \sigma$, we have $J_t = J_t(u, \tau; v, \sigma)$.

We can now argue as in [7], Proposition 3.1, to obtain

$$\text{essinf}_{u\in\mathscr{U}} J_t(u,\tau;v,\sigma) = \bar{\xi} + \int_t^{\tau\wedge\sigma} \text{essinf}_{u\in\mathscr{U}} H(s,x_s,z_s,r_s,u_s,v_s)\mathrm{d}s \\ - \int_t^{\tau\wedge\sigma} z_s \mathrm{d}B_s - \int_t^{\tau\wedge\sigma}\int_E r_s(e)\widetilde{\mu}(\mathrm{d}s,\mathrm{d}e).$$

Actually, this is possible since we can use comparison of solutions of those BSDEs thanks to the properties of β and especially the fact that $\beta > -1$. Therefore, the process $(\text{esssup}_{\sigma\in\mathscr{T}_t}\text{essinf}_{\tau\in\mathscr{T}_t}\text{essinf}_{u\in\mathscr{U}} J_t(u,\tau;v,\sigma))_{t\le T}$ is the value function of the corresponding Dynkin game, i.e., the solution of the RBSDE associated with $(\text{essinf}_{u\in\mathscr{U}} H(t,x_t,z,r,u,v),\xi,L,U)$.

On the other hand, using once more the comparison of solutions of BSDEs with two reflecting barriers we obtain that the process $(\text{essinf}_{u\in\mathscr{U}} Y_t^{u,v})_{t\le T}$ is the solution of the RBSDE associated with $(\text{essinf}_{u\in\mathscr{U}} H(t,x_t,z,r,u,v),\xi,L,U)$. Now by uniqueness we obtain for any $t \le T$,

$$\text{essinf}_{u\in\mathscr{U}} Y_t^{u,v} = \text{esssup}_{\sigma\in\mathscr{T}_t}\text{essinf}_{\tau\in\mathscr{T}_t}\text{essinf}_{u\in\mathscr{U}} J_t(u,\tau;v,\sigma).$$

It follows that $\forall t \le T$,

$$\begin{aligned} Y_t &= \text{essup}_{v\in\mathscr{V}}\text{essinf}_{u\in\mathscr{U}} Y_t^{u,v} = \text{essup}_{v\in\mathscr{V}}\text{esssup}_{\sigma\in\mathscr{T}_t}\text{essinf}_{\tau\in\mathscr{T}_t}\text{essinf}_{u\in\mathscr{U}} J_t(u,\tau;v,\sigma) \\ &= \text{esssup}_{\sigma\in\mathscr{T}_t,v\in\mathscr{V}}\text{essinf}_{\tau\in\mathscr{T}_t,u\in\mathscr{U}} J_t(u,\tau;v,\sigma). \end{aligned}$$

In the same way we can show that

$$\text{esssup}_{v\in\mathscr{V}} Y_t^{u,v} = \text{essinf}_{\tau\in\mathscr{T}_t}\text{esssup}_{\sigma\in\mathscr{T}_t}\text{esssup}_{v\in\mathscr{V}} J_t(u,\tau;v,\sigma),$$

which implies that $Y_t = \text{essinf}_{\tau\in\mathscr{T}_t,u\in\mathscr{U}}\text{esssup}_{\sigma\in\mathscr{T}_t,v\in\mathscr{V}} J_t(u,\tau;v,\sigma),\ t \le T$. Thus, the proof of the claim is complete.

Remark 1. Since the process U (resp. L) can have a predictable positive (resp. negative) jump, then it is not granted that the zero-sum game has a saddle-point.

5 Appendix

We now prove Theorem 1. Therefore, throughout this section, we assume that the assumptions $(i)-(iii)$ of this theorem are fulfilled.

5.1 Preliminaries. Reflected BSDEs with One rcll Barrier

We first focus on uniqueness of the solution.

Proposition 2. *The RBSDE (3) has at most one solution, i.e., if (Y, Z, V, K^+, K^-) and (Y', Z', V', K'^+, K'^-) are two solutions of (3), then $Y = Y'$, $Z = Z'$, $V = V'$, $K^+ = K'^+$ and $K^- = K'^-$.*

Proof. For $k \geq 1$ let us set:

$$\tau_k := \inf\left\{t \geq 0, \int_0^t (|Z_s|^2 + |Z_s'|^2)\mathrm{d}s + \int_0^t \int_E (|V_s(e)|^2 + |V_s'(e)|^2)\lambda(\mathrm{d}e)\mathrm{d}s \geq k\right\} \wedge T.$$

Then the sequence $(\tau_k)_{k\geq 0}$ is nondecreasing of stationary type and converges to T. Next using the localization procedure with $(\tau_k)_{k\geq 1}$, arguing as usual and finally taking the limit as $k \to \infty$ we obtain $Y = Y'$, henceforth we get also $Z = Z'$, $V = V'$ and finally $K^+ = K'^+$ and $K^- = K'^-$. The latter equalities require the complete separation of the barriers which is supposed.

Let us now recall the following result by S. Hamadène and Y. Ouknine [13] (see also [8]) related to BSDEs with one reflecting *rcll* barrier. Let $\bar{\Phi}$ be a function $(t, \omega, y, z, v) \in [0, T] \times \Omega \times R^{1+d} \times L^2(E, \mathscr{E}, \lambda; R) \longmapsto \bar{\Phi}(t, \omega, y, z, v) \in R$ s.t. $(\bar{\Phi}(t, \omega, y, z, v))_{t\leq T}$ is $\mathscr{P}$-measurable for any $(z, v) \in R^d \times L^2(E, \mathscr{E}, \lambda; R))$ and $E[\int_0^T |\bar{\Phi}(t, \omega, 0, 0, 0)|^2\mathrm{d}t] < \infty$. Moreover, we assume that $\bar{\Phi}$ is uniformly Lipschitz with respect to (y, z, v).

Theorem 3. *[13]: The BSDE with one reflecting rcll upper barrier associated with $(\bar{\Phi}, \xi, U)$ has a unique solution, i.e., there exists a unique quadruple of processes $(Y_t, Z_t, V_t, K_t)_{t\leq T}$ s.t. $\forall t \leq T$:*

$$\begin{cases} (i)\ \ Y \in \mathscr{S}^2, Z \in \mathscr{H}^{2,d}, V \in \mathscr{L}^2 \text{ and } K \in \mathscr{A}^2 \\ (ii)\ \ Y_t = \xi + \displaystyle\int_t^T \bar{\Phi}(s, Y_s, Z_s, V_s)\mathrm{d}s - (K_T - K_t) - \int_t^T Z_s\mathrm{d}B_s \\ \qquad\qquad - \displaystyle\int_t^T \int_E V_s(e)\tilde{\mu}(\mathrm{d}s, \mathrm{d}e), \\ (iii)\ \ Y_t \leq U_t, \\ (iv)\ \ \text{if } K = K^c + K^d \text{ where } K^c \text{ (resp. } K^d\text{) is the continuous part} \\ \qquad \text{(resp. purely discontinuous part) then } K^d \text{ is predictable,} \\ \qquad \displaystyle\int_0^T (U_t - Y_t)\mathrm{d}K_t^c = 0 \text{ and } \Delta K_t^d = (Y_t - U_{t-})^+ 1_{[\Delta U_t > 0]}. \end{cases} \tag{8}$$

Moreover, the process Y can be characterized as follows: $\forall t \leq T$,

$$Y_t = \operatorname{essinf}_{\tau\geq t} E\left[\int_t^\tau \bar{\Phi}(s, Y_s, Z_s, V_s)\mathrm{d}s + U_\tau 1_{[\tau<T]} + \xi 1_{[\tau=T]}|\mathscr{F}_t\right].$$

Remark 2. The condition of (8)–(iv) is equivalent to $\int_0^T (U_{s-} - Y_{s-})\mathrm{d}K_s = 0$, and K^d can also be written as: $\forall t \leq T$, $K_t^d = \sum_{0<s\leq t} (Y_t - U_{t-})^+ 1_{[\Delta U_t > 0]}$.

Remark 3. We could also give the notion of a solution for a BSDE with a lower reflecting barrier $(L_t)_{t\le T}$. Moreover, the solution Y can also be characterized as

$$Y_t = \mathrm{esssup}_{\tau\ge t} E\left[\int_t^\tau \bar{\Phi}(s,Y_s,Z_s,V_s)\mathrm{d}s + L_\tau 1_{[\tau<T]} + \xi 1_{[\tau=T]}|\mathscr{F}_t\right].$$

We will now provide a comparison result between solutions of one barrier reflected BSDEs which plays an important role in this paper. So assume there exists another quadruple of processes (Y', Z', V', K') solution for the one upper barrier reflected BSDE associated with $(\bar{\Phi}', \xi', U)$. Then we have:

Theorem 4. *Assume that:*

(i) $\bar{\Phi}$ *is independent of* v*,*
(ii) P*-a.s. for any* $t \le T$*,* $\bar{\Phi}(t, Y'_t, Z'_t) \le \bar{\Phi}'(t, Y'_t, Z'_t, V'_t)$ *and* $\xi \le \xi'$*.*

Then P*-a.s.,* $\forall t \le T$*,* $Y_t \le Y'_t$*. Additionally, if* $\bar{\Phi}'$ *does not depend on* v *then we have also* $K_t - K_s \le K'_t - K'_s$*, for any* $0 \le s \le t \le T$*.*

Proof. The first assertion is classic so we focus only on the second one. In the case where $\bar{\Phi}'$ does not depend on v, the solutions of the BSDEs can be constructed via the following penalization schemes. For $n \ge 0$, (Y^n, Z^n, V^n) and (Y'^n, Z'^n, V'^n) are defined as follows:

$$\begin{aligned}Y_t^n = \xi + &\int_t^T \left\{\bar{\Phi}(s, Y_s^n, Z_s^n) - n(Y_s^n - U_s)^+\right\}\mathrm{d}s - \int_t^T Z_s^n \mathrm{d}B_s\\ &- \int_t^T\int_E V_s^n(e)\tilde{\mu}(\mathrm{d}s,\mathrm{d}e)\end{aligned}$$

and

$$\begin{aligned}Y_t'^n = \xi' + &\int_t^T \left\{\bar{\Phi}'(s, Y_s'^n, Z_s'^n) - n(Y_s'^n - U_s)^+\right\}\mathrm{d}s - \int_t^T Z_s'^n \mathrm{d}B_s\\ &- \int_t^T\int_E V_s'^n(e)\tilde{\mu}(\mathrm{d}s,\mathrm{d}e).\end{aligned}$$

First note that through comparison we have $Y^n \le Y'^n$ for any $n \ge 0$. On the other hand, it has been shown in ([8], Theorem 5.1) that the sequences $(Z^n)_{n\ge0}$ and $(V^n)_{n\ge0}$ (resp. $(Z'^n)_{n\ge0}$ and $(V'^n)_{n\ge0}$) converge in $L^p([0,T]\times\Omega, \mathrm{d}t\otimes\mathrm{d}P)$ and $L^p([0,T]\times\Omega\times U, \mathrm{d}t\otimes\mathrm{d}P\otimes\mathrm{d}\lambda)$ to the processes Z and V (resp. Z' and V') for any $p\in[0,2[$ (see also S. Peng [16] in the case of Brownian filtration). Moreover, for any stopping time τ the sequence $(Y_\tau^n)_{n\ge1}$ and $(Y_\tau'^n)_{n\ge1}$ converge decreasingly to Y_τ and Y'_τ P-$a.s.$. Therefore, at least after extracting a subsequence, the sequences $n\int_0^\tau (Y_s^n - U_s)^+\mathrm{d}s$ and $n\int_0^\tau (Y_s'^n - U_s)^+\mathrm{d}s$ converge in $L^p(\mathrm{d}P)$ to K_τ and K'_τ ($p\in[0,2[$). Henceforth for any $s\le t$ we have:

$$K_t - K_s = \lim_{n\to\infty}\int_s^t n(Y_s^n - U_s)^+\mathrm{d}s \le \lim_{n\to\infty}\int_s^t n(Y'^n_s - U_s)^+\mathrm{d}s = K'_t - K'_s$$

since $Y^n \le Y'^n$. The proof is complete.

Remark 4. Using Remark 2, since $Y \leq Y'$ then we obviously have $P-a.s.$, for any $s \leq t$, $K_t^d - K_s^d \leq K_t'^d - K_s'^d$.

5.2 Penalization Schemes and Local Solutions

We are now going to show the existence of a process Y, which satisfies locally the BSDE (3). For $n \geq 1$, let $(Y_t^n, Z_t^n, V_t^n, K_t^n)_{t\leq T}$ be the quadruple of processes s.t.:

$$\begin{cases} (i)\;\; Y^n \in \mathscr{S}^2, Z^n \in \mathscr{H}^{2,d}, V^n \in \mathscr{L}^2 \text{ and } K^n \in \mathscr{A}^2 \\ (ii)\;\; Y_t^n = \xi + \int_t^T \{g(s) + n(L_s - Y_s^n)^+\}\mathrm{d}s - (K_T^n - K_t^n) - \int_t^T Z_s^n \mathrm{d}B_s \\ \qquad\qquad - \int_t^T \int_E V_s^n(e)\tilde{\mu}(\mathrm{d}s, \mathrm{d}e), \\ (iii)\; Y^n \leq U, \\ (iv)\;\; \text{if } K^{n,c} \text{ (resp. } K^{n,d}\text{) is the continuous (resp. purely discontinuous)} \\ \qquad\quad \text{part of } K^n, \text{ then } \int_0^T (U_s - Y_s^n)\mathrm{d}K_s^{n,c} = 0 \text{ and } K^{n,d} \text{ is predictable and} \\ \qquad\quad \text{satisfies } K_t^{n,d} = \sum_{0<s\leq t}(Y_s^n - U_{s-})^+, \forall t \leq T. \end{cases} \tag{9}$$

The existence of the quadruple $(Y^n, Z^n, V^n, K^{n,-})$ is due to Theorem 3. Now the comparison result given in Theorem 4 implies that for any $n \geq 0$ we have $Y^n \leq Y^{n+1} \leq U$. Therefore, there exists a right lower semicontinuous process $Y = (Y_t)_{t\leq T}$ s.t. P-$a.s.$, for any $t \leq T$, $Y_t = \lim_{n\to\infty} Y_t^n$ and $Y_t \leq U_t$. Additionally, the sequence of processes $(Y^n)_{n\geq 0}$ converges to Y in $\mathscr{H}^{2,1}$.

Next for an arbitrary stopping time τ, let us set:

$$\begin{aligned} \delta_\tau^n :&= \inf\left\{s \geq \tau, K_s^n - K_\tau^n > 0\right\} \wedge T \\ &= \inf\left\{s \geq \tau, K_s^{n,d} - K_\tau^{n,d} > 0\right\} \wedge \inf\left\{s \geq \tau, K_s^{n,c} - K_\tau^{n,c} > 0\right\} \wedge T. \end{aligned}$$

Once more from the comparison theorem (4), $K_t^n - K_\tau^n \leq K_t^{n+1} - K_\tau^{n+1}$, therefore $(\delta_\tau^n)_{n\geq 0}$ is a decreasing sequence of stopping times and converges to $\delta_\tau := \lim_{n\to\infty} \delta_\tau^n$, which is also a stopping time. Besides note that for any $t \in [\tau, \delta_\tau[$, $K_t^{n,d} - K_\tau^{n,d} = 0$ for any $n \geq 0$.

The processes Y satisfies:

Proposition 3. *For any stopping time τ, it holds true:*

$$P-a.s.,\;\; 1_{[\delta_\tau<T]} Y_{\delta_\tau} \geq 1_{[\delta_\tau<T]}(U_{\delta_\tau} - 1_{[\delta_\tau>\tau]}(\triangle U_{\delta_\tau})^+).$$

Proof. By definition of δ_τ^n, $K_{\delta_\tau^n}^{n,c} = K_\tau^{n,c}$, hence from (9), we get that: $\forall t \in [\tau, \delta_\tau^n]$,

$$\begin{aligned} Y_t^n = Y_{\delta_\tau^n}^n &+ \int_t^{\delta_\tau^n} \{g(s) + n(L_s - Y_s^n)^+\}\mathrm{d}s - (K_{\delta_\tau^n}^{n,d} - K_t^{n,d}) \\ &- \int_t^{\delta_\tau^n} Z_s^n \mathrm{d}B_s - \int_t^{\delta_\tau^n}\int_E V_s^n(e)\tilde{\mu}(\mathrm{d}s, \mathrm{d}e). \end{aligned} \tag{10}$$

In this equation, the term $K^{n,d}_{\delta^n_\tau} - K^{n,d}_t$ still remains because the process $K^{n,d}$ could have a jump at δ^n_τ. Moreover, we have:

$$\forall t \in [\tau, \delta^n_\tau],\ K^{n,d}_{\delta^n_\tau} - K^{n,d}_t \le 1_{[t<\delta^n_\tau]\cap[Y^n_{\delta^n_\tau -}=U_{\delta^n_\tau -}]}\left(Y^n_{\delta^n_\tau} - U_{\delta^n_\tau -}\right)^+, \tag{11}$$

since the stopping time δ^n_τ could be not predictable. Next for any $n \ge 0$, we have $Y^0 \le Y^n \le U$ therefore there exists a constant C s.t. $E[\sup_{t\le T}|Y^n_t|^2] \le C$. Additionally, standard calculations (see e.g. [13]) imply:

$$\sup_{n\ge 0} E\left[\int_\tau^{\delta^n_\tau} |Z^n_s|^2 \mathrm{d}s\right] + \sup_{n\ge 0} E\left[\int_\tau^{\delta^n_\tau} \mathrm{d}s \int_E |V^n_s(e)|^2 \lambda(\mathrm{d}e)\right] < \infty. \tag{12}$$

Then from (10) and (11) we deduce that:

$$\begin{aligned} Y^n_{\delta_\tau} 1_{[\delta_\tau<T]} \ge\ & E\left[\left\{Y^n_{\delta^n_\tau} - 1_{[\delta_\tau<\delta^n_\tau]}(Y^n_{\delta^n_\tau} - U_{\delta^n_\tau -})^+\right\} 1_{[\delta_\tau<T]}|\mathscr{F}_{\delta_\tau}\right] \\ & -E\left[\int_{\delta_\tau}^{\delta^n_\tau} |g(s)|\mathrm{d}s|\mathscr{F}_{\delta_\tau}\right] \end{aligned} \tag{13}$$

because the random variable $1_{[\delta_\tau<T]}$ belongs to $\mathscr{F}_{\delta_\tau}$.

But on the set $[\delta^n_\tau < T]$ it holds true that $Y^n_{\delta^n_\tau} \ge U_{\delta^n_\tau} - 1_{[\delta^n_\tau>\tau]}(\triangle U_{\delta^n_\tau})^+$. Actually thanks to Remark 2 on the set $[\delta^n_\tau > \tau] \cap [\delta^n_\tau < T]$ we have either $\{Y^n_{\delta^n_\tau -} = U_{\delta^n_\tau -}$ and $Y^n_{\delta^n_\tau} > U_{\delta^n_\tau -}\}$ or $Y^n_{\delta^n_\tau} = U_{\delta^n_\tau}$, hence $Y^n_{\delta^n_\tau} \ge U_{\delta^n_\tau} \wedge U_{\delta^n_\tau -} = U_{\delta^n_\tau} - (\triangle U_{\delta^n_\tau})^+$. Now on $[\delta^n_\tau = \tau] \cap [\delta^n_\tau < T]$, once more thanks to Remark 2, there exists a decreasing sequence of real numbers $(t^n_k)_{k\ge 0}$ converging to τ s.t. $Y^n_{t^n_k -} = U_{t^n_k -}$. Taking the limit as $k \to \infty$ gives $Y^n_\tau \ge U_\tau$ since U and Y^n are *rcll*, whence the claim.

Next going back to (13) to obtain:

$$\begin{aligned} Y^n_{\delta_\tau} 1_{[\delta_\tau<T]} \ge\ & E\left[\xi 1_{[\delta^n_\tau=T]\cap[\delta_\tau<T]}|\mathscr{F}_{\delta_\tau}\right] - E\left[\int_{\delta_\tau}^{\delta^n_\tau} |g(s)|\mathrm{d}s|\mathscr{F}_{\delta_\tau}\right] \\ +E&\left[\{(U_{\delta^n_\tau} - 1_{[\delta^n_\tau>\tau]}(\triangle U_{\delta^n_\tau})^+)1_{[\delta^n_\tau<T]} - 1_{[\delta_\tau<\delta^n_\tau]}(Y^n_{\delta^n_\tau} - U_{\delta^n_\tau -})^+\}1_{[\delta_\tau<T]}|\mathscr{F}_{\delta_\tau}\right]. \end{aligned} \tag{14}$$

We now examine the terms of the right-hand side of (14). First note that in the space $L^1(\mathrm{d}P)$, as $n \to \infty$, $E[\xi 1_{[\delta^n_\tau=T]\cap[\delta_\tau<T]}|\mathscr{F}_{\delta_\tau}] \to 0$ and from (12) we deduce also that $\int_{\delta_\tau}^{\delta^n_\tau} |g(s)|\mathrm{d}s \to 0$ since $\delta^n_\tau \to \delta_\tau$. On the other hand, let us set

$$A = \cap_{n\ge 0}[\delta_\tau < \delta^n_\tau].$$

For n large enough we have: $1_{[\delta_\tau<\delta_\tau^n]}(Y^n_{\delta^n_\tau}-U_{\delta_\tau^n-})^+=1_A(Y^n_{\delta^n_\tau}-U_{\delta_\tau^n-})^+$. Therefore,

$$\limsup_{n\to\infty}1_{[\delta_\tau<\delta_\tau^n]}(Y^n_{\delta^n_\tau}-U_{\delta_\tau^n-})^+\le 1_A\limsup_{n\to\infty}(Y_{\delta^n_\tau}-U_{\delta_\tau^n-})^+=0.$$

Finally

$$\begin{aligned}\lim_{n\to\infty}[U_{\delta^n_\tau}-1_{[\delta^n_\tau>\tau]}(\triangle U_{\delta^n_\tau})^+]&=U_{\delta_\tau}-1_{A^c}\lim_{n\to\infty}1_{[\delta^n_\tau>\tau]}(\triangle U_{\delta^n_\tau})^+\\&\ge U_{\delta_\tau}-1_{[\delta_\tau>\tau]}(\triangle U_{\delta_\tau})^+,\end{aligned}$$

and $1_{[\delta^n_\tau<T]\cap[\delta_\tau<T]}\to 1_{[\delta_\tau<T]}$ as $n\to\infty$. It follows that on $[\delta_\tau<T]$ we have, at least after extracting a subsequence and taking the limit,

$$Y_{\delta_\tau}\ge U_{\delta_\tau}-1_{[\delta_\tau>\tau]}(\triangle U_{\delta_\tau})^+.$$

The proof is now complete.

Proposition 4. *There exists a 4-uplet $(Z',V',K'^+,K'^{d,-})$ which in combination with the process Y satisfies:*

$$\begin{cases}(a)\ Z'\in\mathscr{H}^{2,d},V'\in\mathscr{L}^2,\ K'^+ \text{ and } K'^{d,-}\in\mathscr{A}^2;\\(b)\ Y_t=Y_{\delta_\tau}+\int_t^{\delta_\tau}g(s)\mathrm{d}s-(K'^{d,-}_{\delta_\tau}-K'^{d,-}_t)+(K'^+_{\delta_\tau}-K'^+_t)\\\qquad-\int_t^{\delta_\tau}Z'_s\mathrm{d}B_s-\int_t^{\delta_\tau}\int_E V'_s(e)\tilde\mu(\mathrm{d}s,\mathrm{d}e),\ \forall t\in[\tau,\delta_\tau]\\(c)\ \forall t\in[0,T],\ L_t\le Y_t\le U_t,\\(d)\ K'^+_\tau=0 \text{ and if } K'^{c,+} \text{ (resp. } K'^{d,+}) \text{ is the continuous}\\\qquad\text{(resp. purely discontinuous) part of } K'^+,\ \text{then } K'^{d,+} \text{ is predictable,}\\\qquad K'^{d,+}_t=\sum_{\tau<s\le t}(L_{s-}-Y_s)^+,\ \forall t\in[\tau,\delta_\tau] \text{ and } \int_\tau^{\delta_\tau}(Y_s-L_s)\mathrm{d}K'^{c,+}_s=0,\\(e)\ K'^{d,-} \text{ is predictable and purely discontinuous, } K'^{d,-}_t=0\ \forall t\in[\tau,\delta_\tau[,\\\qquad\text{and if } K'^{d,-}_{\delta_\tau}>0 \text{ then } Y_{\delta_\tau-}=U_{\delta_\tau-} \text{ and } K'^{d,-}_{\delta_\tau}=(Y_{\delta_\tau}-U_{\delta_\tau-})^+.\end{cases}\tag{15}$$

Proof. It will be divided into three steps.

$\boxed{\textit{Step 1:}}$ Construction of the process $K'^{d,-}$.

For $n\ge 0$ and $t\in[0,T]$ let us set $\varDelta^{n,d}_t:=K^{n,d}_{(t\vee\tau)\wedge\delta_\tau}-K^{n,d}_\tau$. The process $\varDelta^{n,d}$ is purely discontinuous and predictable. We just focus on the latter property. Actually for any inaccessible stopping time ζ the process $\varDelta^{n,d}$ cannot jump at ζ since $K^{n,d}$ cannot. On the other hand for any predictable stopping time η, we have $\varDelta^{n,d}_\eta=1_{[\tau<\eta]}K^{n,d}_{\eta\wedge\delta_\tau}-1_{[\tau<\eta]}K^{n,d}_{\eta\wedge\tau}\in\mathscr{F}_{\eta-}$ since a stopped predictable process remains predictable whence the claim (see *e.g.* [2], pp. 5, Prop. 4.5). Now from Remark 4, we get that for any $n \geq 0$, $\Delta_t^{n,d} \leq \Delta_t^{n+1,d}$, $\forall t \leq T$. On the other hand, $\forall t \in [\tau, \delta_\tau[$, $\Delta_t^{n,d} = 0$, and finally for any $t \in [\tau, \delta_\tau]$,

$$\Delta_t^{n,d} \leq 1_{[t<\delta_\tau]\cap Y^n_{\delta_\tau -} = U_{\delta_\tau -}]}(Y^n_{\delta_\tau} - U_{\delta_\tau -})^+.$$

It follows that $(\Delta^{n,d})_{n\geq 0}$ converges to a nondecreasing purely discontinuous predictable *rcll* process $(K_t^{\prime d,-})_{t\leq T}$ which satisfies $K_\tau^{\prime d,-} = 0$ and for any $t \in [\tau, \delta_\tau[$, $K_t^{\prime d,-} = 0$. Suppose now that ω is s.t. $K_{\delta_\tau}^{\prime d,-}(\omega) > 0$ (which implies that we must have $\tau(\omega) < \delta_\tau(\omega)$). Therefore, there exists $n_0(\omega)$ such for any $n \geq n_0$ we have $\Delta_{\delta_\tau}^{n,d}(\omega) > 0$. Using Remark 2, it follows that for any $n \geq n_0$ we have $Y^n_{\delta_\tau -}(\omega) = U_{\delta_\tau -}(\omega)$ and $\Delta_{\delta_\tau}^{n,d}(\omega) = (Y^n_{\delta_\tau} - U_{\delta_\tau -})^+(\omega)$. Consequently, we have also $K_{\delta_\tau}^{\prime d,-}(\omega) = (Y_{\delta_\tau} - U_{\delta_\tau -})^+(\omega)$ and $Y_{\delta_\tau -}(\omega) = U_{\delta_\tau -}(\omega)$ since $Y^n \leq Y \leq U$ and then the left limit of $Y(\omega)$ at $\delta_\tau(\omega)$ exists. We have established the claim (e).

$\boxed{\textit{Step 2:}}$ Y is *rcll* on $[\tau, \delta_\tau]$ and $Y \geq L$.

From (10), since $\delta_\tau \leq \delta_\tau^n$ then we have: $\forall t \in [\tau, \delta_\tau]$,

$$\begin{aligned} Y_t^n = Y^n_{\delta_\tau} &+ \int_t^{\delta_\tau} (g(s) + n(L_s - Y_s^n)^+)\mathrm{d}s - \left(K^{n,d}_{\delta_\tau} - K_t^{n,d}\right) \\ &- \int_t^{\delta_\tau} Z_s^n \mathrm{d}B_s - \int_t^{\delta_\tau}\int_E V_s^n(e)\widetilde{\mu}(\mathrm{d}s, \mathrm{d}e). \end{aligned}$$

So if for $t \in [\tau, \delta_\tau]$ we set $\bar{Y}_t^n = Y_t^n - \Delta_t^{n,d} = Y_t^n - (K_t^{n,d} - K_\tau^{n,d}) + \int_\tau^t g(s)\mathrm{d}s$ then $\bar{Y}^n$ satisfies:

$$\bar{Y}_t^n = \bar{Y}^n_{\delta_\tau} + \int_t^{\delta_\tau} n(L_s - Y_s^n)^+ \mathrm{d}s - \int_t^{\delta_\tau} Z_s^n \mathrm{d}B_s - \int_t^{\delta_\tau}\int_E V_s^n(e)\widetilde{\mu}(\mathrm{d}s, \mathrm{d}e).$$

Writing the latter forwardly, we get that on $[\tau, \delta_\tau]$, $\bar{Y}^n$ is a supermartingale for any n. Next it holds true that $P - a.s.$, $\forall\ t \in [\tau, \delta_\tau]$, $\bar{Y}_t^n \leq \bar{Y}_t^{n+1}$.

Actually, if $\tau = \delta_\tau$, then the claim is obvious since $\bar{Y}_t^n = Y_\tau^n$. Now if $t \in [\tau, \delta_\tau[\cap [\tau < \delta_\tau]$, the claim is also obvious since for any $n \geq 0$,

$$\bar{Y}_t^n = Y_t^n + \int_\tau^t g(s)\mathrm{d}s$$

and we know that $Y^n \leq Y^{n+1}$. Finally, let us consider the case of $t = \delta_\tau(\omega)$ when $\tau(\omega) < \delta_\tau(\omega)$.

First note that $\bar{Y}^n_{\delta_\tau} = Y^n_{\delta_\tau} - (K^{n,d}_{\delta_\tau} - K_\tau^{n,d}) + \int_\tau^{\delta_\tau} g(s)\mathrm{d}s$. So we are going to consider two cases.

$\boxed{\textit{Case 1:}}$ If $K^{n+1,d}_{\delta_\tau}(\omega) - K^{n+1,d}_{\tau}(\omega) = 0$, then thanks to comparison (see Remark 4) we have also $K^{n,d}_{\delta_\tau}(\omega) - K^{n,d}_{\tau}(\omega) = 0$, therefore

$$\bar{Y}^n_{\delta_\tau}(\omega) = Y^n_{\delta_\tau}(\omega) + \int_\tau^{\delta_\tau} g(s)\mathrm{d}s \leq Y^{n+1}_{\delta_\tau}(\omega) + \int_\tau^{\delta_\tau} g(s)\mathrm{d}s = \bar{Y}^{n+1}_{\delta_\tau}(\omega).$$

$\boxed{\textit{Case 2:}}$ If $K^{n+1,d}_{\delta_\tau}(\omega) - K^{n+1,d}_{\tau}(\omega) > 0$, then δ_τ is a stopping time s.t. the pair $(\omega, \delta_\tau(\omega))$ element of the graph of δ_τ, i.e., $[\![\delta_\tau]\!]$, does not belong to the graph $[\![\theta]\!] := \{(\omega, \theta(\omega)), \omega \in \Omega\}$ of any inaccessible stopping time θ. This is due to the fact that the process $K^{n+1,d}$ is predictable and its jumping times are exhausted by a countable set of disjunctive graphs of predictable stopping times (see *e.g.* [5], pp. 128). Next as

$$K^{n+1,d}_{\delta_\tau}(\omega) - K^{n+1,d}_{\tau}(\omega) = (Y^{n+1}_{\delta_\tau} - U_{\delta_\tau -})^+ 1_{[Y^{n+1}_{\delta_\tau -} = U_{\delta_\tau -}]}(\omega)$$

then

$$\bar{Y}^{n+1}_{\delta_\tau}(\omega) = Y^{n+1}_{\delta_\tau -}(\omega) + \int_\tau^{\delta_\tau} g_s(\omega)\mathrm{d}s = U_{\delta_\tau -}(\omega) + \int_\tau^{\delta_\tau} g_s(\omega)\mathrm{d}s.$$

So if $K^{n,d}_{\delta_\tau}(\omega) - K^{n,d}_{\tau}(\omega) > 0$ then it is equal to

$$(Y^n_{\delta_\tau} - U_{\delta_\tau -})^+ 1_{[Y^n_{\delta_\tau -} = U_{\delta_\tau -}]}(\omega)$$

and

$$\bar{Y}^n_{\delta_\tau} = Y^n_{\delta_\tau -} + \int_\tau^{\delta_\tau} g(s)\mathrm{d}s = U_{\delta_\tau -} + \int_\tau^{\delta_\tau} g(s)\mathrm{d}s = \bar{Y}^{n+1}_{\delta_\tau}.$$

Now if $K^{n,d}_{\delta_\tau}(\omega) - K^{n,d}_{\tau}(\omega) = 0$ then $Y^n_{\delta_\tau}(\omega) = Y^n_{\delta_\tau -}(\omega)$ since $\delta_\tau(\omega)$ cannot be equal to $\theta(\omega)$ for any inaccessible stopping time θ, therefore $Y^n(\omega)$ is continuous at $\delta_\tau(\omega)$. It follows that $\bar{Y}^n_{\delta_\tau}(\omega) = Y^n_{\delta_\tau -}(\omega) + \int_\tau^{\delta_\tau} g_s(\omega)\mathrm{d}s \leq U_{\delta_\tau -}(\omega) + \int_\tau^{\delta_\tau} g_s(\omega)\mathrm{d}s = \bar{Y}^{n+1}_{\delta_\tau}(\omega)$. Therefore the sequence $(\bar{Y}^n)$ is non-decreasing.

Now for any $t \in [\tau, \delta_\tau]$, let us set $\bar{Y}_t = \lim_{n\to\infty} \nearrow \bar{Y}^n_t$. As $\bar{Y}^n$ is a supermartingale then $\bar{Y}$ is also a *rcll* supermartingale on $[\tau, \delta_\tau]$ (see e.g. [5], pp. 86). But from the definition of $\bar{Y}^n$ we obtain that $\bar{Y}_t = Y_t - K'^d_t + \int_\tau^t g_s \mathrm{d}s$ and since $K'^{d,-}$ is *rcll* then so is Y.

We now focus on the second property. We know that,

$$Y^n_\tau = Y^n_{\delta_\tau} + \int_\tau^{\delta_\tau} g(s)\mathrm{d}s + \int_\tau^{\delta_\tau} n(L_s - Y^n_s)^+\mathrm{d}s - \left(K^{n,d}_{\delta_\tau} - K^{n,d}_t\right)$$
$$- \int_\tau^{\delta_\tau} Z^n_s \mathrm{d}B_s - \int_\tau^{\delta_\tau}\int_E V^n_s(e)\tilde{\mu}(\mathrm{d}s, \mathrm{d}e).$$

After taking expectation dividing by n and letting $n \to \infty$, we get

$$E[\int_\tau^{\delta_\tau} (L_s - Y_s^n)^+ \mathrm{d}s] \to 0$$

since the other terms in both hand-sides are bounded by Cn^{-1}. Therefore when $\tau(\omega) < \delta_\tau(\omega)$,$\forall t \in [\tau(\omega), \delta_\tau(\omega)[$, $Y_t(\omega) \geq L_t(\omega)$ since Y is *rcll* on $[\tau, \delta_\tau]$. Finally, let us consider the case where $\tau(\omega) = \delta_\tau(\omega)$. From the previous proposition we have:

$$\begin{aligned}
1_{[\tau=\delta_\tau]}Y_\tau &= 1_{[\tau=\delta_\tau]\cap[\delta_\tau<T]}Y_{\delta_\tau} + 1_{[\tau=\delta_\tau]\cap[\delta_\tau=T]}Y_T \\
&\geq 1_{[\tau=\delta_\tau]\cap[\delta_\tau<T]}(U_{\delta_\tau} - 1_{[\delta_\tau>\tau]}(\triangle U_{\delta_\tau})^+) + 1_{[\tau=\delta_\tau]\cap[\delta_\tau=T]}\xi \\
&\geq 1_{[\tau=\delta_\tau]\cap[\delta_\tau<T]}L_{\delta_\tau} + 1_{[\tau=\delta_\tau]\cap[\delta_\tau=T]}L_T \\
&= 1_{[\tau=\delta_\tau]}L_\tau.
\end{aligned}$$

It follows that for any $t \in [\tau, \delta_\tau]$, $Y_t \geq L_t$. Actually we cannot have

$$P[L_{\delta_\tau} > Y_{\delta_\tau}] > 0$$

because if so we obtain a contradiction in making the same reasoning after replacing τ by δ_τ. Hence, for any stopping time τ we have $Y_\tau \geq L_\tau$ and, since Y and L are optional processes, from optional section theorem (see e.g. [4], pp. 220), we conclude that $P-a.s.$, $\forall t \leq T$, $Y_t \geq L_t$.

$\boxed{\textit{Step 3:}}$ Y satisfies (15).

For $n \geq 0$, let us introduce the process $\tilde{Y}^n$ defined by:

$$\forall t \in [\tau, \delta_\tau], \tilde{Y}_t^n = Y_t^n - \Delta_t^{n,d} = Y_t^n - \left(K_t^{n,d} - K_\tau^{n,d}\right).$$

First note that for any $t \in [\tau, \delta_\tau[$, $K_t^{n,d} - K_\tau^{n,d} = 0$. Therefore making the substitution in (16), we obtain: $\forall t \in [\tau, \delta_\tau]$,

$$\begin{aligned}
\tilde{Y}_t^n = Y_{\delta_\tau}^n - \Delta_{\delta_\tau}^{n,d} + \int_t^{\delta_\tau} (g(s) + n(\tilde{L}_s^n - \tilde{Y}_s^n)^+)\mathrm{d}s - \int_t^{\delta_\tau} (Z_s^n \mathrm{d}B_s \\
+ \int_E V_s^n(e)\widetilde{\mu}(\mathrm{d}s, \mathrm{d}e)),
\end{aligned}$$

where $\tilde{L}_t^n := L_t - \Delta_t^{n,d}$. On the other hand, it holds true that:

$$\forall t \in [\tau, \delta_\tau], \tilde{Y}^n \geq \tilde{Y}^n \wedge \tilde{L}^n$$

and $\int_\tau^{\delta_\tau}(\tilde{Y}_s^n - \tilde{Y}_s^n \wedge \tilde{L}_s^n)\mathrm{d}K_s^n = 0$, where $K_t^n = \int_\tau^t n(\tilde{L}_s - \tilde{Y}_s^n)^+\mathrm{d}s$. Hence, thanks to Remark 3, we have: $\forall t \in [\tau, \delta_\tau]$,

$$\begin{aligned}\tilde{Y}_t^n &= \operatorname{esssup}_{t\le\sigma\le\delta_\tau} E\Big[1_{[\sigma=\delta_\tau]}\left(Y_{\delta_\tau}^n - \Delta_{\delta_\tau}^{n,d}\right) + 1_{[\sigma<\delta_\tau]}\left(\tilde{L}_\sigma^n \wedge \tilde{Y}_\sigma^n\right) \\ &\qquad + \int_t^\sigma g(s)\mathrm{d}s|\mathscr{F}_t\Big] \\ &= \operatorname{esssup}_{t\le\sigma\le\delta_\tau} E\Big[1_{[\sigma=\delta_\tau]}\left(Y_{\delta_\tau}^n - \Delta_{\delta_\tau}^{n,d}\right) + 1_{[\sigma<\delta_\tau]}\left(L_\sigma \wedge Y_\sigma^n\right) \\ &\qquad + \int_t^\sigma g(s)\mathrm{d}s|\mathscr{F}_t\Big].\end{aligned}$$

Let us now consider the following BSDE: $\forall t \in [0, \delta_\tau]$,

$$\begin{cases}\tilde{Y} \in \mathscr{S}^2, \tilde{Z} \in \mathscr{H}^{2,d}, \tilde{V} \in \mathscr{L}^2 \text{ and } \tilde{K}^+ \in \mathscr{A}^2; \\ \tilde{Y}_t = Y_{\delta_\tau} - K'^{d,-}_{\delta_\tau} + \int_t^{\delta_\tau} g(s)\mathrm{d}s + (\tilde{K}_{\delta_\tau}^+ - \tilde{K}_t^+) - \int_t^{\delta_\tau}\tilde{Z}_s\mathrm{d}B_s - \int_t^{\delta_\tau}\int_E \tilde{V}_s(e)\tilde{\mu}(\mathrm{d}s,\mathrm{d}e), \\ \tilde{Y}_t \ge L_t - K'^{d,-}_t := \tilde{L}_t, \text{ and } \tilde{K}_t^+ = \tilde{K}_t^{c,+} + \tilde{K}_t^{d,+} \text{ satisfies:} \\ \int_\tau^{\delta_\tau}(\tilde{Y}_s - \tilde{L}_s)\mathrm{d}\tilde{K}_s^{c,+} = 0,\ \tilde{K}^{d,+} \text{ is predictable and } \tilde{K}_t^{d,+} = \sum_{0<s\le t}(\tilde{L}_{s-} - \tilde{Y}_s)^+.\end{cases}$$

The existence of the solution $(\tilde{Y}_t, \tilde{Z}_t, \tilde{V}_t, \tilde{K}_t)_{t\le\delta_\tau}$ is guaranteed by Theorem 3 and Remark 3. Additionally, we have the following characterization for $\tilde{Y}$: $\forall t \in [\tau, \delta_\tau]$,

$$\tilde{Y}_t = \operatorname{esssup}_{t\le\sigma\le\delta_\tau} E\left[1_{[\sigma=\delta_\tau]}(Y_{\delta_\tau} - K'^{d,-}_{\delta_\tau}) + 1_{[\sigma<\delta_\tau]}L_\sigma + \int_t^\sigma g(s)\mathrm{d}s|\mathscr{F}_t\right].$$

We are going now to prove that $P-a.s.$ for any $t \in [\tau, \delta_\tau]$, $\tilde{Y}_t^n \nearrow \tilde{Y}_t$. Actually, $P-a.s.$, for any $t \in [\tau, \delta_\tau]$ we have:

$$1_{[\tau\le t<\delta_\tau]}L_t \wedge Y_t^n + 1_{[t=\delta_\tau]}(Y_{\delta_\tau}^n - \Delta_{\delta_\tau}^{n,d}) \nearrow 1_{[\tau\le t<\delta_\tau]}L_t + \wedge 1_{[t=\delta_\tau]}(Y_{\delta_\tau} - K'^{d,-}_{\delta_\tau}).$$

Note that the increasing convergence of $(Y_{\delta_\tau}^n - \Delta_{\delta_\tau}^{n,d})$ to $Y_{\delta_\tau} - K'^{d}_{\delta_\tau}$ is obtained from *Step 2*. Using now the fact that the Snell envelope operator is continuous through nondecreasing sequence of *rcll* processes (see e.g. Appendix in [13]) we obtain that: $\tilde{Y}^n \nearrow \tilde{Y}$, i.e., for any $t \in [\tau, \delta_\tau]$, $Y_t^n - \Delta_t^{n,d} \nearrow \tilde{Y}_t$. It follows that for any $t \in [\tau, \delta_\tau]$, $Y_t = \tilde{Y}_t + K'^{d,-}_t$. Taking now into account the equation satisfied by $\tilde{Y}$ we obtain: $\forall t \in [\tau, \delta_\tau]$,

$$\begin{aligned}Y_t &= Y_{\delta_\tau} - (K'^{d,-}_{\delta_\tau} - K'^{d,-}_t) + \int_t^{\delta_\tau} g(s)\mathrm{d}s + (\tilde{K}_{\delta_\tau}^{c,+} - \tilde{K}_t^{c,+}) + (\tilde{K}_{\delta_\tau}^{d,+} - \tilde{K}_t^{d,+}) \\ &\quad - \int_t^{\delta_\tau}\tilde{Z}_s\mathrm{d}B_s - \int_t^{\delta_\tau}\int_E \tilde{V}_s(e)\widetilde{\mu}(\mathrm{d}s,\mathrm{d}e).\end{aligned} \tag{16}$$

Next let us set $K_t'^{c,+} = (\tilde{K}^{c,+}_{(t\vee\tau)\wedge\delta_\tau} - \tilde{K}^{c,+}_\tau)$, $t \le T$ (and then $K_\tau'^{c,+} = 0$). Then the process $K'^{c,+}$ is non-decreasing continuous and satisfies

$$\int_\tau^{\delta_\tau} (Y_s - L_s)\mathrm{d}K_s'^{c,+} = 0$$

since $Y_t - L_t = \tilde{Y}_t - \tilde{L}_t$ for any $t \in [\tau, \delta_\tau]$. Next we set $K_t'^{d,+} = (\tilde{K}^{d,+}_{(t\vee\tau)\wedge\delta_\tau} - \tilde{K}^{d,+}_\tau)$, $t \le T$ (and then $K_\tau'^{d,+} = 0$). Then $K'^{d,+}$ is nondecreasing predictable and purely discontinuous since $\tilde{K}^{d,+}$ is so. Finally for $t \le T$ let us set $Z'_t = \tilde{Z}_t 1_{[\tau,\delta_\tau]}(t)$ and $V'_t = \tilde{V}_t 1_{[\tau,\delta_\tau]}(t)$. Therefore, using (16) we obtain that the 5-uplet $(Y, Z', V', K'^{c,+}, K'^{d,+}, K'^{d,-})$ satisfies (b). It remains now to show property (d).

Let η be a predictable stopping time s.t. $\eta < \delta_\tau$ and $\Delta K_\eta'^{d,+} > 0$. Therefore, $\Delta K_\eta'^{d,+} = \Delta\tilde{K}_\eta^{d,+} = (\tilde{L}_{\eta-} - \tilde{Y}_\eta)^+ = (L_{\eta-} - Y_\eta)^+$ since $K_t'^{d,-} = 0$ for any $t \in [\tau, \delta_\tau[$. Suppose now that $\eta = \delta_\tau$ and $\Delta K_\eta'^{d,+} > 0$. Therefore thanks to (16) we have $0 < \Delta K_\eta'^{d,+} = \Delta\tilde{K}_\eta^{d,+} = Y_{\eta-} - Y_\eta + K_\eta'^{d,-} = \tilde{Y}_{\eta-} - Y_\eta + K_\eta'^{d,-} = L_{\eta-} - Y_\eta + K_\eta'^{d,-}$. Recall here that the Poisson part in (16) have only inaccessible jumps and η is predictable. But if $K_\eta'^{d,-} > 0$ then $Y_{\eta-} = U_{\eta-}$ and $K_\eta'^{d,-} = Y_\eta - U_{\eta-}$, then $0 < \Delta K_\eta'^{d,+} = \Delta\tilde{K}_\eta^{d,+} = L_{\eta-} - Y_\eta + Y_\eta - U_{\eta-} \le 0$, which is contradictory. It follows that $K_\eta'^{d,-} = 0$ and then $\Delta K_\eta'^{d,+} = L_{\eta-} - Y_\eta = (L_{\eta-} - Y_\eta)^+$. The proof is now complete.

We now consider the following decreasing penalization scheme:

$$\begin{cases} (i)\ \ Y'^n \in \mathscr{S}^2,\ Z'^n \in \mathscr{H}^{2,d},\ V'^n \in \mathscr{L}^2,\ K'^n \in \mathscr{A}^2, \\ (ii)\ \ Y_t'^n = \xi + \int_t^T \{g(s) - n(Y_s'^n - U_s)^+\}\mathrm{d}s + (K_T'^n - K_t'^n) \\ \qquad\qquad - \int_t^T Z_s'^n \mathrm{d}B_s - \int_t^T \int_E V_s'^n(e)\tilde{\mu}(\mathrm{d}s, \mathrm{d}e),\ \ \forall t \in [0,T] \\ (iii)\ Y'^n \ge L, \\ (iv)\ \text{if } K'^{n,c} \text{ (resp. } K'^{n,d}\text{) is the continuous (resp. purely discontinuous)} \\ \qquad \text{part of } K'^n,\ \text{then } \int_0^T (Y_s'^n - L_{s-})\mathrm{d}K_s'^{n,c} = 0 \text{ and } K'^{n,d} \text{ is predictable} \\ \qquad \text{and satisfies } K_t'^{n,d} = \sum_{0<s\le t}(L_{s-} - Y_s'^n)^+, \forall t \le T. \end{cases}$$

For any $n \ge 0$, the quadruple (Y'^n, Z'^n, V'^n, K'^n) exists through Theorem 3. Using once more the comparison result Theorem 4, we have for any $n \ge 0$ P-$a.s.$, $L \le Y'^{n+1} \le Y'^n$ therefore there exists a process $Y' := (Y'_t)_{t\le T}$ s.t. P-$a.s.$, $Y' \ge L$ and for any $t \le T$, $Y'_t = \lim_{n\to\infty} Y_t'^n$. Additionally, thanks to the Lebesgue dominated convergence theorem the sequence $(Y'^n)_{n\ge 0}$ converges to Y' in $\mathscr{H}^{2,1}$.

Next for any stopping time τ and $n \ge 0$, let us set:

$$\begin{aligned} \lambda_\tau^n :&= \inf\left\{s \ge \tau, K_s'^n - K_\tau'^n > 0\right\} \wedge T \\ &= \inf\left\{s \ge \tau : K_s'^{n,d} - K_\tau'^{n,d} > 0\right\} \wedge \inf\left\{s \ge \tau, K_s'^{n,c} - K_\tau'^{n,c} > 0\right\} \wedge T. \end{aligned}$$

The same analysis reveals that $(\lambda_\tau^n)_{n\geq 0}$ is a nondecreasing sequence of stopping times and converges to another stopping time $\lambda_\tau := \lim_{n\to\infty} \lambda_\tau^n$. The following properties related to Y', which can be proved analogously as in Prop. 3–4, hold true:

Proposition 5. *(i) P-a.s.,* $1_{[\lambda_\tau<T]} Y'_{\lambda_\tau} \leq 1_{[\lambda_\tau<T]}(L_{\lambda_\tau} + 1_{[\lambda_\tau>\tau]}(\triangle L_{\lambda_\tau})^-)$. *(ii) There exists a 4-uplet of processes* $(Z", V", K"^{,-}, K"^{d,+})$ *which in association with* Y' *satisfies:*

$$
\begin{cases}
(a)\ (Z", V", K"^{,-}, K"^{d,+}) \in \mathscr{H}^{2,d} \times \mathscr{L}^2 \times \mathscr{A}^2 \times \mathscr{A}^2, \\
(b)\ Y'_t = Y'_{\lambda_\tau} + \int_t^{\lambda_\tau} g(s)\mathrm{d}s - (K"^{,-}_{\lambda_\tau} - K"^{,-}_t) + (K"^{d,+}_{\lambda_\tau} - K"^{d,+}_t) \\
\qquad - \int_t^{\lambda_\tau} Z"_s \mathrm{d}B_s - \int_t^{\lambda_\tau} \int_E V"_s(e)\tilde{\mu}(\mathrm{d}s, \mathrm{d}e),\ \ \forall t \in [\tau, \lambda_\tau] \\
(c)\ \forall t \in [0,T], L_t \leq Y'_t \leq U_t, \\
(d)\ K"^{,-}_\tau = 0 \textit{ and if } K"^{c,-} (\textit{resp. } K"^{d,-}) \textit{is the continuous part} \\
\qquad (\textit{resp. purely discontinuous part of}) K"^{,-},\ \textit{then } K"^{d,-} \textit{ is predictable,} \\
\qquad K"^{d,-}_t = \sum_{\tau<s\leq t} (Y'_s - U_{s-})^+,\ \forall t \in [\tau, \delta_\tau] \textit{ and } \int_\tau^{\lambda_\tau} (U_s - Y'_s)\mathrm{d}K"^{c,-}_s = 0, \\
(e)\ K"^{d,+} \textit{ is predictable and purely discontinuous, } K"^{d,+}_t = 0\, \forall t \in [\tau, \lambda_\tau[, \\
\qquad \textit{and if } K"^{d,+}_{\lambda_\tau} > 0 \textit{ then } Y'_{\lambda_\tau-} = L_{\lambda_\tau-} \textit{ and } K"^{d,+}_{\lambda_\tau} = (L_{\lambda_\tau-} - Y'_{\lambda_\tau})^+.
\end{cases}
$$

Remark 5. The process Y' is *rcll* on the interval $[\tau, \lambda_\tau]$.

5.3 Existence of the Local Solution

Recall that Y (resp. Y') is the limit of the increasing (resp. decreasing) approximating scheme. In fact, the processes Y and Y' are undistinguishable as we show it now.

Proposition 6. *P-a.s., for any* $t \leq T$, $Y_t = Y'_t$. *Additionally* Y *is rcll.*

Proof. First let us point out that for any $n, m \geq 0$ and all $t \in [0,T]$ we have $Y_t^n \leq Y_t'^m$. Actually to prove this claim, we just need to apply Itô-Meyer's formula as in Theorem 4 with $\psi(Y^n - Y'^m)$ where $\psi(x) = (x^+)^2$, $x \in R$, and to remark that:

$$
\int_t^T \psi'\left(Y_s^n - Y_s'^m\right) m \left(Y_s'^m - U_s\right)^+ \mathrm{d}s
$$
$$
= \int_t^T \psi'\left(Y_s^n - Y_s'^m\right) n \left(L_s - Y_s^n\right)^+ \mathrm{d}s = 0.
$$

Then we argue as in Theorem 4 to obtain that for any $t \leq T$ we have $Y_t^n \leq Y_t'^m$. Therefore $P - a.s., \forall t \leq T, Y_t \leq Y'_t$.

Next let τ be a stopping time and μ_τ^p another stopping time defined by:

$$
\mu_\tau^p := \inf\{s \geq \tau : Y_s \geq U_s - p^{-1} \text{ or } Y'_s \leq L_s + p^{-1}\} \wedge T,
$$

where $p \geq 1$ is a real constant. First notice that for all $s \in [\tau, \mu_\tau^p] \cap [\tau < \mu_\tau^p]$ and all n we have $Y_{s-}^n < U_{s-}$ and $Y_{s-}^{\prime n} > L_{s-}$. Therefore for any $s \in [\tau, \mu_\tau^p]$ we have $d(K_s^n + K_s^{\prime n}) = 0$. Now using Itô's formula with $(Y_t^{\prime n} - Y_t^n)^2$, $t \in [\tau, \mu_\tau^p]$, then taking expectation in both hand-sides yield:

$$E\left[(Y_\tau^{\prime n} - Y_\tau^n)^2\right] \leq E\left[(Y_{\mu_\tau^p}^{\prime n} - Y_{\mu_\tau^p}^n)^2\right]. \tag{17}$$

Finally take the limit as $n \to \infty$ to obtain, $E[(Y'_\tau - Y_\tau)^2] \leq E[(Y'_{\mu_\tau^p} - Y_{\mu_\tau^p})^2]$. Here, note that we are not allowed to apply Itô's formula with $Y - Y'$ because we do not know whether $Y - Y'$ is a semimartingale on $[\tau, \mu_\tau^p]$.

We now show that $E\left[(Y'_{\mu_\tau^p} - Y_{\mu_\tau^p})^2\right] \to 0$ as $p \to \infty$. First notice that

$$0 \leq (Y'_{\mu_\tau^p} - Y_{\mu_\tau^p})1_{[\tau < \mu_\tau^p]} \leq \frac{1}{p}$$

since $U \geq Y' \geq Y \geq L$. Let us now focus on the case when $\tau = \mu_\tau^p$. First we have:

$$\begin{aligned} 1_{[\tau=\mu_\tau^p]}\left(Y'_{\mu_\tau^p} - Y_{\mu_\tau^p}\right) = {} & 1_{[\tau=\mu_\tau^p]\cap[\tau<\delta_\tau\wedge\lambda_\tau]}\left(Y'_\tau - Y_\tau\right) \\ & + 1_{[\tau=\mu_\tau^p]\cap[\tau=\delta_\tau\wedge\lambda_\tau]}\left(Y'_\tau - Y_\tau\right). \end{aligned} \tag{18}$$

Suppose that $\omega \in [\tau = \mu_\tau^p] \cap [\tau < \delta_\tau \wedge \lambda_\tau]$. Then there exists a sequence of real numbers $(t_k)_{k\geq 0}$ which depends on p and ω s.t. $t_k \searrow \tau$ as $k \to \infty$ and $Y_{t_k} \geq U_{t_k} - \frac{1}{p}$ or $Y'_{t_k} \leq L_{t_k} + \frac{1}{p}$. So assume we have $Y_{t_k} \geq U_{t_k} - \frac{1}{p}$. Then taking the limit as $k \to \infty$ implies that $Y_\tau \geq U_\tau - \frac{1}{p}$ since $\omega \in [\tau < \delta_\tau]$ and we know that Y is *rcll* on $[\tau, \delta_\tau]$. It follows that $U_\tau \geq Y'_\tau \geq Y_\tau \geq U_\tau - \frac{1}{p}$. In the same way we can show that if $Y'_{t_k} \leq L_{t_k} + \frac{1}{p}$ then $L_\tau \leq Y_\tau \leq Y'_\tau \leq L_\tau + \frac{1}{p}$. Therefore, $1_{[\tau=\mu_\tau^p]\cap[\tau<\delta_\tau\wedge\lambda_\tau]}(Y'_\tau - Y_\tau) \leq \frac{1}{p}$. Finally, we deal with the second term of (18). We have:

$$\begin{aligned} & 1_{[\tau=\delta_\tau\wedge\lambda_\tau]}(Y'_\tau - Y_\tau) \\ & \quad = 1_{[\tau=\delta_\tau\wedge\lambda_\tau]\cap[\tau<T]}(Y'_\tau - Y_\tau) \\ & \quad = 1_{[\tau=\delta_\tau]\cap[\tau<T]\cap[\delta_\tau\leq\lambda_\tau]}(Y'_{\delta_\tau} - Y_{\delta_\tau}) + 1_{[\tau=\lambda_\tau]\cap[\tau<T]\cap[\lambda_\tau<\delta_\tau]}(Y'_{\lambda_\tau} - Y_{\lambda_\tau}) \\ & \quad = 1_{[\tau=\delta_\tau]\cap[\tau<T]\cap[\delta_\tau\leq\lambda_\tau]}(Y'_{\delta_\tau} - U_{\delta_\tau}) + 1_{[\tau=\lambda_\tau]\cap[\tau<T]\cap[\lambda_\tau<\delta_\tau]}(L_{\lambda_\tau} - Y_{\lambda_\tau}) \leq 0, \end{aligned}$$

because in that case, taking into account 3 - 5-(i), we have either $Y_{\delta_\tau} = U_{\delta_\tau}$ or $Y'_{\lambda_\tau} = L_{\lambda_\tau}$ and we know that $U \geq Y' \geq Y \geq L$.

It follows that $0 \leq (Y'_{\mu_\tau^p} - Y_{\mu_\tau^p})^2 \leq 1/p^2$, then taking the limit as $p \to \infty$ in (17) we deduce that $Y_\tau = Y'_\tau$. As τ is an arbitrary stopping time then $P - a.s.$, $Y = Y'$.

We are now going to deal with the second property. For any $t \leq T$, we have: $U_t \geq Y_t \geq Y_t^n$ and $L_t \leq Y_t^{\prime} \leq Y_t^{\prime n}$, hence from the right continuity of Y^n and $Y^{\prime n}$ we have:

$$\liminf_{s\downarrow t} Y_s \geq \liminf_{s\downarrow t} Y_s^n = Y_t^n \text{ and } \limsup_{s\downarrow t} Y_s = \limsup_{s\downarrow t} Y_s^{\prime} \leq \limsup_{s\downarrow t} Y_s^{\prime n} = Y_t^{\prime n}.$$

Letting $n \to \infty$ we get the right continuity of Y since $Y = Y'$. Let us now show that Y has left limits. Define the predictable processes $\bar{Y}$ and $\tilde{Y}$ as following: $\bar{Y}_t = \liminf_{s\uparrow t} Y_s$ and $\tilde{Y}_t = \limsup_{s\uparrow t} Y_s$. Then, we only need to prove that for any predictable stopping time τ, we have $\bar{Y}_\tau = \tilde{Y}_\tau$. Let $(s_k)_k$ be a sequence of stopping times which announce τ. Then we have:

$$\tilde{Y}_\tau = \limsup_{s_k\uparrow\tau} Y_{s_k} = \limsup_{s_k\uparrow\tau} Y'_{s_k} \leq \limsup_{s_k\uparrow\tau} Y'^n_{s_k} = \lim_{s_n\uparrow\tau} Y'^n_{s_k} = Y'^n_{\tau-} = Y'^n_\tau$$
$$+(L_{\tau-} - Y'^n_\tau)^+.$$

Letting now $n \to \infty$, we obtain, $\tilde{Y}_\tau \leq Y_\tau + (L_{\tau-} - Y_\tau)^+$. Similarly, we can also get that $\bar{Y}_\tau \geq Y_\tau - (Y_\tau - U_{\tau-})^+$. Since we obviously have $L_{\tau-} \leq \bar{Y}_\tau \leq \tilde{Y}_\tau \leq U_{\tau-}$ then combining the three inequalities yields:

$$L_{\tau-} \vee (Y_\tau - (Y_\tau - U_{\tau-})^+) \leq \bar{Y}_\tau \leq \tilde{Y}_\tau \leq U_{\tau-} \wedge (Y_\tau + (L_{\tau-} - Y_\tau)^+).$$

Note that both hand sides are equal to $L_{\tau-}1_{[Y_\tau < L_{\tau-}]} + Y_\tau 1_{[L_{\tau-}\leq Y_\tau \leq U_{\tau-}]} + U_{\tau-}1_{[Y_\tau > U_{\tau-}]}$. Therefore for any predictable stopping time τ, $\tilde{Y}_\tau = \bar{Y}_\tau$, due to the predictable section theorem (see e.g. [4], pp. 220), $\tilde{Y}$ and $\bar{Y}$ are undistinguishable. It follows that $\lim_{s\nearrow t} Y_s$ exists for any $t \leq T$ and then Y has left limits.

Through Propositions 3 and 4 and the previous one we have:

Corollary 1. *The process Y satisfies:*

$$Y_{\delta_\tau} \geq U_{\delta_\tau} - 1_{[\tau<\delta_\tau]}(\Delta U_{\delta_\tau})^+ \quad \text{on } [\delta_\tau < T]$$
$$\text{and } Y_{\lambda_\tau} \leq L_{\lambda_\tau} + 1_{[\tau<\lambda_\tau]}(\Delta L_{\lambda_\tau})^- \text{ on } [\lambda_\tau < T].$$

Summing up now the results obtained in Propositions 3, 4, and 5, we have the following result related to the existence of local solutions for the BSDE (3).

Theorem 5. *There exists a process $Y := (Y_t)_{t\in[0,T]}$ s.t.:*

(1) *Y is $\mathscr{P}$-measurable, rcll and satisfies: $Y_T = \xi$,*
(2) *for any stopping time τ there exists a stopping time $\theta_\tau \geq \tau$, P-a.s., and a quadruple of processes $(Z^\tau, V^\tau, K^{\tau,+}, K^{\tau,-}) \in \mathscr{H}^{2,d} \times \mathscr{L}^2 \times \mathscr{A}^2 \times \mathscr{A}^2$ ($K^{\tau,\pm}_\tau = 0$) s.t. P-a.s.,*

$$\begin{cases} (i) \quad Y_t = Y_{\theta_\tau} + \int_t^{\theta_\tau} g(s)\mathrm{d}s + (K^{\tau,+}_{\theta_\tau} - K^{\tau,+}_t) - (K^{\tau,-}_{\theta_\tau} - K^{\tau,-}_t) \\ \qquad -\int_t^{\theta_\tau} Z^\tau_s \mathrm{d}B_s - \int_t^{\theta_\tau}\int_E V^\tau_s \tilde{\mu}(\mathrm{d}s,\mathrm{d}e),\ \forall t \in [\tau,\theta_\tau]; \\ (ii) \quad \forall t \in [0,T],\ L_t \leq Y_t \leq U_t, \\ (iii)\ \int_\tau^{\theta_\tau}(U_s - Y_s)\mathrm{d}K^{\tau c,-}_s = \int_\tau^{\theta_\tau}(Y_s - L_s)\mathrm{d}K^{\tau c,+}_s = 0, \text{ where } K^{\tau c,\pm} \text{ is the} \\ \qquad \text{continuous part of } K^{\tau,\pm}, \\ (iv) \quad \text{the purely discontinuous part } K^{\tau d,\pm} \text{ are predictable and } \forall t \in [\tau,\theta_\tau], \\ \qquad K^{\tau d,+}_t = \sum_{\tau<s\leq t}(L_{s-} - Y_s)^+ \text{ and } K^{\tau d,-}_t = \sum_{\tau<s\leq t}(Y_s - U_{s-})^+. \end{cases}$$

Hereafter, we say that Y is the solution of $\mathscr{BL}(g,\xi,L,U)$.

Proof. Let $Y := (Y_t)_{t \le T}$ be the adapted process defined as the limit of the increasing (or decreasing) scheme. Obviously it is *rcll* and satisfies, $L \le Y \le U$ and $Y_T = \xi$, P-$a.s.$.

Let us now focus on (2). Let τ be a stopping time, let δ_τ be the stopping time defined in the previous section and finally let us set $\theta_\tau = \lambda_{\delta_\tau}$. Thanks to Proposition 5, there exists $(Z^{"\delta_\tau}, V^{"\delta_\tau}, K^{"\delta_\tau d,\pm})$ (which we still denote $(Z^{"}, V^{"}, K^{"d,\pm})$) s.t.:

$$\begin{cases} (a)\ (Z^{"}, V^{"}, K^{",-}, K^{"d,+}) \in \mathscr{H}^{2,d} \times \mathscr{L}^2 \times \mathscr{A}^2 \times \mathscr{A}^2 \\ (b)\ Y_t = Y_{\theta_\tau} + \int_t^{\theta_\tau} g(s)\mathrm{d}s - (K^{",-}_{\theta_\tau} - K^{",-}_t) + (K^{"d,+}_{\theta_\tau} - K^{"d,+}_t) \\ \qquad - \int_t^{\theta_\tau} Z^{"}_s \mathrm{d}B_s - \int_t^{\theta_\tau}\int_E V^{"}_s(e)\tilde{\mu}(\mathrm{d}s,\mathrm{d}e),\ \ \forall t \in [\delta_\tau, \theta_\tau] \\ (c)\ K^{",-}_{\delta_\tau} = 0 \text{ and if } K^{"c,-} \text{ (resp. } K^{"d,-}\text{)is the continuous part} \\ \qquad \text{(resp. purely discontinuous part) of } K^{",-},\ \text{then } K^{"d,-} \text{ is predictable,} \\ \quad K^{"d,-}_t = \sum_{\delta_\tau < s \le t} (Y_s - U_{s-})^+,\ \forall t \in [\delta_\tau, \theta_\tau] \text{ and } \int_{\delta_\tau}^{\theta_\tau} (U_s - Y_s)\mathrm{d}K^{"c,-}_s = 0 \\ (d)\ K^{"d,+} \text{ is predictable and purely discontinuous, } K^{"d,+}_t = 0\ \forall t \in [\delta_\tau, \theta_\tau[, \\ \qquad \text{and if} K^{"d,+}_{\theta_\tau} > 0 \text{ then } Y_{\theta_\tau -} = L_{\theta_\tau -} \text{ and } K^{"d,+}_{\theta_\tau} = (L_{\theta_\tau -} - Y_{\theta_\tau})^+. \end{cases} \tag{19}$$

Now for any $t \le T$, let us set:

(i) $Z^\tau_t := Z^{'}_t 1_{[\tau \le t \le \delta_\tau]} + Z^{''}_t 1_{[\delta_\tau < t \le \theta_\tau]}$ and $V^\tau_t := V^{'}_t 1_{[\tau \le t \le \delta_\tau]} + V^{''}_t 1_{[\delta_\tau < t \le \theta_\tau]}$;

(ii) $K^{\tau,+}_t = K^{\tau c,+}_t + K^{\tau d,+}_t := K^{'c,+}_{(t\wedge\delta_\tau)\vee\tau} + K^{'d,+}_{(t\wedge\delta_\tau)\vee\tau} + K^{"d,+}_{(t\wedge\theta_\tau)\vee\delta_\tau}$ and $K^{\tau,-}_t = K^{\tau c,-}_t + K^{\tau d,-}_t := K^{"c,-}_{(t\wedge\theta_\tau)\vee\delta_\tau} + K^{'d,-}_{(t\wedge\delta_\tau)\vee\tau} + K^{"d,-}_{(t\wedge\theta_\tau)\vee\delta_\tau}$.

The process Z^τ (resp. V^τ) belongs to $\mathscr{H}^{2,d}$ (resp. $\mathscr{L}^2$) and, through their definitions, the processes $K^{\tau d,\pm} \in \mathscr{A}^2$ are purely discontinuous and predictable and $K^{\tau c,\pm} \in \mathscr{A}^2$ are predictable and continuous.

Next we show that Y, Z^τ, V^τ and $K^{\tau,\pm}$ satisfy the relation (2). (2)–(i) and (2)–(iii) follow directly from the definitions of the processes. Now let η be a predictable stopping time s.t. $\tau \le \eta \le \theta_\tau$. Therefore thanks to relation $(2.i)$ we have:

$$\Delta Y_\tau = \Delta K^{\tau d,-}_\eta - \Delta K^{\tau d,+}_\eta.$$

But $\{\Delta K^{\tau d,-} > 0\} \subset \{Y \ge U_-\}$ and $\{\Delta K^{\tau d,+} > 0\} \subset \{Y \le L_-\}$. As $L_- < U_-$ then $\Delta K^{\tau d,-}$ and $\Delta K^{\tau d,+}_\eta$ cannot jump in the same time. Hence, the positive (resp. negative) predictable jumps of Y are the same as the ones of $K^{\tau d,-}$ (resp. $K^{\tau d,+}$).

Assume now that $\Delta K^{\tau d,+}_\eta > 0$. Therefore, the definitions of $K^{\tau d,+}$, $K^{'d,+}$ and $K^{"d,+}$ imply that

$$\begin{aligned} \Delta K^{\tau d,+}_\eta &= \Delta K^{'d,+}_\eta 1_{[\tau < \eta \le \delta_\tau]} + \Delta K^{"d,+}_\eta 1_{[\eta = \theta_\tau]} \\ &= (L_{\eta -} - Y_\eta)^+ 1_{[\tau < \eta \le \delta_\tau]} + 1_{[\eta = \theta_\tau]}(L_{\theta_\tau -} - Y_{\theta_\tau})^+ = (L_{\eta -} - Y_\eta)^+, \end{aligned}$$

because from (19) we deduce that on the interval $]\delta_\tau, \theta_\tau[$ the process Y has no predictable negative jump. Similarly for any predictable stopping time η s.t. $\tau \le \eta \le \theta_\tau$ and $\Delta K_\eta^{\tau d,-} > 0$, $\Delta K_\eta^{\tau d,-} = (Y_\eta - U_{\eta-})$. Thus we have proved $(2.iv)$.

Remark 6. When the process Y is fixed, from Proposition 2 we deduce that the quadruple $(Z^\tau, V^\tau, K^{\tau,+}, K^{\tau,-})$ is unique on $[\tau, \theta_\tau]$.

5.4 Existence of a Global Solution for the BSDE with two Completely Separated rcll Barriers

Theorem 6. *The BSDE with two reflecting rcll barriers associated with* (g, ξ, L, U) *has a unique solution.*

Proof. Let Y be the *rcll* process defined in Theorem 5. Then for any $n \ge 1$, there exists a stopping time γ_n, defined recursively as $\gamma_0 = 0, \gamma_n = \theta_{\gamma_{n-1}}$, and a unique quadruple $(Z^n, V^n, K^{n,+}, K^{n,-})$ which belongs to $\mathscr{H}^{2,d} \times \mathscr{L}^2 \times \mathscr{A}^2 \times \mathscr{A}^2$ and which with the process Y satisfy $\mathscr{BL}(\xi, g, L, U)$ on $[\gamma_{n-1}, \gamma_n]$.

First note that since for any $t \le T$ $U_t > L_t$ and $U_{t-} > L_{t-}$ then for any $n \ge 1$ we have $P[(\gamma_{n-1} = \gamma_n) \cap (\gamma_n < T)] = 0$ and the sequence $(\gamma_n)_{n\ge1}$ is of stationary type, i.e., $P[\omega, \gamma_n(\omega) < T, \forall n \ge 1] = 0$. In other words for ω fixed there exists an integer rank $n_0(\omega)$ s.t. for $n \ge n_0(\omega)$ $\gamma_n(\omega) = \gamma_{n+1}(\omega) = T$.

Next let us introduce the following processes $Z, V, K^\pm$: $P-a.s.$, for any $t \le T$, one sets:

$$\begin{aligned}
Z_t &= Z_t^1 1_{[0,\gamma_1]}(t) + \textstyle\sum_{n\ge1} Z_t^{n+1} 1_{]\gamma_n,\gamma_{n+1}]},\\
V_t &= V_t^1 1_{[0,\gamma_1]}(t) + \textstyle\sum_{n\ge1} V_t^{n+1} 1_{]\gamma_n,\gamma_{n+1}]},\\
K_t^\pm &= K_t^{c,\pm} + K_t^{d,\pm} = \begin{cases} K_t^{1c,\pm} + K_t^{1d,\pm} & \text{if } t \in [0,\gamma_1] \\ K_{\gamma_n}^{c,\pm} + K_t^{(n+1)c,\pm} + K_{\gamma_n}^{d,\pm} + K_t^{(n+1)d,\pm} & \text{if } t \in]\gamma_n, \gamma_{n+1}]. \end{cases}
\end{aligned}$$

Then the 5-uplet $(Y, Z, V, K^\pm)$ verify the BSDE and the uniqueness of the solution has been shown in Proposition 2.

Remark 7. The sequence of stopping times $(\gamma_k)_{k\ge0}$ will be called associated with the solution $(Y, Z, V, K^\pm)$. Also note that for any k, we have the following local integrability of the processes Z, V and $K^\pm$:

$$E\left[\int_0^{\gamma_k} \mathrm{d}s\{|Z_s|^2 + \int_E |V_s(e)|^2 \lambda(de)\} + (K_{\gamma_k}^+)^2 + (K_{\gamma_k}^-)^2\right] < \infty.$$

Let us now investigate under which conditions Mokobodski's condition introduced in (4) is verified at least locally.

Proposition 7. *There exists a sequence* $(\gamma_k)_{k\geq 0}$ *of stopping times s.t.:*

(i) for any $k \geq 0$, $\gamma_k \leq \gamma_{k+1}$ *and the sequence is of stationary type, i.e.*

$$P[\gamma_k < T, \forall k \geq 0] = 0(\gamma_0 = 0);$$

(ii) for any $k \geq 0$*, there exists a pair* (h^k, h'^k) *of nonnegative supermartingales which belong to* $\mathscr{S}^2$ *such that:*

$$P - a.s., \forall t \leq \gamma_k, L_t \leq h_t^k - h_t'^k \leq U_t. \tag{20}$$

Proof. Let $(Y, Z, V, K^{\pm})$ be the solution of the RBSDE associated with $(0, \xi, L, U)$, which exists thanks to Theorem 6. Let $(\gamma_k)_{k\geq 0}$ be the sequence of stopping times associated with this solution (see Remark 7). By construction, this sequence satisfies the claim (i). Let us focus on (ii). For $k \geq 1$ and $t \leq T$ one sets:

$$\begin{aligned} h^k_{t\wedge\gamma_k} &= E\left[Y^+_{\gamma_k} + (K^+_{\gamma_k} - K^+_{t\wedge\gamma_k})|\mathscr{F}_{t\wedge\gamma_k}\right] \text{ and } h'^k_{t\wedge\gamma_k} \\ &= E\left[Y^-_{\gamma_k} + (K^-_{\gamma_k} - K^-_{t\wedge\gamma_k})|\mathscr{F}_{t\wedge\gamma_k}\right]. \end{aligned}$$

Then h^k, h'^k are supermartingales of $\mathscr{S}^2$ which satisfy $L_t \leq h_t^k - h_t'^k \leq U_t$ for any $t \leq \gamma_k$ since $E[\int_0^{\gamma_k} \mathrm{d}s\{|Z_s|^2 + \int_E |V_s(e)|^2\lambda(de)\} + (K^+_{\gamma_k})^2 + (K^-_{\gamma_k})^2] < \infty$. Thus, we have the desired result.

We are now ready to establish the main result of this section which is the proof of Theorem 1.

Theorem 7. *The BSDE (3) with jumps and two reflecting rcll barriers associated with* (Φ, ξ, L, U) *has a unique solution, i.e., there exits a unique 5-uplet* (Y, Z, V, K^+, K^-) *which satisfies the BSDE (3).*

Proof. Let $(\gamma_k)_{k\geq 0}$ be the sequence of stopping times defined in Proposition 7. Therefore (20) is satisfied. Consider the following scheme: $(Z^0, V^0) = (0, 0)$ and for all $j \geq 1$,

$$\begin{cases} (i) \quad Y^j \in \mathscr{S}^2, Z^j \in \mathscr{H}^d, V^j \in \mathscr{L}, K^{\pm} \in \mathscr{A}, \\ (ii) \quad Y_t^j = \xi + \int_t^T 1_{[s\leq\gamma_j]}\Phi(s, Z_s^{j-1}, V_s^{j-1})\mathrm{d}s + \int_t^T d(K_s^{+,j} - K_s^{-,j}) \\ \qquad\qquad - \int_t^T Z_s^j \mathrm{d}B_s - \int_t^T \int_E V_s^j(e)\tilde{\mu}(\mathrm{d}s, \mathrm{d}e), \quad \forall t \in [0, T] \\ (iii) \; \forall t \leq T, \; L_t \leq Y_t^j \leq U_t \text{ and if } K^{c,\pm,j} \text{ is the continuous part of } K^{\pm,j} \\ \qquad\quad \text{then } \int_0^T (Y_t - L_t)\mathrm{d}K_t^{c,+,j} = \int_0^T (U_t - Y_t)\mathrm{d}K_t^{c,-,j} = 0, \\ (iv) \; \text{the purely discontinuous part } K^{d,\pm,j} \text{ of } K^{\pm,j} \text{ are predictable and} \\ \qquad\quad K_t^{d,-,j} = \sum_{0<s\leq t}(Y_s^j - U_{s-})^+ 1_{[\Delta U_s>0]} \text{ and} \\ \qquad\quad K_t^{d,+,j} = \sum_{0<s\leq t}(Y_s^j - L_{s-})^- 1_{[\Delta L_s<0]}\;, \\ (v) \quad \forall k \geq 0, \; E[(K^{\pm}_{\gamma_k})^2 + \int_0^{\gamma_k}(|Z_s^j|^2 + \|V_s^j\|^2)\mathrm{d}s] < \infty. \end{cases}$$

This sequence of reflected BSDEs defined recursively is well-posed. Let us prove it by induction. First note that the indicator $1_{[.\leq \gamma_j]}$ is in place in order to have a coefficient which belongs to $\mathscr{H}^{2,1}$.

For $j=1$, the quintuple $(Y^1, Z^1, V^1, K^{\pm,1})$ exists thanks to Theorem 6. Now for any $k \geq 0$, on $[0,\gamma_k]$, thanks to Proposition 7, Mokobodski's assumption is satisfied and $(Y^1, Z^1, V^1, K^{\pm,1})$ is a solution for the BSDE with two *rcll* barriers associated with $(1_{[t\leq \gamma_1]}\Phi(t,0,0), Y^1_{\gamma_k}, L, U)$. Therefore uniqueness implies that $E[\int_0^{\gamma_k}(|Z^1_s|^2 + \|V^1_s\|^2)\mathrm{d}s] < \infty$ since under Mokobodski's assumption we have existence, uniqueness and square integrability (see e.g. [13]) of the solution of the BSDEs of type (3).

Assume now that for some j the property is satisfied. In the same way as previously we can show that $(Y^{j+1}, Z^{j+1}, V^{j+1}, K^{\pm,j+1})$ exists. Thus we have proved the well-posedness of the sequence.

The rest of the proof is mainly the same as the one given in ([10], Theorem 4.2, Step 2) even if in this paper the obstacles have only inaccessible jumps, therefore it is omitted.

References

1. Barles, G., Buchdahn, R., Pardoux, E.: Backward stochastic differential equations and integral-partial differential equations. Stochastics and Stochastics Reports **60**, 57–83 (1997)
2. Bass, R.: General Theory of Processes, www.math.uconn.edu/bass/gtp.pdf
3. Bensoussan, A., Lions, J.L.: Contrôle impulsionnel et inéqualitions quasi variationnelles. Dunod, Paris (1982)
4. Dellacherie, C., Meyer, P.A.: Probabilités et Potentiel. Chap. 1–4, Hermann, Paris (1975)
5. Dellacherie, C., Meyer, P.A.: Probabilités et Potentiel. Chap. 5–8, Hermann, Paris (1975)
6. El-Karoui, N., Hamadène, S.: BSDEs and Risk-Sensitive Control, Zero-sum and Non zero-sum Game Problems of Stochastic Functional Differential Equations. Stochastic Processes and their Applications **107**, 145–169 (2003)
7. El-Karoui, N., Peng, S., Quenez, M.C.: Backward differential equations in finance. Mathematical Finance **7**(1), 1–71 (1997)
8. Essaky, E.H.: Reflected BSDEs with jumps and RCLL obstacle. Bulletin des Sciences Mathématiques **132**(8), 690–710 (2008)
9. Hamadène, S.: Mixed Zero-sum differential game and American game options. SIAM Journal Control Optimization **45**, 496–518 (2006)
10. Hamadène, S., Hassani, M.: BSDEs with two reflecting barriers driven by a Brownian motion and an independent Poisson noise and related Dynkin game. EJP **11**, 121–145, paper no.5 (2006). http://www.math.washington.edu/ejpecp/
11. Hamadène, S., Hassani, M., Ouknine, Y.: BSDEs with two general discontinuous reflecting barriers without Mokobodski's hypothesis. In: Bulletin des Sciences Mathématiques, doi:10.1016/j.bulsci.2010.03.001 (2010)
12. Hamadène, S., Lepeltier, J.-P.: Reflected BSDEs and mixed game problem. Stochastic Processes and their Applications **85**, 177–188 (2000)
13. Hamadène, S., Ouknine, Y.: Refelected Backward SDEs with general jumps. Preprint, Université du Maine, Le Mans (F.) and available at arXiv:0812.3965v1
14. Li, X., Tang, S.: Necessary condition for optimal control of stochastic systems with random jumps. SIAM JCO **332**, 1447–1475 (1994)

15. Øksendal, B., Sulem, A.: Applied Stochastic Control of Jump Diffusions. Springer Universitext (2005)
16. Peng, S.: Monotonic limit theorem of BSDE and nonlinear decomposition theorem of Doob-Meyer's type. Probability Theory and Related Fields **113**, 473–499 (1999)
17. Lepeltier, J.P., Xu, M.: Reflected BSDEs with two rcll barriers. ESAIM: PS **11**, 3–22 (2007)
18. Protter, E.P.: Stochastic Integration and Differential Equations. 2nd Edition, Version 2.1., Springer, B.H.N-Y (2000)

Discontinuous Value Function in Time-Optimal Differential Games

Liudmila Kamneva

Abstract The article is devoted to the study of the value function in time-optimal differential games. Suppose that some function to be tested is constructed. It is required to prove that this function coincides with the value function of the game. A theorem on sufficient conditions for a tested function to coincide with the value function of the time-optimal differential game under consideration is proved. The theorem covers the case of a discontinuous value function. An application of the theorem is illustrated by an example of a time-optimal second-order differential game with the dynamics of conflict-controlled material point.

1 Introduction

Differential game theory studies control problems under conditions of uncertainty and disturbances. Investigations of differential games began in the 1950–1960s with an analysis of mathematical models of conflict situations in dynamic systems. In these models, a motion of a control system is described by ordinary differential equations, which contain controls in the right-hand sides. A useful control is considered as an action of the first player, which minimizes a certain functional over the set of trajectories of the system, and the disturbance is assumed to be a result of a control action of the second player, whose aim is to maximize the same functional. The players' controls satisfy geometric constraints.

This paper considers differential games, in which the pay-off functional is the time required for the phase point to reach a given closed terminal set $M \subset \mathbb{R}^n$. Such games are called time-optimal differential games. They include, for example, pursuit-evasion games and time-optimal problems of control theory, which can be considered as game problems with a zero constraint on the second player's control.

L. Kamneva (✉)
Institute of Mathematics and Mechanics, S. Kovalevskaya str., 16, 620219, GSP-384, Ekaterinburg, Russia
e-mail: kamneva@imm.uran.ru

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_6,

We follow the positional formalization of a differential game described in books [1,2] by Krasovskii and Subbotin. Within the framework of the positional formalization, an important question in studying a differential game is searching for a value function, which puts the optimal guaranteed result of a game starting at some point in correspondence to this point. On the basis of the value function, optimal feedback control strategies can be constructed.

In the general case, the value function of a time-optimal differential game may be nonsmooth and discontinuous and, moreover, can admit the improper value ∞.

Suppose that some function to be tested is constructed. It is required to prove that this function coincides with the value function of the game.

The statement of this problem completely coincides with that considered in [12]. But the result given in [12] is mainly of theoretical interest and is difficult to apply to examples. The sufficient conditions of this chapter do not depend on the result formulated in [12].

The problem under consideration is closely related to characterization of the value function.

A differentiable value function is the unique classical solution of a boundary value problem (Dirichlet problem) for the first-order partial differential equation (PDE) associated with a time-optimal differential game (Isaacs–Bellman equation) [3].

If the value function is nonsmooth but continuous, then the main role in its characterization is played by notions of continuous u- and v-stable functions [2, p. 145], which arose from the theory of positional differential games. In this case, u-stable (v-stable) functions majorize (minorize) the value function of a differential game under the corresponding boundary condition, and the value function is a unique function possessing the properties of u- and v-stability. Let us remark here that the notions of u- and v-stable functions are equivalent to the notions of upper and lower viscosity solutions, see e.g. [4].

Characterization of a discontinuous value function is considerably more complicated and coincides with the description of a discontinuous minimax solution [5, p. 223] of the Dirichlet problem for the Isaacs–Bellman equation. Namely, in time-optimal problems, the value function is a unique lower semicontinuous u-stable function satisfying the null condition on the boundary of the terminal set, to which a sequence of upper semicontinuous v-stable functions converges pointwise. The v-stable functions satisfy the same boundary condition and are continuous on the boundary of M. Verification of existence of such a sequence and constructing it are difficult even for problems in the plane.

Let us remark that much attention was paid lately to investigations of discontinuous value functions in control problems and differential games as well as its characterization as solutions of appropriate PDEs. Many works are devoted to problems with fixed termination and discontinuous solutions of the corresponding initial (or final) value problems (Cauchy problems) for the first-order PDEs, see e.g. [6–8]. Termination of a time-optimal differential game is nonfixed, and if such a game has an autonomous dynamics, then it corresponds to a boundary value problem (Dirichlet problem) for the first-order PDE. Time-optimal control problems (with convex Hamiltonians) were considered e.g. in [4, 9]. Time-optimal differential games (with more general Hamiltonians) were investigated e.g. in [4, 5]. Within the theory of viscosity solutions, the notion of a discontinuous minimax solution of the Dirichlet problem is equivalent to the notion of envelope viscosity solution (e-solution), see Introduction and Sect. 1 in [10]; Chap. 5, Sect. 3 in [4], and the references therein.

This chapter gives a theorem on sufficient conditions for a function to coincide with the value function of the time-optimal differential game under consideration. The theorem covers the case of a discontinuous value function. These conditions require verification of properties similar to those of a discontinuous minimax solution but in arbitrarily small neighborhoods of subsets, into which the boundaries of the level sets of the function are decomposed. In many cases, consideration of several neighborhoods renders the conditions more convenient in practice than the direct use of the definition of a discontinuous minimax solution. The application of the theorem is illustrated by an example of a time-optimal differential game in the plane.

2 Statement of the Problem

Consider a control system, whose motion is described by the equation

$$\dot{x}(t) = f(x(t), u(t), v(t)), \quad t \geq 0. \tag{1}$$

Here, $x(t) \in \mathbb{R}^n$ is the phase state at the instant t; $u(t) \in P$ and $v(t) \in Q$ are the controls of the first (minimizing) and second (maximizing) players, P and Q are compact sets. Following [1, 2], we assume that the function f is continuous with respect to the set of variables and satisfies the sublinear growth condition and the Lipschitz condition with respect to x. Suppose that

$$H(x, p) := \min_{u \in P} \max_{v \in Q} \langle p, f(x, u, v) \rangle = \max_{v \in Q} \min_{u \in P} \langle p, f(x, u, v) \rangle. \tag{2}$$

Here, $x, p \in \mathbb{R}^n$ and the angle brackets denote the inner product of vectors. The function H is called the Hamiltonian of system (1).

Positional strategies of the first and second players are arbitrary functions $U : [0, \infty) \times \mathbb{R}^n \to P$ and $V : [0, \infty) \times \mathbb{R}^n \to Q$ [1, 2]. Strategies U and V generate bunches $X_1(x_0, U)$ and $X_2(x_0, V)$ of constructive motions [1, p. 33], which emanate from an initial position x_0 at the instant $t = 0$.

The aim of the first player is to make the point $x(t)$ approach the given closed set $M \subset \mathbb{R}^n$ as fast as possible. The second player tries either to prevent this point from reaching M or to maximize the reaching time. Thus, the pay-off functional for the time-optimal game problem has the form

$$J(x(\cdot); M) = \begin{cases} \infty, & \text{if } x(t) \notin M \text{ for any } t \geq 0, \\ \min\{t \geq 0 : \ x(t) \in M\} & \text{otherwise.} \end{cases}$$

If at the point x_0

$$\inf_{U} \sup_{x(\cdot)\in X_1(x_0,U)} J(x(\cdot);M) = \sup_{V} \inf_{x(\cdot)\in X_2(x_0,V)} J(x(\cdot);M) =: T(x_0;M),$$

then $T(x_0;M) \in [0,\infty]$ is called the value of the game at the point x_0.

Under the above conditions onto the function f, the value of the game exists for any $x_0 \in \mathbb{R}^n$ [1, 2]. The function $T(\,\cdot\,;M) : \mathbb{R}^n \to [0,\infty]$ is called the value function of the game.

Suppose given a function $\varphi(\,\cdot\,) : \Omega \to [0,\infty]$ defined on a closed set $\Omega \subseteq \mathbb{R}^n$. The problem consists in finding conditions on the function $\varphi(\cdot)$, under which

$$\varphi(x) = T(x;M), \quad x \in \Omega.$$

3 Properties of *u*- and *v*-Stable Functions

The value function is closely associated with the notions of u- and v-stable functions [2, 5].

Definition 1. A function $\omega(\cdot) : G \to [0,\infty]$ is u-stable on an open set $G \subseteq \mathbb{R}^n$ if it is lower semicontinuous and, for any $y_0 \in G$ and $v_* \in Q$, there exist an instant $\tau > 0$ and a solution $y(\cdot) : [0,\tau] \to G$ of the differential inclusion

$$\dot{y}(t) \in \mathrm{co}\{f(y(t),u,v_*) : u \in P\}, \quad y(0) = y_0, \tag{3}$$

such that

$$\omega(y(t)) \le \omega(y_0) - t, \quad t \in [0,\tau]. \tag{4}$$

For a finite function $\omega(\cdot)$, inequality (4) is equivalent to the inclusion

$$(\omega(y_0) - t, y(t)) \in \mathrm{epi}\,\omega, \quad t \in [0,\tau],$$

where $\mathrm{epi}\,\omega = \{(z,x) : \; z \ge \omega(x), \; x \in G\}$ is an epigraph of $\omega(\cdot)$.

Remark 1. The value function $T(\cdot;M)$ in time-optimal differential games can take the improper value ∞. To apply the theory of generalized (minimax or viscosity) solutions of first-order partial differential equations, one usually uses the Kruzhkov transformation [11] to a function $\omega(\cdot) : G \to [0,\infty]$:

$$\mathbf{w}(x) = 1 - \exp(-\omega(x)) : \; G \to [0,1]. \tag{5}$$

The Isaacs–Bellman equation [3]

$$H(x,\omega(x)) + 1 = 0$$

transforms into the form

$$H(x, \nabla \mathbf{w}(x)) + 1 - \mathbf{w}(x) = 0. \tag{6}$$

One of the equivalent properties, which define an upper (minimax or viscosity) solution to (6), is the following [5, Sects. 4.2, 19.4, 19.5]: for any $v \in Q$ the epigraph of the function $\mathbf{w}(\cdot)$ is weakly invariant with respect to the differential inclusion

$$\begin{aligned}
&(\dot{x}(t), \dot{z}(t)) \in \mathbf{E}^{+}(x(t), z(t), v),\\
&\mathbf{E}^{+}(x, z, v) = \mathrm{co}\,\{(f(x, u, v), g) \in \mathbb{R}^n \times \mathbb{R} : \ u \in P, \ g = z - 1\}.
\end{aligned}$$

Taking into account (5), one can see that the property of upper solution to (6) is equivalent to the property of u-stability given in Definition 1.

Definition 2. A function $\widetilde{\omega}(\cdot) : G \to [0, \infty]$ is v-stable on an open set $G \subseteq \mathbb{R}^n$ if it is upper semicontinuous and, for any $y_0 \in G$ and $u_* \in P$, there exists an instant $\tau > 0$ and a solution $y(\cdot) : [0, \tau] \to G$ of the differential inclusion

$$\dot{y}(t) \in \mathrm{co}\,\{f(y(t), u_*, v) : v \in Q\}, \quad y(0) = y_0, \tag{7}$$

such that

$$\widetilde{\omega}(y(t)) \geq \widetilde{\omega}(y_0) - t, \quad t \in [0, \tau]. \tag{8}$$

For a finite function $\widetilde{\omega}(\cdot)$, inequality (8) is equivalent to the inclusion

$$(\widetilde{\omega}(y_0) - t, y(t)) \in \mathrm{hypo}\,\omega, \quad t \in [0, \tau],$$

where $\mathrm{hypo}\,\widetilde{\omega} = \{(z, x) : \ z \leq \widetilde{\omega}(x), \ x \in G\}$ is a hypograph of $\widetilde{\omega}(\cdot)$.

Remark 2. In the same way as in Remark 1, one can see that one of equivalent properties, which define a lower solution to (6), is equivalent to the property of v-stability given in Definition 2.

Let us formulate lemmas, which relate u- and v-stable functions with the value function of the game for an appropriate terminal set.

Definition 3. *An admissible trajectory* is a solution of (1), where $u(\cdot) : [0, \infty) \to P$ and $v(\cdot) : [0, \infty) \to Q$ are some measurable control functions.

Lemma 1. *Let $G \subseteq \mathbb{R}^n$ be an open set, $D \subset \mathbb{R}^n$ be a closed set, a function $\omega(\cdot)$: $G \to [0, \infty]$ be u-stable on the set $G \setminus D$, $t_* \geq 0$, and*

$$D \cap G = \{x \in G : \ \omega(x) \leq t_*\} \neq \varnothing.$$

Assume that $x_* \in G \setminus D$ and any admissible trajectory with the initial point x_* does not leave the set G on $[0, \omega(x_*) - t_*]$. Then $\omega(x_*) - t_* \geq T(x_*; D)$.

Lemma 2. *Let $G \subseteq \mathbb{R}^n$ be an open set, $D \subset \mathbb{R}^n$ be a closed set, a function $\widetilde{\omega}(\cdot): G \to [0, \infty]$ be v-stable on the set $G \setminus D$, $t_* \geq 0$,*

$$D \cap G = \{x \in G: \widetilde{\omega}(x) \leq t_*\} \neq \varnothing,$$

and the function $\widetilde{\omega}(\cdot)$ be continuous at points of the set $D \cap G$. Assume that $x_* \in G \setminus D$, $\vartheta \in [0, \widetilde{\omega}(x_*) - t_*)$, and any admissible trajectory with the initial point x_* does not leave the set G on $[0, \vartheta]$. Then $\vartheta < T(x_*; D)$.

In the case $G = \mathbb{R}^n$, the truth of Lemmas 1 and 2 follows from [1, pp. 49–65] and is based on properties of strategies, which are extremal to the epigraph (hypograph) of u-stable (v-stable) function. One more proof procedure results from [5, pp. 243–251] provided that the right-hand side of the dynamic system is globally Lipschitzian with respect to the variable x. In the case $G \subset \mathbb{R}^n$, the idea of the proofs given in the Appendix is not changed in comparison with [1] because admissible trajectories do not leave G on the considered time interval.

4 The Theorem on Sufficient Conditions

Hereafter, we need the following lemma from [12].

Lemma 3. *Suppose closed sets $D_\tau \subset \mathbb{R}^n$, $\tau > 0$, decrease monotonically by inclusion as $\tau \to +0$ and $\cap_{\tau > 0} D_\tau = M$. Then*

$$\lim_{\tau \to +0} T(x; D_\tau) = T(x; M), \quad x \in \mathbb{R}^n.$$

Assume that $\overline{A}$ denotes the closure of a set $A \subset \mathbb{R}^n$ and $O(r)$ is an open ball in $\mathbb{R}^n$ of the radius $r > 0$ centered at the origin.

Theorem 1. *Suppose that $\Omega \subseteq \mathbb{R}^n$ and $M \subseteq \Omega$ are closed sets, a function*

$$\varphi(\cdot): \Omega \to [0, \infty]$$

is given, and

$$\Theta = \sup_{z \in \Omega} \varphi(z), \quad D(t) = \{x \in \Omega : \varphi(x) \leq t\}, \quad t \in [0, \Theta),$$
$$F(t) = \{x \in \partial D(t) : \varphi(x) = t\}, \quad B(t) = \{x \in \partial D(t) : \varphi(x) < t\},$$
$$S(t) = \overline{F(t)} \cap \overline{B(t)}, \quad G(t, \varepsilon) = \begin{cases} \varnothing, & S(t) = \varnothing, \\ S(t) + O(\varepsilon), & S(t) \neq \varnothing, \end{cases} \quad \varepsilon > 0.$$

Suppose also that the function $\varphi(\cdot)$ is lower semicontinuous, $D(0) = M$; $\mathscr{T} \subset (0, \Theta)$ is a finite (possibly, empty) set, and the following conditions hold.

(1) *For any* $t \in (0, \Theta) \setminus \mathcal{T}$ *such that* $S(t) \neq \varnothing$*, there exist a number* $\varepsilon_0 > 0$ *and a set* $G_\infty \subset G(t, \varepsilon_0) \setminus D(t)$ *such that*

(a) *the relations* $G(t, \varepsilon_0) \subset G_\infty \cup \Omega$ *and*

$$\lim_{\varepsilon \to +0} \sup\{\varphi(x) : x \in G(t, \varepsilon) \setminus G_\infty\} = t \tag{9}$$

hold, and the function

$$\omega(x) = \begin{cases} \varphi(x) - t, & x \in G(t, \varepsilon_0) \setminus (G_\infty \cup D(t)), \\ 0, & x \in D(t) \cap G(t, \varepsilon_0), \\ \infty, & x \in G_\infty \end{cases} \tag{10}$$

is u-stable on the set $G(t, \varepsilon_0) \setminus D(t)$*;*

(b) *there exists a sequence of functions*

$$\omega_k(\cdot) : G(t, \varepsilon_0) \to [0, \infty], \quad k \in \mathbb{N},$$

which are v-stable on the set $G(t, \varepsilon_0) \setminus D(t)$*, vanishing and continuous at the points of the set* $D(t) \cap G(t, \varepsilon_0)$*, and*

$$\lim_{k \to \infty} \omega_k(x) = \omega(x), \quad x \in G(t, \varepsilon_0). \tag{11}$$

(2) *For any* $t \in (0, \Theta) \setminus \mathcal{T}$ *such that* $F(t) \setminus S(t) \neq \varnothing$ *and any arbitrarily small* $\varepsilon > 0$*, there is a number* $\delta > 0$ *such that for the set*

$$G^F(t, \varepsilon, \delta) = F(t) \setminus G(t, \varepsilon) + O(\delta)$$

the inclusion $G^F(t, \varepsilon, \delta) \subset \Omega$ *holds and the function* $\varphi(\cdot)$ *is finite and u- and v-stable on the set* $G^F(t, \varepsilon, \delta)$.

(3) *For any* $t \in (0, \Theta) \setminus \mathcal{T}$ *such that* $B(t) \setminus S(t) \neq \varnothing$ *and any arbitrarily small* $\varepsilon > 0$*, there are a number* $\delta > 0$ *and a sequence of functions*

$$\omega_k^\infty(\cdot) : G^B(t, \varepsilon, \delta) \to [0, \infty], \quad k \in \mathbb{N},$$

where

$$G^B(t, \varepsilon, \delta) = B(t) \setminus G(t, \varepsilon) + O(\delta),$$

such that the functions $\omega_k^\infty(\cdot)$*,* $k \in \mathbb{N}$*, are v-stable on the set* $G^B(t, \varepsilon, \delta) \setminus D(t)$*, vanish and are continuous at the points of the set* $D(t) \cap G^B(t, \varepsilon, \delta)$*, and*

$$\lim_{k \to \infty} \omega_k^\infty(x) = \infty, \quad x \in G^B(t, \varepsilon, \delta) \setminus D(t). \tag{12}$$

(4) For any $x_0 \in \Omega \setminus M$ *such that* $\varphi(x_0) = \Theta < \infty$*, there exists a sequence* $\{x_k\}_1^\infty \subset \Omega$*, for which* $\varphi(x_k) < \varphi(x_0)$ *and* $x_k \to x_0$ *as* $k \to \infty$*.*
Then,

$$\varphi(x) = T(x; M), \quad x \in \Omega.$$

Proof. Suppose that $\mathscr{T} = \varnothing$.

If $\Theta = 0$, then $\varphi(x) = 0$, $x \in \Omega$. By the condition $D(0) = M$ of the theorem, we get $\Omega = M$ and the conclusion of the theorem is evident. From now on, we assume that $\Theta > 0$.

1°. Set $t_* \in (0, \Theta)$. For brevity, we will use the following notations:

$$D_* = D(t_*), \quad F_* = F(t_*), \quad B_* = B(t_*), \quad S_* = S(t_*).$$

Let us prove that there exists such a value $\vartheta > 0$, for which the equality is true:

$$W(t; D_*) = D(t_* + t), \quad t \in [0, \vartheta]. \tag{13}$$

Here, $W(t; D_*) = \{x \in \mathbb{R}^n : T(x; D_*) \leq t\}$ is a level set of the value function.

The three cases are possible: $S_* \neq \varnothing$; $S_* = \varnothing$ and $F_* \neq \varnothing$; $S_* = \varnothing$ and $B_* \neq \varnothing$.

Case $S_* \neq \varnothing$. For the instant t_*, let us choose the value $\varepsilon_0 > 0$, the set G_∞, and the functions $\omega_k(\cdot)$, $k \in \mathbb{N}$, which are spoken about in condition (1) of the theorem. Define the function $\omega(\cdot)$ by formula (10), where $t = t_*$.

a) Let us show that there exists such a value $\delta_* \in (0, \varepsilon_0)$ that

$$\varphi(x) - t_* = T(x; D_*), \quad x \in \big(\overline{F}_* + O(\delta_*)\big)\backslash(G_\infty \cup D_*). \tag{14}$$

Denote

$$G_*(\varepsilon) = G(t_*, \varepsilon).$$

Fix $\varepsilon_1 \in (0, \varepsilon_0)$. There exists such a value $\vartheta_1 > 0$ that for any initial point $x_0 \in G_*(\varepsilon_1)$ admissible trajectories do not leave the set $G_*(\varepsilon_0)$ on $[0, \vartheta_1]$. In virtue of (9), there exists such a value $\varepsilon_2 \in (0, \varepsilon_1]$ that

$$\omega(x) = \varphi(x) - t_* < \vartheta_1, \quad x \in G_*(\varepsilon_2)\backslash G_\infty.$$

Choose $x_0 \in G_*(\varepsilon_2) \setminus (G_\infty \cup D_*)$. Since any admissible trajectory with the initial point x_0 does not leave the set $G_*(\varepsilon_0)$ on $[0, \vartheta_1]$ and $\omega(x_0) < \vartheta_1$, and the function $\omega(\cdot)$ is u-stable, using Lemma 1 we get the inequality

$$\omega(x_0) \geq T(x_0; D_*). \tag{15}$$

Note that $\omega(x_0) > 0$. Choose an arbitrary value $\tau \in (0, \omega(x_0)]$. In virtue of condition (11), there exists such a number $k \in \mathbb{N}$ that the following inequalities are true:

$$\omega(x_0) - \tau < \omega_k(x_0) < \vartheta_1.$$

In addition, $\omega(x_0) - \tau \geq 0$. The function $\omega_k(\cdot)$ is v-stable on the set $G_*(\varepsilon_0)$, and is equal to zero and continuous at points of the set $G_*(\varepsilon_0) \cap D_*$. In virtue of Lemma 2, the following inequality is true:

$$\omega(x_0) - \tau < T(x_0; D_*), \quad \tau \in (0, \omega(x_0)].$$

Letting $\tau \to 0$ and using relation (15), we obtain equality (14) at any point of the set $G_*(\varepsilon_2) \setminus (G_\infty \cup D_*)$.

Suppose $F_* \setminus S_* \neq \varnothing$. Without loss of generality, we can assume $F_* \setminus G_*(\varepsilon_2) \neq \varnothing$.

Using condition (2) of the theorem, we have that there exists such a value $\delta \in (0, \varepsilon_2]$ that the function $\varphi(\cdot)$ is continuous and u- and v-stable on the set $G^F(t_*, \varepsilon_2, \delta)$. Let $\delta_1 \in (0, \delta)$. Choose such an instant $\vartheta_2 \in (0, \vartheta_1]$ that for any initial point $x_0 \in G^F(t_*, \varepsilon_2, \delta_1)$ admissible trajectories do not leave the set $G^F(t_*, \varepsilon_2, \delta)$ on $[0, \vartheta_2]$. In virtue of continuity and u-stability of the function $\varphi(\cdot)$ on the set $G^F(t_*, \varepsilon_2, \delta)$, there exists such a value $\delta_* \in (0, \delta_1]$ that

$$\varphi(x) - t_* \leq \vartheta_2, \quad x \in G^F(t_*, \varepsilon_2, \delta_*).$$

Using Lemmas 1 and 2, we see that for any point $x_0 \in G^F(t_*, \varepsilon_2, \delta_*)$ the following inequalities are valid:

$$\varphi(x_0) - t_* - \tau < T(x_0; D_*) \leq \varphi(x_0) - t_*, \quad \tau \in (0, \varphi(x_0) - t_*].$$

Letting $\tau \to 0$, we obtain equality (14) at any point of the set $G^F(t_*, \varepsilon_2, \delta_*)$.

Since

$$\overline{F}_* + O(\delta_*) \subset G_*(\varepsilon_2) \cup G^F(t_*, \varepsilon_2, \delta_*),$$

relation (14) is true.

In the case $F_* \setminus S_* = \varnothing$, we can take $\delta_* = \varepsilon_2$. In addition, let us set $\vartheta_2 = \vartheta_1$.

b) Let us prove that

$$W(\vartheta_2; D_*) \cap G_\infty \cap G_*(\delta_*) = \varnothing. \tag{16}$$

Fix $x_0 \in G_\infty \cap G_*(\delta_*)$. Then $\omega(x_0) = \infty$. By condition (11), we can choose such a number $k \in \mathbb{N}$ that $\omega_k(x_0) > \vartheta_2$. The function $\omega_k(\cdot)$ is v-stable on the set $G_*(\varepsilon_0)$, is equal to zero and continuous at points of the set $D_* \cap G_*(\varepsilon_0)$, and for any initial point x_0 admissible trajectories do not leave the set $G_*(\varepsilon_0)$ on $[0, \vartheta_2]$. Using Lemma 2, we have

$$\vartheta_2 < T(x_0; D_*), \quad x_0 \in G_\infty \cap G_*(\delta_*).$$

Thus, relation (16) is proved.

c) First, suppose that $B_* \setminus S_* \neq \varnothing$. Without loss of generality, we can assume that $B_* \setminus G_*(\delta_*) \neq \varnothing$. For the value δ_*, let us find a value $\delta_2 \in (0, \delta_*]$ and a sequence of functions

$$\omega_k^\infty(\cdot) : G^B(t_*, \delta_*, \delta_2) \to [0, \infty], \quad k \in \mathbb{N},$$

which satisfy the properties described in condition (3) of the theorem.

Fix $\delta_3 \in (0, \delta_2)$. Choose such an instant $\vartheta_3 \in (0, \vartheta_2]$ that, for any initial point $x_0 \in G^B(t_*, \delta_*, \delta_3)$, admissible trajectories do not leave the set $G^B(t_*, \delta_*, \delta_2)$ on $[0, \vartheta_3]$.

Let $x_0 \in G^B(t_*, \delta_*, \delta_3) \setminus D_*$. In virtue of (12), choose such a number $k \in \mathbb{N}$ that $\omega_k^\infty(x_0) > \vartheta_3$. The function $\omega_k^\infty(\cdot)$ is v-stable on the set $G^B(t_*, \delta_*, \delta_2)$, is equal to zero and continuous at the points of the set $D_* \cap G^B(t_*, \delta_*, \delta_2)$. Using Lemma 2, we get

$$\vartheta_3 < T(x_0; D_*), \quad x_0 \in G^B(t_*, \delta_*, \delta_3) \setminus D_*.$$

Thus,

$$W(\vartheta_3; D_*) \cap (G^B(t_*, \delta_*, \delta_3) \setminus D_*) = \varnothing. \tag{17}$$

In the case $B_* \setminus S_* = \varnothing$, let us set $\vartheta_3 = \vartheta_2$ and $\delta_3 = \delta_*$.

d) Find such an instant $\vartheta \in (0, \vartheta_3]$ that for any initial point $x_0 \notin D_* + O(\delta_3)$ admissible trajectories do not attain the set D_* on $[0, \vartheta]$. We have $T(x_0; D_*) > \vartheta$ for $x_0 \notin D_* + O(\delta_3)$. Therefore, $W(\vartheta; D_*) \subset D_* + O(\delta_3)$.

e) Taking into account (16) and (17) (the last is used for the case $B_* \setminus S_* \neq \varnothing$), we obtain

$$W(\vartheta; D_*) \setminus D_* \subset \big(\overline{F}_* + O(\delta_3)\big) \setminus (G_\infty \cup D_*).$$

Thus, in virtue of (14) and inequality $\delta_3 < \delta_*$, relation (13) is fulfilled.

Case $S_* = \varnothing$ and $F_* \neq \varnothing$. In this case, $\partial D_* = F_*$, and for any $\varepsilon > 0$ and $\delta > 0$ we have $G^F(t_*, \varepsilon, \delta) = F_* + O(\delta)$.

By condition (2) of the theorem, there exists such a value $\delta_0 > 0$ that the function $\varphi(\cdot)$ is continuous and has properties of u- and v-staibity on the set $F_* + O(\delta_0)$.

Assume $\delta_1 \in (0, \delta_0)$. Choose such an instant $\vartheta_1 > 0$ that for any initial point $x_0 \in F_* + O(\delta_1)$ admissible trajectories do not leave the set $F_* + O(\delta_0)$ on $[0, \vartheta_1]$. By continuity and u-stability of the function $\varphi(\cdot)$, there exists such a value $\delta_* \in (0, \delta_1]$ that

$$\varphi(x) - t_* \leq \vartheta_1, \quad x \in F_* + O(\delta_*).$$

In virtue of Lemmas 1 and 2, for any point $x_0 \in F_* + O(\delta_*)$ the following inequalities are true:

$$\varphi(x_0) - t_* - \tau < T(x_0; D_*) \leq \varphi(x_0) - t_*, \quad \tau \in (0, \varphi(x_0) - t_*].$$

Letting $\tau \to 0$, we obtain the equality

$$\varphi(x) - t_* = T(x; D_*), \quad x \in F_* + O(\delta_*). \tag{18}$$

Let us find such an instant $\vartheta \in (0, \vartheta_1]$ that for any initial point $x_0 \notin D_* + O(\delta_*)$ admissible trajectories do not attain the set D_* on $[0, \vartheta]$. We have $T(x_0; D_*) > \vartheta$ for $x_0 \notin D_* + O(\delta_*)$. Therefore, $W(\vartheta; D_*) \subset D_* + O(\delta_*)$. By equality (18), relation (13) is fulfilled.

Case $S_* = \varnothing$ *and* $B_* \neq \varnothing$. In this case, $\partial D_* = B_*$ and for any $\varepsilon > 0$ and $\delta > 0$ we have $G^B(t_*, \varepsilon, \delta) = B_* + O(\delta)$. In addition, there exists such an instant $\vartheta_1 > 0$ that the following equation is true:

$$D(t_* + t) = D_*, \quad t \in [0, \vartheta_1]. \tag{19}$$

Let us find a value $\delta_0 > 0$ and a sequence of functions

$$\omega_k^\infty(\cdot) : B_* + O(\delta_0) \to [0, \infty], \quad k \in \mathbb{N},$$

which satisfy the properties given in condition (3) of the theorem.

Fix $\delta_1 \in (0, \delta_0)$. Choose such a value $\vartheta \in (0, \vartheta_1]$ that for any initial point $x_0 \in B_* + O(\delta_1)$ admissible trajectories do not leave the set $B_* + O(\delta_0)$ on $[0, \vartheta]$, and, in addition, for any initial point $x_0 \notin B_* + O(\delta_1)$ admissible trajectories do not attain the set D_* on $[0, \vartheta]$. We have $W(\vartheta; D_*) \subset B_* + O(\delta_1)$.

Let $x_0 \in (B_* + O(\delta_1)) \setminus D_*$. In virtue of (12), let us choose such a number $k \in \mathbb{N}$ that $\omega_k^\infty(x_0) > \vartheta$. Since the function $\omega_k^\infty(\cdot)$ is v-stable on the set $B_* + O(\delta_0)$ and is equal to zero and continuous at points of the set $D_* \cap (B_* + O(\delta_0))$, by Lemma 2 we have the estimate

$$\vartheta < T(x_0; D_*), \quad x_0 \in (B_* + O(\delta_1)) \setminus D_*.$$

Thus,

$$W(t; D_*) \cap ((B_* + O(\delta_1)) \setminus D_*) = \varnothing, \quad t \in [0, \vartheta].$$

Therefore, $W(t; D_*) = D_*$, $t \in [0, \vartheta]$. Relation (13) follows from equality (19).

2°. Choose $\tau \in (0, \Theta)$. Define the function $\varphi_\tau(\cdot) : \Omega \to [0, \infty]$ as follows:

$$\varphi_\tau(x) = \begin{cases} \varphi(x) - \tau, & x \in \Omega \setminus D(\tau), \\ 0, & x \in D(\tau). \end{cases}$$

For brevity, assume

$$D_\tau = D(\tau), \quad E(t) = D(\tau + t).$$

Define

$$\gamma = \sup\{\vartheta \in [0, \Theta - \tau) : W(t; D_\tau) = E(t),\ t \in [0, \vartheta]\}.$$

By the result proved in item 1°, we have the inequality $\gamma > 0$.

a) Let us show that

$$T(x; D_\tau) = \varphi_\tau(x), \quad x \in \cup_{\vartheta \in [0,\gamma)} E(\vartheta). \tag{20}$$

Choose

$$\vartheta \in [0, \gamma), \quad x \in E(\vartheta), \quad t = T(x; D_\tau), \quad t' = \varphi_\tau(x).$$

We have $t' \in [0, \vartheta]$. From the equality $W(\vartheta; D_\tau) = E(\vartheta)$, we obtain $t \in [0, \vartheta]$. Therefore,

$$W(t; D_\tau) = E(t), \quad W(t'; D_\tau) = E(t').$$

Since $x \in W(t; D_\tau) \cap E(t')$, we have $x \in W(t'; D_\tau) \cap E(t)$. This implies that

$$\varphi_\tau(x) \le t, \quad T(x; D_\tau) \le t',$$

and equality (20) is proved.

b) Let us prove that

$$T(x; D_\tau) = \varphi_\tau(x), \quad x \in \Omega. \tag{21}$$

Consider the following cases.

Case 1: $\gamma = \infty$. If $\Omega = \cup_{\vartheta \ge 0} E(\vartheta)$, then relation (21) coincides with the proved relation (20). Assume now that $\Omega \setminus \cup_{\vartheta \ge 0} E(\vartheta) \ne \varnothing$. Choose $x \in \Omega \setminus \cup_{\vartheta \ge 0} E(\vartheta)$. We have $\varphi_\tau(x) = \infty$. Using the definition of the value γ, we obtain that $x \notin \cup_{\vartheta \ge 0} W(\vartheta; D_\tau)$ and, therefore, $T(x; D_\tau) = \infty$. Thus, $\varphi_\tau(x) = T(x; D_\tau) = \infty$. In virtue (20), relation (21) for Case 1 is proved.

Case 2: $0 < \gamma < \infty$. Choose $x \in E(\gamma) \setminus \bigcup_{\vartheta \in [0,\gamma)} E(\vartheta)$. We have $\varphi_\tau(x) = \gamma$. Using the definition of the value γ, we find that $x \notin \cup_{\vartheta \in [0,\gamma)} W(\vartheta; D_\tau)$ and, therefore, $T(x; D_\tau) \ge \gamma$.

If $0 < \varphi(x) < \Theta$, then by the property of u-stability from conditions (1) and (2) there exists such a sequence $\{x_k\}_1^\infty$ that $\varphi(x_k) < \varphi(x)$ and $x_k \to x$ as $k \to \infty$. If $\varphi(x) = \Theta$, existence of the sequence $\{x_k\}_1^\infty$ is given by condition (4).

We have

$$t_k := \varphi_\tau(x_k) = \varphi(x_k) - \tau < \varphi(x) - \tau = \gamma, \quad k \in \mathbb{N}.$$

From lower semicontinuity of the function $\varphi(\cdot)$, we get that $\varphi(x) \le \lim_{k\to\infty} \varphi(x_k)$ and, therefore, $t_k \to \gamma$ as $k \to \infty$. Since $W(t_k; D_\tau) = E(t_k)$ and $x_k \in E(t_k)$, the inequality $T(x_k; D_\tau) \le t_k$ is true. Because of lower semicontinuity of the function $T(\cdot; D_\tau)$ [1,5], we get

$$\gamma = \lim_{k\to\infty} t_k \ge \lim_{k\to\infty} T(x_k; D_\tau) \ge T(x; D_\tau) \ge \gamma.$$

Therefore, $T(x; D_\tau) = \gamma$.

By relation (20), the following equality is proved:

$$T(x; D_\tau) = \varphi_\tau(x), \quad x \in E(\gamma). \tag{22}$$

Let us show that

$$W(\gamma; D_\tau) = E(\gamma). \tag{23}$$

The inclusion $E(\gamma) \subseteq W(\gamma; D_\tau)$ follows from (22). Let us prove the inverse inclusion.

Let $\hat{x} \in W(\gamma; D_\tau)$ and $\hat{t} = T(\hat{x}; D_\tau)$. We have $\hat{t} \leq \gamma$.

Assume that $\hat{t} < \gamma$. Using the definition of the value γ, we find that

$$\hat{x} \in W(\hat{t}; D_\tau) = E(\hat{t}) \subseteq E(\gamma).$$

Suppose $\hat{t} = \gamma$. Since $\gamma > 0$, we have $\hat{t} > 0$ and $\hat{x} \notin D_\tau$. By u-stability of the function $T(\cdot; D_\tau)$ [1, 5], there exists such a sequence $\{y_k\}_1^\infty$ that $t_k = T(y_k; D_\tau) < \gamma$ and $y_k \to \hat{x}$ as $k \to \infty$. Since $W(t_k; D_\tau) = E(t_k)$ and $y_k \in W(t_k; D_\tau)$, it follows that $y_k \in E(t_k) \subseteq E(\gamma)$. By closure of the set $E(\gamma)$, we get the inclusion $\hat{x} \in E(\gamma)$.

Thus, equality (23) is proved.

The results of item 1°, equality (23), and definition of the value γ imply that $\gamma = \Theta - \tau$. By equality (22) and $E(\gamma) = D(\Theta) = \Omega$, relation (21) is true for Case 2.

c) Let $x_* \in \Omega \setminus M$. Then for rather small $\tau > 0$, we have the equality $\varphi_\tau(x_*) = \varphi(x_*) - \tau$. By Lemma 3 and equality (21), we obtain

$$T(x_*; M) = \lim_{\tau \to +0} T(x_*; D(\tau)) = \lim_{\tau \to +0} \varphi_\tau(x_*) = \varphi(x_*).$$

This implies the statement of the theorem for the case $\mathscr{T} = \varnothing$.

Suppose now that $\mathscr{T} = \{a_1\}$, $a_1 \in (0, \Theta)$. A proof similar to the case $\mathscr{T} = \varnothing$ but applied to the function $\varphi(\cdot) : D(a_1) \to [0, \infty)$ gives the equality $\varphi(x) = T(x; M)$ for $x \in D(a_1)$. Further, introducing the notations

$$M_1 = D(a_1), \quad \varphi_1(x) = \begin{cases} \varphi(x) - a_1, & x \in \Omega \setminus M_1, \\ 0, & x \in M_1, \end{cases}$$

we apply a similar proof to the function $\varphi_1(\cdot) : \Omega \to [0, \infty]$ and to the differential game with the terminal set M_1. We obtain the equality $\varphi_1(x) = T(x; M_1)$, $x \in \Omega$. Using the relation $T(x; M_1) = T(x; M) - a_1$ [1,5], we find that $\varphi(x) = T(x; M)$, $x \in \Omega$.

Similar arguments are used in the case of any finite set $\mathscr{T} \subset (0, \Theta)$. □

5 An Example

5.1 A Model Time-Optimal Differential Game

Consider a model example of a game of pursuit on a straight line. The motion equations and constraints on the controls of the pursuer and evader have the form

$$\ddot{y}_P = u, \quad \dot{y}_E = v, \quad |u| \le \alpha, \quad 0 \le v \le \beta, \quad \alpha > 0. \tag{24}$$

The pursuer aims the point y_P be coincident with the point y_E for the minimal time in such a way that at the instant of coincidence the velocity $\dot{y}_P$ of the point y_P is equal to a value a given in advance, $a > \beta$.

Such a problem of pursuit can be formulated as a time-optimal differential game problem if the change of variables is introduced: $x_1 = y_P - y_E$, $x_2 = \dot{y}_P$. In virtue of system (24), we have

$$\dot{x}_1 = x_2 - v, \quad \dot{x}_2 = u, \quad |u| \le \alpha, \quad 0 \le v \le \beta. \tag{25}$$

The first player minimizes the transfer time of the phase point $x = (x_1, x_2)$ from a given initial position x_0 to the terminal set $M = (0, a)$, $a > \beta$. The interests of the second player are opposite.

If $\beta = 0$, then we get an optimal control problem and its solution is known (see, for example, [13, 14]). Figure 1 shows trajectories composed by parts of parabolas, which constitute two families. The left family corresponds to the control $u = \alpha$, the right one is referred to the control $u = -\alpha$. The time $\varphi(x_0)$ of motion from the point x_0 up to the point M along the mentioned curves is the optimal time in the control problem. The function $\varphi(\cdot)$ is discontinuous on the curve composed by the arcs $\mathscr{B}^+$ and $\mathscr{B}^-$, and it is nondifferentiable on the arc $\mathscr{E}$ of a parabola of the left family.

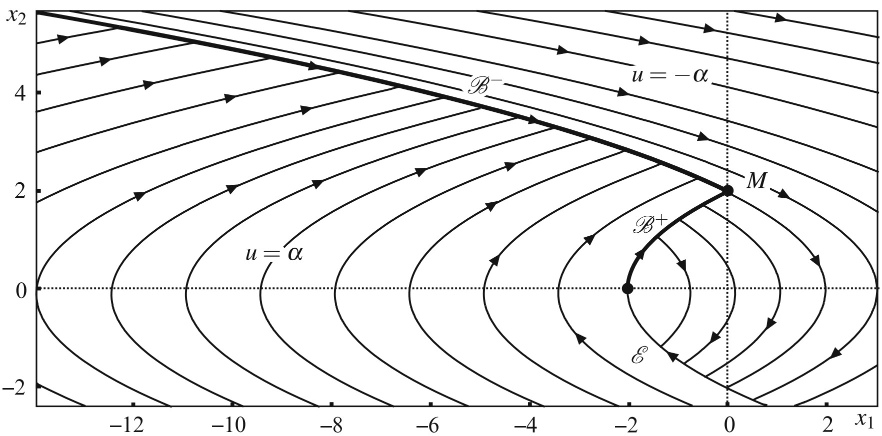


Fig. 1 Construction of the function $\varphi(\cdot)$ for $a = 2$, $\alpha = 1$, $\beta = 0$

5.2 Construction of a Function to Be Tested

In the case $\beta > 0$, let us describe a construction of trajectories that allows one to define the function $\varphi(\cdot)$. Such trajectories were investigated earlier in [15, 16].

In the set $x_1 \leq 0$, $x_2 \geq 0$, consider the curve $\mathscr{B}^{\pm}$, along which the trajectory of system (25) arrives at the point M for $v(t) = 0$ and $u(t) = \pm\alpha$. The parabola $\mathscr{B}^{+}$ intersects the x_1 axis at the point $\widetilde{x} = (-a^2/(2\alpha), 0)$. In the lower half-plane, we construct a curve $\mathscr{E}$ satisfying the differential equation

$$\frac{\mathrm{d}x_1}{\mathrm{d}x_2} = \frac{x_2}{\alpha} + \frac{2\beta x_2}{\alpha\left(\sqrt{\beta^2 - 4\beta x_2 - 4\alpha x_1 + 2x_2^2 + 2a^2} - \beta - 2x_2\right)}$$

and passing through the point $\widetilde{x}$. We denote the open part of the plane placed at the left (at the right) of the composite curve $\mathscr{B}^{-}\mathscr{B}^{+}\mathscr{E}$ by Ω^{+} (respectively, by Ω^{-}) (Fig. 2).

Suppose that the initial point is on the curve $\mathscr{B}^{-}$ and the players use the controls $v = \beta$ and $u^{-}(x) = \alpha(1 - \beta/x_2)$. Then, the motion goes along $\mathscr{B}^{-}$ and is the slowest motion of system (25) along this curve to the point M. We denote the time of motion from a point $x \in \mathscr{B}^{-}$ to the point M by $\varphi(x)$.

For a point $x \in \Omega^{+} \cup \mathscr{B}^{+}\mathscr{E}$ (resp., $x \in \Omega^{-}$), let the value $\varphi(x)$ be equal to the time of motion of system (25) from the initial point x to the intersection point with the curve $\mathscr{B}^{-}$ (resp., $\mathscr{E}$) for $u = \alpha$ and $v = \beta$ (resp., $u = -\alpha$ and $v = 0$) plus the value of the function $\varphi(\cdot)$ at the intersection point.

Let us underline the principal difference of synthesis corresponding to the game problem (Fig. 2) and the synthesis corresponding to the control problem (Fig. 1). In case of the game problem, the switching control line $\mathscr{E}$ in the lower part is not a parabola arc. It is described by a differential equation, which is not integrable

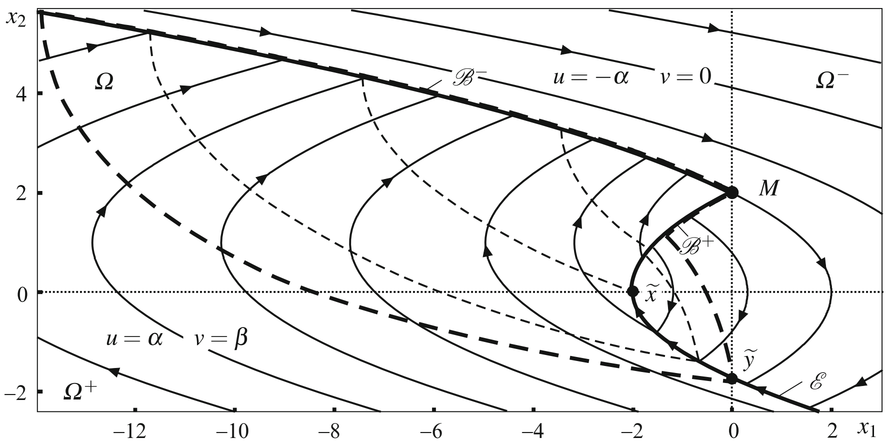


Fig. 2 Construction of the function $\varphi(\cdot)$ for $a = 2, \alpha = \beta = 1$

in explicit form. In addition, the accepted value of the function $\varphi(\cdot)$ on the upper parabola $\mathscr{B}^-$ corresponds to the "slowest" motion of the system along this curve for the control $v = \beta$, although the parabola $\mathscr{B}^-$ becomes a trajectory for $u = -\alpha$ and $v = 0$.

Let $\widetilde{y}$ be an intersection point of the curve $\mathscr{E}$ with the line $x_1 = 0$. Consider the function $\varphi(\cdot)$ on a bounded closed set $\Omega = \{x \in \mathbb{R}^2 : \varphi(x) \leq \varphi(\widetilde{y})\}$, on which the function is discontinuous on the arc $\overline{(\mathscr{B}^+ \cap \operatorname{int} \Omega)} \setminus \{\widetilde{x}\}$. In Fig. 2, a sketch boundary of the set Ω is shown by the thick dashed line. The figure also shows sketches of the level lines of the function $\varphi(\cdot)$ on Ω.

5.3 Verification of the Conditions

Let us show that the function $\varphi(\cdot) : \Omega \to [0, \varphi(\widetilde{y})]$ satisfies the conditions of the theorem.

The way of definition of the function $\varphi(\cdot)$ on the sets $\Omega^\pm$ realizes the well-known method of characteristics [17] for construction of twice continuously differentiable solution to the equation

$$H^\pm(x, \nabla\varphi(x)) = -1, \quad x \in \Omega^\pm,$$

with an appropriate boundary condition on the curves $\mathscr{B}^-$ or $\mathscr{E}$. Here,

$$H^+(x, p) = p_1(x_2 - \beta) + p_2\alpha, \quad H^-(x, p) = p_1 x_2 - p_2\alpha.$$

Hamiltonian (2) of the differential game is of the form

$$H(x, p) = |p_1|\beta/2 + p_1(x_2 - \beta/2) - |p_2|\alpha.$$

It is not difficult to verify that

$$H^\pm(x, \nabla\varphi(x)) = H(x, \nabla\varphi(x)) = -1, \quad x \in \Omega^\pm. \tag{26}$$

By definition, the function $\varphi(\cdot) : \Omega \to [0, \infty)$ is lower semicontinuous and $D(0) = M$. Condition (4) of the theorem is obviously fulfilled. Define $\mathscr{T} = \{\varphi(\widetilde{x})\}$. Let us check on conditions (1) – (3) for $t_* \in (0, \Theta) \setminus \mathscr{T}$, where $\Theta = \varphi(\widetilde{y})$.

(1) We have

$$S(t_*) = \{b^-(t_*), b^+(t_*)\}, \quad b^\pm(t_*) = \overline{F}(t_*) \cap \mathscr{B}^\pm.$$

Since the set $S(t_*)$ is finite, condition (1) of the theorem can be checked for each point $\widetilde{b} \in S(t_*)$ separately. The three cases are possible:

A. $\widetilde{b} = b^-(t_*)$, $\quad t_* \in (0, \Theta) \setminus \mathscr{T}$;
B. $\widetilde{b} = b^+(t_*)$, $\quad t_* \in (0, \varphi(\widetilde{x}))$;
C. $\widetilde{b} = b^+(t_*)$, $\quad t_* \in (\varphi(\widetilde{x}), \Theta)$.

Let us give a proof only for case **A**. Assume $\widetilde{b} = b^-(t_*)$, $t_* \in (0, \Theta) \setminus \mathscr{T}$. Choose such a value $\varepsilon_0 > 0$ that the set $\widetilde{G} = \{\widetilde{b}\} + O(\varepsilon_0)$ can be represented in the form

$$\widetilde{G} = G_+ \cup \Gamma \cup G_\infty,$$

where $\Gamma \subset \mathscr{B}^-$, $\varphi(\cdot) \in C^1(G_+)$, $G_+ \subset \Omega^+$, $G_\infty \subset \Omega^-$ (see the left part of Fig. 3).

(a) We have

$$\widetilde{G} \subset G_\infty \cup \Omega, \quad \lim_{\varepsilon \to 0} \sup\{\varphi(x) : x \in \widetilde{G} \setminus G_\infty\} = t_*.$$

Define the function $\omega(\cdot) : \widetilde{G} \to [0, \infty]$ by formula (10), where the notations $D(t)$ and $G(t, \varepsilon_0)$ are replaced by $D(t_*)$ and $\widetilde{G}$.

Using infinitesimal criteria [5] for stable functions, one can prove u-stability of the function $\omega(\cdot)$ on the set $\widetilde{G} \setminus D(t_*)$.

(b) Define a sequence of the functions

$$\omega_k(\cdot) : \widetilde{G} \to [0, \infty], \quad k \in \mathbb{N},$$

as follows. Let us shift the parabola $\mathscr{B}^-$ for the value $1/k$ to the right along the axis x_1 and denote it by $\mathscr{B}_k^-$. Extend the line $F(t_*)$ by a smooth line l into the set G_∞ in such a way that trajectories of the first family continued above the line $\mathscr{B}^-$ intersect the curve l.

The line $F(t_*)$ has a vertical tangent at the point $\widetilde{b}$ (it can be shown by investigation the vector $\nabla\varphi(x)$, $x \in \Omega^+$). Let us take a segment of the vertical straight line as l. Let us denote by l_k the parabola of the first family, which passes through the point $l \cap \mathscr{B}_k^-$ (see the right part of Fig. 3).

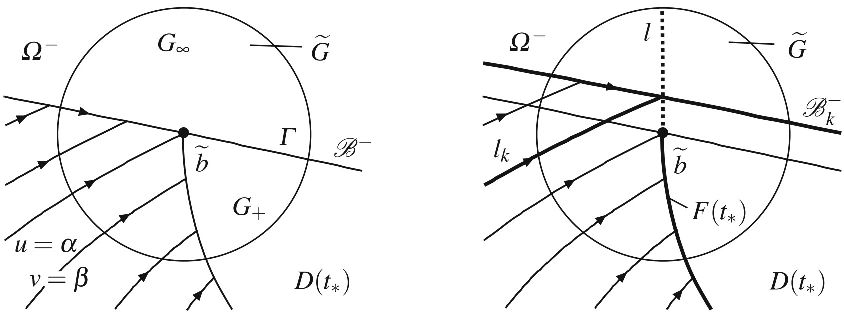


Fig. 3 Checking the conditions of the theorem in case **A**

Define $\omega_k(x) = \infty$ for the points $x \in \widetilde{G}$, which lie above the parabola $\mathscr{B}_k^-$ inclusively; and $\omega_k(x) = 0$ for the points $x \in \widetilde{G}$, which lie strictly below the curve $\mathscr{B}_k^-$ and to the right from the curve $F(t_*) \cup l$ inclusively. For a point $x \in \widetilde{G}$, which lies strictly to the left from the curve $F(t_*) \cup l$ and not above the parabola l_k, let us define $\omega_k(x)$ as the time of the motion from the initial point x along a parabola of the first family under $u = \alpha$ and $v = \beta$ until the instant of attaining the set $F(t_*) \cup l$. For a point $x \in \widetilde{G}$, which lies between the curves $\mathscr{B}_k^-$ and l_k, let us define $\omega_k(x)$ as the time of the motion from the initial point x along a parabola of the first family under $u = \alpha$ and $v = \beta$ until the instant of intersecting the curve $\mathscr{B}_k^-$ plus the time of the slowest motion along the parabola $\mathscr{B}_k^-$ from the point of the intersection to the point $l_k \cap \mathscr{B}_k^-$.

The functions $\omega_k(\cdot)$, $k \in \mathbb{N}$, are upper semicontinuous, equal to zero, and continuous at the points of the set $D(t_*) \cap \widetilde{G}$. It is not difficult to see that the following limit relation is fulfilled:

$$\lim_{k \to \infty} \omega_k(x) = \omega(x), \quad x \in \widetilde{G}.$$

Using infinitesimal criteria [5] for stable functions, one can prove v-stability of the function $\omega_k(\cdot)$, $k \in \mathbb{N}$, on the set $\widetilde{G} \setminus D(t_*)$.

(2) For any $\varepsilon > 0$, there exists such a value $\delta > 0$ that the inclusion is true:

$$G^F(t_*, \varepsilon, \delta) \subset \Omega \cap (\Omega^+ \cup \mathscr{E} \cup \Omega^-).$$

The set $G^F(t_*, \varepsilon, \delta)$ for $t_* \in (0, \varphi(\widetilde{x}))$ is shown at the left part of Fig. 4.

From equalities (26), we have that points of the set $\Omega^\pm$ are regular points [18] of the function $\varphi(\cdot)$, and points of the curve $\mathscr{E}$ are simplest singular points of equivocal type [18]. In virtue of the main result of [18], the function $\varphi(\cdot)$ has the properties of u- and v-stability on the set $G^F(t_*, \varepsilon, \delta)$.

(3) For any $\varepsilon > 0$, let us choose such a value $\delta > 0$ that the following relations are fulfilled:

$$G^B(t_*, \varepsilon, \delta) \cap F(t_*) = \varnothing, \quad G^B(t_*, \varepsilon, \delta) \cap \{x \in \mathbb{R}^2 : x_2 = 0\} = \varnothing.$$

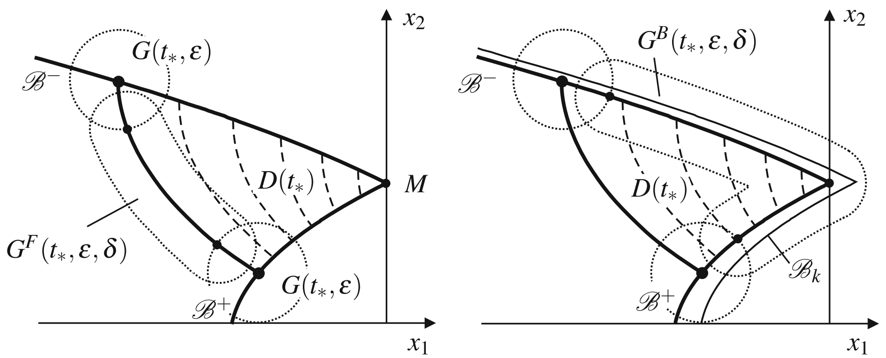


Fig. 4 Checking the conditions of the theorem for the sets $F(t_*)$ and $B(t_*)$

The set $G^B(t_*, \varepsilon, \delta)$ for $t_* \in (0, \varphi(\widetilde{x}))$ is shown at the right part of Fig. 4.

Define a sequence of the functions

$$\omega_k^\infty(\cdot) : G^B(t_*, \varepsilon, \delta) \to [0, \infty], \quad k \in \mathbb{N},$$

as follows. Let us shift the composed curve $\mathscr{B}^-\mathscr{B}^+$ to the right along the axis x_1 for the value $1/k$ and denote it by $\mathscr{B}_k$ (see the right part of Fig. 4). Define $\omega_k^\infty(x) = 0$ at the points $x \in G^B(t_*, \varepsilon, \delta)$, which lies strictly on the left from the curve $\mathscr{B}_k$, and define $\omega_k^\infty(x) = \infty$ for the other points of the set $G^B(t_*, \varepsilon, \delta)$.

One can check v-stability of $\omega_k^\infty(\cdot)$ on the set $G^B(t_*, \varepsilon, \delta) \setminus D(t_*)$.

The following limit relation is obviously true:

$$\lim_{k\to\infty} \omega_k^\infty(x) = \infty, \quad x \in G^B(t_*, \varepsilon, \delta) \setminus D(t_*).$$

Cases **B** and **C** can be analyzed in the same way.

Acknowledgements The work was fulfilled within the framework of the program "Fundamental problems of nonlinear dynamics" (Presidium of the Russian Academy of Sciences) and was supported by the Russian Foundation for Basic Research, project no. 09-01-00436-a.

Appendix

Proof of Lemma 1. Define

$$W_* = \{(t, x) \in [0, \omega(x_*) - t_*] \times G : \omega(x) \le \omega(x_*) - t\}.$$

Note that the set

$$\{(z, x) : \ z = \omega(x_*) - t, \ (t, x) \in W_*\}$$

is a part of epigraph of the function $\omega(\cdot)$, which is sandwiched between two values $z = t_*$ and $z = \omega(x_*)$. Actually,

$$z = \omega(x_*) - t, (t, x) \in W_* \Leftrightarrow z = \omega(x_*) - t, t \in [0, \omega(x_*) - t_*], \omega(x) \le \omega(x_*) - t$$
$$\Leftrightarrow \quad z \in [t_*, \omega(x_*)], \quad z \ge \omega(x).$$

Since the function $\omega(\cdot)$ is u-stable, the set W_* is u-stable with respect to the set $\widetilde{D}$ [1, p. 52], where

$$\widetilde{D} = [0, \omega(x_*) - t_*] \times (D \cup (\mathbb{R}^n \setminus G)).$$

This means that for any point $(t_0, y_0) \in W_*$, any instant $\tau > 0$, and $v_* \in Q$, there exists such a solution $y(\cdot) : [t_0, t_0 + \tau] \to \mathbb{R}^n$ to differential inclusion (3) with the initial condition $y(t_0) = y_0$, that either $(t_0 + \tau, y(t_0 + \tau)) \in W_*$ or $(\tau_*, y(\tau_*)) \in \widetilde{D}$ for some $\tau_* \in [t_0, t_0 + \tau]$.

We have $(0, x_*) \in W_*$ and

$$\{x \in G : (\omega(x_*) - t_*, x) \in W_*\} = D \cap G.$$

Since admissible trajectories with the initial point x_* do not leave G on the interval $[0, \omega(x_*) - t_*]$, the first player's strategy of extremal aiming to the set W_* [1, p. 57] keeps a trajectory of the system in W_* and, therefore, guarantees intersection with the set D on $[0, \omega(x_*) - t_*]$. This implies inequality $\omega(x_*) - t_* \geq T(x_*; D)$. □

Proof of Lemma 2. Define

$$W_\vartheta^* = \{(t, x) \in [0, \vartheta] \times G : \widetilde{\omega}(x) \geq \widetilde{\omega}(x_*) - t\}.$$

Note that the set

$$\{(z, x) : \ z = \widetilde{\omega}(x_*) - t, \ (t, x) \in W_\vartheta^*\}$$

is a part of hypograph of the function $\widetilde{\omega}(\cdot)$, which is sandwiched between two values $z = \widetilde{\omega}(x_*) - \vartheta$ and $z = \widetilde{\omega}(x_*)$. Actually,

$$\begin{aligned} z = \widetilde{\omega}(x_*) - t, (t, x) \in W_\vartheta^* \Leftrightarrow z = \widetilde{\omega}(x_*) - t, \quad t \in [0, \vartheta], \quad \widetilde{\omega}(x) \geq \widetilde{\omega}(x_*) - t \\ \Leftrightarrow \quad z \in [\widetilde{\omega}(x_*) - \vartheta, \widetilde{\omega}(x_*)], \quad z \leq \widetilde{\omega}(x). \end{aligned}$$

Since the function $\widetilde{\omega}(\cdot)$ is v-stable, the set W_ϑ^* is v-stable set [1, p. 52], i.e., for any point $(t_0, y_0) \in W_\vartheta^*$, any instant $\tau > 0$, and $u_* \in P$ there exists such a solution $y(\cdot) : [t_0, t_0 + \tau] \to \mathbb{R}^n$ to differential inclusion (7) with the initial condition $y(t_0) = y_0$ that inequality $t_0 + \tau \leq \vartheta$ implies $(t_0 + \tau, y(t_0 + \tau)) \in W_\vartheta^*$. Continuity of the function $\widetilde{\omega}(\cdot)$ at the points of the set $D \cap G$ implies that the set W_ϑ^* is closed.

We have $(0, x_*) \in W_\vartheta^*$. Since $(\vartheta, x) \in W_\vartheta^*$ implies $\widetilde{\omega}(x) \geq \widetilde{\omega}(x_*) - \vartheta$ and the inequality $\widetilde{\omega}(x_*) - t_* > \vartheta$ is true, the following relation holds:

$$\{x \in G : (\vartheta, x) \in W_\vartheta^*\} \cap D = \varnothing.$$

Since admissible trajectories with the initial point x_* do not leave G on $[0, \vartheta]$, the second player's strategy of extremal aiming to the set W_ϑ^* [1, p. 57] keeps a trajectory of the system in W_ϑ^* and, therefore, guarantees avoidance of intersection with the set D on $[0, \vartheta]$. This implies inequality $\vartheta < T(x_*; D)$. □

References

1. Krasovskii, N.N., Subbotin, A.I.: Positional Differential Games. Nauka, Moscow (1974) (in Russian)
2. Krasovskii, N.N., Subbotin, A.I.: Game-Theoretical Control Problems. Springer, New York (1988)

3. Isaacs, R.: Differential Games. Wiley, New York (1965)
4. Bardi, M., Capuzzo-Dolcetta, I.: Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations. Birkhäuser, Boston (1997)
5. Subbotin, A.I.: Generalized Solutions of First-Order PDEs: the Dynamical Optimization Perspective. Birkhäuser, Boston (1995)
6. Subbotin, A.I.: Minimax and Viscosity Solutions of Hamilton – Jacobi Equations. Nauka, Moscow (1991) (in Russian)
7. Plaskacz, S., Quincampoix, M.: Discontinuous Mayer problem under state-constraints. Topol. Methods Nonlinear Anal. **15**, 91–100 (2000)
8. Serea, O.S.: Discontinuous differential games and control systems with supremum cost. J. Math. Anal. Appl. **270**(2), 519–542 (2002)
9. Cannarsa, P., Frankowska, H., Sinestrari, C.: Optimality conditions and synthesis for the minimum time problem. Set-Valued Anal. **8**, 127–148 (2000)
10. Bardi, M., Bottacin, S., Falcone, M.: Convergence of discrete schemes for discontinuous value functions of Pursuit-Evasion Games. In: Olsder, G.J. (ed.) New Trends in Dynamic Games and Applications, Annals of ISDG **3** pp. 273-304. Birkhäuser, Boston (1995)
11. Kruzhkov, S.N.: Generalized solutions of the Hamilton-Jacobi equations of eikonal type. I. Math. USSR Sb. **27**(3), 406–446 (1975)
12. Kamneva, L.V.: The conditions for a discontinuous function to be identical with the value function of a game in a time-optimal problem. J. Appl. Math. Mech. **70**, 667–679 (2006)
13. Fedorenko, R.P.: On the Cauchy problem for Bellman's equation of dynamic programming. Ž. Vyčisl. Mat. i Mat. Fiz. **9**(2), 426–432 (1969) (in Russian)
14. Leitmann, G.: The Calculus of Variations and Optimal Control. Plenum, New York etc. (1981)
15. Patsko, V.S.: A model example of game pursuit problem with incomplete information. I, II. Differents. Uravneniya **7**(3), 424–435 (1971); **8**(8), 1423–1434 (1972) (in Russian)
16. Filimonov, M. Yu.: Conjugacy of singular lines in a differential game. In: Investigations of minimax control problems, pp. 117–124. Akad. Nauk SSSR, Ural. Nauchn. Tsentr, Sverdlovsk (1985) (in Russian)
17. Courant, R., Hilbert, D.: Methoden der Mathematischen Physik. Springer, Berlin (1931)
18. Kamneva, L.V.: The sufficient conditions of stability for the value function of a differential game in terms of singular points. J. Appl. Math. Mech. **67**(3), 329–343 (2003)

On Grid Optimal Feedbacks to Control Problems of Prescribed Duration on the Plane

Nina N. Subbotina and Timofey B. Tokmantsev

Abstract We consider optimal control problems of prescribed duration. A new numerical method is suggested to solve the problems of controlling until a given instant. The solution is based on a generalization of the method of characteristics for Hamilton–Jacobi–Bellman equations. Constructions of optimal grid synthesis are suggested and numerical algorithms solving the problems on the plane are created. Efficiency of the grid feedback is estimated. Results of simulations using the numerical algorithms are exposed.

1 Introduction

We consider nonlinear optimal control problems, where values of controls are restricted by geometric constraints. The quality of an admissible open-loop control is equal to the minimum of a running cost of the Bolza type with the upper time limit varying in the time interval of prescribed duration. The goal is to minimize the quality. The problems are actual to the optimal control theory and applications, see, for example, [1–3, 5, 6, 11, 13].

A new numerical method is suggested to solve the problems in the class of feedbacks. We construct a universal grid feedback, which is near optimal for all initial phase states of a controlled system.

A universal optimal feedback is called the optimal synthesis. Numerous examples in references mentioned above show that the optimal synthesis is discontinuous, as a rule. The basis for constructions of optimal feedbacks is the value function, which brings the correspondence between any initial phase state of the controlled system and the optimal result (minimal cost) at this state [1, 2, 5, 6, 11].

N.N. Subbotina (✉) and T.B. Tokmantsev
Institute of Mathematics and Mechanics UrB RAS, S. Kovalevskaja str., 16, Ekaterinburg, 620990, Russia
e-mail: subb@uran.ru; tokmantsev@imm.uran.ru

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_7,

Therefore, we first have to calculate the nonsmooth value function. It has been shown [12] that the value function in optimal control problems of prescribed duration is the unique minimax/viscosity solution of the Cauchy problem for the Bellman equation, with additional restrictions on values. The value function has a representation in terms of solutions of the Hamiltonian differential inclusions. The result is based on necessary and sufficient optimality conditions of the first order [11]. Hence, the value function can be constructed via a backward procedure of integration of the Hamiltonian differential inclusions [3, 9, 14, 15]. Solutions of the Hamiltonian inclusions are called the generalized characteristics. An advantage of the reconstruction of the value function is that the phase and conjugate characteristics have boundary conditions at the same instants.

The second part, and an additional advantage of the backward procedure is the reconstruction of optimal grid feedbacks during the procedure. Note that the procedure provides motions of the phase and conjugate characteristics, which coincide with optimal trajectories of the controlling system and evaluations of supergradients of the value function, respectively. Optimal feedbacks are aimed along the supergradients (see [24]).

Algorithms realizing the procedure on the plane are proposed in the paper [24]. In contrast with known grid methods for uniform grids in the phase space(see, for example, [21, 26]), the constructions below deal with adaptive grids. Efficiency of the grid feedbacks to optimal control problems of prescribed duration can be estimated in the same manner as in the paper [25], where optimal control problems are considered for the Bolza costs with fixed upper limits. The estimations show that applications of adaptive grids to approximations of the value function provide the speed of convergence better than that for uniform grids.

These researches used the approach to constructions and applications of feedbacks to problems of optimal control, observations, estimations, positional reconstruction suggested by N.N. Krasovskii [6] and developed by his scientific school [7,8,10,13,17–19,22,23]. A key notion in researchers is the cost of discontinuous feedbacks. A second important point is the notion of generalized characteristics for the Bellman equation [14]. The suggested construction of optimal grid synthesis is based on known results of the modern theory of minimax/viscosity solutions to Hamilton–Jacobi equations [1, 5, 12, 16].

The proposed grid scheme for solving optimal control problems can be applied to models arising in mechanics, mathematical economics, biology, and other areas.

2 Control Problems of Prescribed Duration

2.1 Statement of the Problem

We consider a controlled system:

$$\begin{aligned} \dot{x}(t) &= f(t, x, u), \quad t \in [0, T],\ x \in R^n, \\ x(t_0) &= x_0, \quad (t_0, x_0) \in [0, T] \times R^n = \mathrm{cl}\, \Pi_T, \end{aligned} \tag{1}$$

where controls u belong to a given compact set $P \subset R^m$ and a time interval $[0, T]$ is fixed. Let us define the sets $U_{t_0,T}$ of admissible controls as follows

$$U_{t_0,T} = \{u(\cdot)\colon [t_0, T] \mapsto P \text{ — is measurable}\}.$$

Let the cost functional be of the Bolza type:

$$I\big(t_0,\ x_0; u(\cdot)\big) = \min_{\theta\in[t_0,T]} \left\{ \sigma\big(\theta,\ x(\theta)\big) + \int\limits_{\theta}^{t_0} g(t, x(t), u(t))\mathrm{d}t \right\}, \tag{2}$$

where $x(\cdot) = x\big(\cdot; t_0, x_0, u(\cdot)\big)\colon [t_0, T] \mapsto R^n$ is a trajectory of the system (1) started at the state $x(t_0) = x_0$ under an admissible control $u(\cdot)$. We define the optimal result as follows:

$$V(t_0, x_0) = \inf_{u(\cdot)\in U_{t_0,T}} I\big(t_0,\ x_0; u(\cdot)\big). \tag{3}$$

The function

$$\operatorname{cl} \varPi_T \ni (t_0, x_0) \mapsto V(t_0, x_0) \in R$$

is called the value function of the problem (1)–(3).

We shall solve the problem (1)–(3) in the set of all feedbacks $(t, x) \to u(t, x)$: $\operatorname{cl} \varPi_T \mapsto P$ and consider discontinuous feedbacks, too. Let us recall the notion of cost $C\big(t_0, x_0; u(t, x)\big)$ of a feedback $u(t, x)$ at an initial state $(t_0, x_0) \in \operatorname{cl} \varPi_T$ (see [6]).

Definition 1. The cost $C\big(t_0, x_0; u(t, x)\big)$ of a feedback $u(t, x)$ at an initial state $(t_0, x_0) \in \operatorname{cl} \varPi_T$ is defined as follows

$$C\big(t_0, x_0; u(t, x)\big) = \limsup_{\operatorname{diam}(\varGamma)\to 0} C_\varGamma\big(t_0, x_0;\ u(t, x)\big),$$

where

$$\begin{aligned}
C_\varGamma\big(t_0, x_0;\ u(t, x)\big) &= I\big(t_0,\ x_0; u_\varGamma(\cdot)\big),\\
\varGamma &= \{\, t_i,\ i = 0, 1, \dots, N\} \subset [t_0, T = t_N],\\
\operatorname{diam}(\varGamma) &= \max_{i=1,\dots,N} (t_i - t_{i-1});\\
u_\varGamma(t) &= u_{i-1} = u\big(t_{i-1}, x_\varGamma(t_{i-1})\big) = \text{const}, \quad \forall t \in [t_{i-1}, t_i);\\
\dot{x}_\varGamma(t) &= f\big(t, x_\varGamma(t), u_{i-1}\big), \quad \forall t \in [t_{i-1}, t_i), \quad i \in \overline{1, N};\\
x_\varGamma(t_0) &= x_0.
\end{aligned}$$

We shall construct a feedback $u^0(t, x)$, which is optimal for all initial states (t_0, x_0).

Definition 2. A feedback $(t,x) \mapsto u^0(t,x)$ is called the optimal synthesis to the problem (1)–(3) if the relations

$$C\big(t_0, x_0; u^0(t,x)\big) = V(t_0, x_0), \quad \forall (t_0, x_0) \in \operatorname{cl} \Pi_T.$$

are true.

2.2 Assumptions

We assume that data of problem (1)–(3) satisfy the following conditions.

A1. Functions $f(t,x,u)$ and $g(t,x,u)$ in (1), (2) are Lipschitz continuous on the set $\operatorname{cl} \Pi_T$, i.e.

$$\begin{aligned} ||f(t',x',u) - f(t'',x'',u)|| &\le L_1(|t'-t''| + ||x'-x''||), \\ |g(t',x',u) - g(t'',x'',u)| &\le L_1(|t'-t''| + ||x'-x''||), \end{aligned}$$

where $L_1 > 0,\ (t',x'),\ (t'',x'') \in \operatorname{cl} \Pi_T,\ u \in P.$

A2. The function $\sigma(t,x)$ in (2) is Lipschitz continuous on $[0,T] \times R^n$; for each $(t,x) \in \operatorname{cl} \Pi_T$, there exists the superdifferential $\partial\sigma(t,x)$:

$$\partial\sigma(t,x) = \{(a,p) \in R \times R^n: \quad \forall (t+\Delta t, x+\Delta x) \in B_\varepsilon(t,x), \sigma(t+\Delta t, x+\Delta x) \\ -\sigma(t,x) \le a\Delta t + < p, \Delta x > +o(|\Delta t| + ||\Delta x||)\},$$

where $\varepsilon > 0$, $o(|\Delta t| + ||\Delta x||)/(|\Delta t| + ||\Delta x||) \to 0$, as $|\Delta t| + ||\Delta x|| \to 0$, and

$$\exists L_2 > 0: \quad |\alpha| + ||p|| \le L_2 \quad \forall (\alpha, p) \in \partial\sigma(t,x).$$

A3. For all $(t,x) \in \operatorname{cl} \Pi_T,\ p \in R^n$, the set

$$\operatorname{Arg} \min_{(f,g) \in E(t,x)} [\langle p, f\rangle + g] = \{\big(f^0(t,x,p), g^0(t,x,p)\big)\}$$

is a singleton. The symbol $\langle \cdot, \cdot \rangle$ denotes the inner product, and

$$E(t,x) = \{\big(f(t,x,u), g(t,x,u)\big): u \in P\}.$$

It is known [3, 12] that assumptions A1–A2 provide local Lipschitz continuity of the value function $V(t,x)$ (3). Condition A2 is important also for existence of the superdifferential of the value function $\partial V(t,x)$ at each point $(t,x) \in \Pi_T = (0,T) \times R^n$ [14]. Condition A3 implies existence of an admissible measurable optimal open-loop control $u^0(\cdot) \in U_{t_0,T}$, and provides the construction of optimal synthesis below.

2.3 Preliminaries

It is proven in the monograph [12] that the value function $V(t, x)$ (3) coincides with the unique minimax/viscosity solution of the Cauchy problem for the Bellman equation

$$\partial V(t,x)/\partial t + \min_{u \in P}[\langle D_x V(t,x), f(t,x,u)\rangle + g(t,x,u)] = 0, \tag{4}$$

$$(t,x) \in \Pi_T = (0,T) \times R^n;$$

$$V(T,x) = \sigma(T,x), \quad \forall x \in R^n, \tag{5}$$

with the additional restrictions

$$V(t,x) \le \sigma(t,x), \quad \forall (t,x) \in \mathrm{cl}\,\Pi_T = [0,T] \times R^n. \tag{6}$$

Here,

$$D_x V(t,x) = \left(\frac{\partial V(t,x)}{\partial x_1}, \ldots, \frac{\partial V(t,x)}{\partial x_n}\right).$$

Existence of the superdifferential $\partial V(t,x)$ at each point $(t,x) \in \Pi_T$ [14], and known results of nonsmooth analysis [3,4] imply that the unique minimax/viscosity solution $V(t,x)$ of boundary problem (4)–(6) satisfies the following equality

$$\min_{(\alpha,p) \in \partial V(t,x)} \min_{u \in P} [\langle (\alpha,p), (1, f(t,x,u))\rangle + g(t,x,u)] = 0$$

at each point $(t,x) \in \Pi_T$, where $V(t,x) < \sigma(t,x)$. Note that the equality turns off to the Bellman equation (4) at all points (t,x), where the value function $V(t,x) \neq \sigma(t,x)$ is differentiable and, hence, the set $\partial V(t,x)$ is the singleton

$$\partial V(t,x) = \{(\partial V(t,x)/\partial t, \quad D_x V(t,x))\} \in R^{n+1}.$$

According to the Rademacher theorem [20], the local Lipschitz continuous value function $V(t,x)$ is differentiable almost everywhere in $\mathrm{cl}\,\Pi_T$.

2.4 The Method of Characteristics

It follows from assumption A3 that the Hamiltonian $H(t,x,p)$ in problem (1)–(3) has the form

$$H(t,x,p) = \min_{u \in P}\{\langle p, \ f(t,x,u)\rangle + g(t,x,u)\} = \langle p, \ f^0(t,x,p)\rangle + g^0(t,x,p),$$

and the following relations hold

$$D_p H(t,x,p) = f^0(t,x,p),$$

$$\langle p, D_p H(t,x,p)\rangle - H(t,x,p) = -g^0(t,x,p),$$

where $\big(f^0(t,x,p), g^0(t,x,p)\big) \in E(t,x)$. The Hamiltonian is Lipschitz continuous relative to (t,x) and continuous in the whole space. Recall that a Lipschitz continuous function is differentiable almost everywhere.

Consider a generalized characteristic system for the Bellman equation (4) involving the Clarke's superdifferential $\partial_x^{\mathrm{cl}} H(t,x,p)$ of $H(t,x,p)$ with respect to variable x [3,20]:

$$\partial_x^{\mathrm{cl}} H(t,x,p) = \mathrm{co}\left\{ \lim_{x_i \to x} D_x H(t,x_i,p) \right\},$$

where (t,x_i,p) are the points of differentiability for $H(t,\cdot,p)$.

We introduce the generalized characteristic system for the Bellman equation (4) as the system of the Hamiltonian differential inclusions

$$\begin{cases} \dfrac{\mathrm{d}\widehat{x}}{\mathrm{d}t} = & D_{\widehat{p}} H(t,\widehat{x},\widehat{p}) = f^0(t,\widehat{x},\widehat{p}), \\ \dfrac{\mathrm{d}\widehat{p}}{\mathrm{d}t} \in & -\partial_{\widehat{x}}^{\mathrm{cl}} H(t,\widehat{x},\widehat{p}), \\ \dfrac{\mathrm{d}\widehat{z}}{\mathrm{d}t} = & \langle \widehat{p}, D_{\widehat{p}} H(t,\widehat{x},\widehat{p})\rangle - H(t,\widehat{x},\widehat{p}) = -g^0(t,\widehat{x},\widehat{p}). \end{cases} \tag{7}$$

Due to boundary condition (5) and restriction (6), the following boundary conditions are considered:
Boundary conditions

$$\begin{aligned} &\widehat{x}(T,y) = y,\ \widehat{p}(T,y) \in \partial_y \sigma(T,y),\ \widehat{z}(T,y) = \sigma(T,y),\ y \in R^n, \\ &\min_{y^* \in R^n : \widehat{x}(t,y^*) = x} \widehat{z}(t,y^*) = \sigma(t,x) \Rightarrow \quad \widehat{p}(t,y^*) \in \partial_y \sigma(t,x). \end{aligned} \tag{8}$$

Here

$$\partial_y \sigma(t,x) = \{p \in R^n : \quad \exists\, \alpha \in R,\ (\alpha,p) \in \partial\sigma(t,x)\}.$$

Definition 3. Absolutely continuous functions

$$\widehat{x}(\cdot,y),\ \widehat{p}(\cdot,y),\ \widehat{z}(\cdot,y): \quad [0,T] \to R^n \times R^n \times R$$

satisfied Hamiltonian differential inclusions (7) and boundary conditions (8) are called the generalized characteristics.

It is known [3,11,15] that necessary optimality conditions in problem (1)–(3) can be expressed in the form of the Hamiltonian differential inclusions. It means that extremals and coextremals for each initial point $(t_0,x_0) \in \mathrm{cl}\,\Pi_T$ can be considered as generalized characteristics $\widehat{x}(\cdot,y^*)$, $\widehat{p}(\cdot,y^*)$ (7), (8), respectively, that cross at the initial state, i.e. $\widehat{x}(t_0,y^*) = x_0$. The following theorems (see [14, 24]) can be considered as a generalization of the classical method of characteristics.

Theorem 1. *If conditions* A1–A3 *are satisfied in problem (1)–(3), then the value function* $(t_0, x_0) \mapsto V(t_0, x_0)$ *has the representation*

$$V(t_0, x_0) = \min\left\{\sigma(t_0, x_0),\ \min_{y\in R^n : \widehat{x}(t_0,y)=x_0} \widehat{z}(t_0, y)\right\}$$
$$= \min\left\{\sigma(t_0, x_0),\ \min_{y\in R^n : \widehat{x}(t_0,y)=x_0} \left\{\sigma(T, y) + \int_{t_0}^{T} g^0\big(\tau, \widehat{x}(\tau, y), \widehat{p}(\tau, y)\big)\mathrm{d}\tau\right\}\right\},$$

where $\big(\widehat{x}(t_0, y), \widehat{p}(t_0, y), \widehat{z}(t_0, y)\big)$ *are generalized characteristics (7), (8).*

Theorem 2. *If conditions* A1–A3 *are satisfied in problem (1)–(3), then an optimal synthesis* $u^0(t, x)$: $\mathrm{cl}\, \Pi_T \mapsto \mathrm{P}$ *can be defined as follows*

$$(t, x) \to\ u^0(t, x) \in \operatorname*{Arg}_{u\in P}\{f(t, x, u) = f^0,\ g(t, x, u) = g^0\},$$

where the relations

$$(f^0, g^0) = \big(f^0(t, x, p_*), g^0(t, x, p_*)\big) \in E(t, x),$$
$$H(t, x, p_*) = \langle p_*, f^0(t, x, p_*)\rangle + g^0(t, x, p_*),$$
$$(-H(t, x, p_*),\ p_*) \in \partial V(t, x),$$

are true , for a vector $p_* \in R^n$.

Note that the suggested below numerical algorithms solving optimal control problems (1)–(3) are the realizations of Theorems 1–2.

3 Algorithms

We introduce a numerical method solving problem (1)–(3), which consists of two parts: a backward procedure constructing an approximation $\widetilde{V}(t, x)$ for the value function, and a construction of an optimal grid synthesis $u^0(t_i, \widehat{x}_i^j)$.

We consider a partition $\Gamma = \{t_0 < t_1 < \ldots < t_N = T\} \subset [0, T]$ with step $\Delta t = \dfrac{T - t_0}{N}$. Let $D \subset R^n$ be a compact domain in the phase space. Let $Q_N = \{y^j \in D\}$, $j = (j_1, \ldots, j_n)$ be a uniform grid for phase variable x with step Δx.

Consider a time interval $[t_i, t_{i+1}]$. Let the symbol $Q_{i+1} \subset R^n$ mean the current grid of nodes on phase variables $\widehat{x}(\cdot, y^j)$ of generalized characteristics at instant $t = t_{i+1}$, $i \in \overline{N-1, 0}$. Let $\big(\widehat{x}^j(\cdot), \widehat{p}^j(\cdot), \widehat{z}^j(\cdot)\big) : [t_i, t_{i+1}] \mapsto R^n \times R^n \times R$ be a numerical solution of generalized characteristic system (7), which satisfies the boundary conditions

$$\widehat{x}^j(t_{i+1}) = y^j \in Q_{i+1},\ \widehat{p}^j(t_{i+1}) = p^j,\ \widehat{z}^j(t_{i+1}) = z^j = \widetilde{V}(t_{i+1}, y^j);$$
$$\widehat{x}^j(t_N) = y^j \in Q_N,\ \widehat{p}^j(t_N) = p^j \in \partial_x\sigma(t_N, y^j),\ \widehat{z}^j(t_N) = z^j = \sigma(t_N, y^j). \quad (9)$$

We consider results of a backward integrating procedure at instant t_i, and denote the results by $(\widehat{x}_i^j, \widehat{p}_i^j, \widehat{z}_i^j) = \big(\widehat{x}^j(t_i), \ \widehat{p}^j(t_i), \widehat{z}^j(t_i)\big)$. We form the next current grid Q_i as the set of points $\widehat{x}_i^j$.

Let $\widehat{x}_i^{j^0}$ be a node of the grid Q_i. According to [14], we consider the approximation of the value function:

$$\widetilde{V}(t_i, \widehat{x}_i^{j^0}) = \min\left\{\sigma(t_i, \widehat{x}_i^{j^0}) - \rho_z, \min_{j: \|\widehat{x}_i^j - \widehat{x}_i^{j^0}\| \le \rho_x} \widehat{z}_i^j\right\}. \tag{10}$$

Here, $i \in \overline{N-1, 0}$, and $\rho_x > 0, \ \rho_z > 0$ are approximation parameters. The inequality $\|\widehat{x}_i^j - \widehat{x}_i^{j^0}\| \le \rho_x$ in (10) means that the numerical approximations of phase characteristics $\widehat{x}(\cdot, y^j)$ and $\widehat{x}(\cdot, y^{j^0})$ cross at instant t_i. Parameter ρ_z estimates accuracy of approximations of restriction (6) $V(t, x) = \sigma(t, x)$.

We form boundary conditions (9) at instant t_i for the next step of the backward procedure on interval $[t_{i-1}, t_i]$ in the following way

$$\begin{aligned}
&\widehat{x}^{j^0}(t_i) = \widehat{x}_i^{j^0} = y^{j^0} \in Q_i; \quad z^{j^0} = \widehat{z}^{j^0}(t_i) = \widetilde{V}(t_i, \widehat{x}^{j^0});\\
&p^{j^0} = \widehat{p}^{j^0}(t_i) = \widehat{p}_i^j, \quad \text{if} \quad \widetilde{V}(t_i, \widehat{x}_i^{j^0}) = \widehat{z}_i^j, \quad \text{otherwise } \ p^{j^0} \in \partial\sigma(t_i, \widehat{x}_i^{j^0}).
\end{aligned}$$

Optimal grid synthesis $u^0(t_i, \widehat{x}_i^{j^0})$ is defined according formulae in Theorem 2, where we put $p_* = p^{j^0}$.

4 Estimations

We estimate efficiency of the optimal grid synthesis in a compact domain $G \subset \operatorname{cl} \varPi_T$, which contains all nodes (t_i, y^j): $y^j = \widehat{x}_i^j \in Q_i, t_i \in \varGamma$. We assume that

$$V(t, x) \le \min\{\widehat{z}(t, y) \colon y \in D, \ \widehat{x}(t, y) = x\} \quad \forall (t, x) \in G.$$

Let us consider an initial state $(t_0, x_0) \in G$ and define the realization $u_*^0(\cdot)$ of the grid feedback $u^0(t_i, \widehat{x}_i^j)$. Let

$$\begin{aligned}
&x^{j^0} \in \operatorname{Arg} \min_{x^j \in Q_0} \|x_0 - x^j\|;\\
&\dot{x}_\varGamma(t) = f\big(t, x_\varGamma(t), u_{i-1}\big), \quad \forall t \in [t_{i-1}, t_i), \quad x_\varGamma(t_0) = x^{j^0};\\
&u_{i-1} = u^0\big(t_{i-1}, x_\varGamma(t_{i-1})\big),\\
&u_\varGamma^0(t) = u_{i-1} = \text{const}, \quad \forall t \in [t_{i-1}, t_i), \quad i = 1, \dots, N;\\
&\dot{x}_\varGamma^0(t) = f\big(t, x_\varGamma^0(t), u_{i-1}^0\big), \quad \forall t \in [t_{i-1}, t_i), \quad x_\varGamma^0(t_0) = x_0;
\end{aligned}$$

$$u^0_{i-1} \in \operatorname{Arg}\min_{u\in P} \| f\big(t_{i-1}, x_\Gamma(t_{i-1}), u_{i-1}\big) - f\big(t_{i-1}, x^0_\Gamma(t_{i-1}), u\big)\|;$$
$$u^0_*(t) = u^0_{i-1} = \text{const}, \quad \forall t \in [t_{i-1}, t_i], \quad i = 1, \dots, N.$$

Note that the motion $x_\Gamma(\cdot)$ is reconstructed as an approximation of an optimal trajectory $x^0(\cdot)$, $x^0(T) \in D$.

The following estimate can be obtained for the difference between the optimal result $V(t_0,\ x_0)$ and the cost $\widetilde{C}_\Gamma\big(t_0, x_0; u^0_*(\cdot)\big) = I\big(t_0,\ x_0; u^0_*(\cdot)\big)$ of the grid feedback $u^0(t_i, \widehat{x}^j_i)$:

$$|V(t_0,\ x_0) - \widetilde{C}_\Gamma\big(t_0, x_0; u^0_*(\cdot)\big)| \le C_1\Delta t + C_2\Delta t\omega(F\Delta t), \tag{11}$$

where F, C_1, C_2 are constants, Δt is the time-step of Γ, $\omega(\cdot)$ is the modulus of continuity for functions $f^0(t, x, p)$, $g^0(t, x, p)$ defined in A3, for $(t, x) \in G$, $\|p\| \le L_V$; $L_V > 0$ is the Lipschitz constant of the value function $V(t, x)$ in G.

The inequality follows from the corresponding estimates obtained in the paper [25] for optimal control problems with the Bolza functional (2), where $\theta \equiv T$.

It is easy to see that estimate (11) is the sum of three terms $\Delta I_1, \Delta I_2, \Delta I_3$:

$$\Delta I_1 = |V(t_0, x_0) - V(t_0, x^0(t_0))| \le L_V \|x^0(t_0) - x_0\|;$$
$$\Delta I_2 = |I_{t_0, x^0(t_0)}(x^0(\cdot), u^0(\cdot)) - I_{t_0, x^{j^0}}(x_\Gamma(\cdot), u^0_\Gamma(\cdot))|;$$
$$\Delta I_3 = |I(t_0, x^{j^0}; u^0_\Gamma(\cdot)) - I\big(t_0,\ x_0; u^0_*(\cdot)\big)|.$$

The estimate for the first term ΔI_1 is based on local Lipschitz continuity of the value function $V(t, x)$. The estimate ΔI_3 can be obtained by means of estimations for difference between two motions of system (1) started at $t = t_0$ from points x_0 and x^{j^0} and generated by the same control $u^0_*(\cdot)$ on the interval $[t_0, T]$. The estimate ΔI_2 for difference between an optimal characteristic and its numerical approximation can be obtained recursively with the help of a backward procedure of integrating characteristic system from the state

$$x^{j^0}(T) = y^{j^0} \in Q_N, \quad p^{j^0}(T) = p^{j^0} \in \partial\sigma(t, y^{j^0}), \quad z^{j^0}(T) = z^{j^0} = \sigma(t, y^{j^0}),$$

where necessary and sufficient optimality condition are applied. See [25] for details.

5 Examples

We implemented the numerical algorithms realizing the suggested numerical method to the following model optimal control problems on the plane.

5.1 Example 1

Dynamics

$$\begin{aligned}\dot{x}_1 &= x_2,\\ \dot{x}_2 &= u,\\ |u| &\le 1;\ t \in [0,1].\end{aligned}$$

Payoff functional

$$I\big(t_0, x_0; u(\cdot)\big) = \min_{\theta\in[t_0,1]} \left\{ \frac{(x_1(\theta)+\theta-1)^2}{2} + \frac{x_2^2(\theta)}{2} + \int_{t_0}^{\theta} u^2(t)\mathrm{d}t \right\}.$$

Hamiltonian

$$H(x,p) = p_1 x_2 - \frac{p_2^2}{4}.$$

Characteristic system

$$\begin{cases} \dfrac{\mathrm{d}\widehat{x}_1}{\mathrm{d}t} = \widehat{x}_2,\ \dfrac{\mathrm{d}\widehat{x}_2}{\mathrm{d}t} = -\dfrac{\widehat{p}_2}{2}, \\ \dfrac{\mathrm{d}\widehat{p}_1}{\mathrm{d}t} = 0,\quad \dfrac{\mathrm{d}\widehat{p}_2}{\mathrm{d}t} = -\widehat{p}_1,\ \dfrac{\mathrm{d}\widehat{z}}{\mathrm{d}t} = -\dfrac{\widehat{p}_2^{\,2}}{4}. \end{cases}$$

Boundary conditions

$$\widehat{x}_1(T,y) = y_1,\ \widehat{x}_2(T,y) = y_2,$$

$$\widehat{p}_1(T,y) = y_1,\ \widehat{p}_2(T,y) = y_2,\ \widehat{z}(T,y) = \frac{y_1^2+y_2^2}{2},$$

$$\widehat{z}(t,y) = \frac{(\widehat{x}_1(t,y)+t-1)^2}{2} + \frac{\widehat{x}_2^2(t,y)}{2} \Rightarrow \begin{cases} \widehat{p}_1(t,y) = \widehat{x}_1(t,y)+t-1, \\ \widehat{p}_2(t,y) = \widehat{x}_2(t,y). \end{cases}$$

The approximation parameters are: $\Delta t = 0.02$, $\rho_x = 0.0004$, $D = [-2,2] \times [-2,2]$, $\Delta x = 0.027$.

The results of simulations are presented in Fig. 1. The functions on the picture are defined at nodes $y^j = (y_1^j, y_2^j)$ of the same adaptive grid $Q_0 \subset R^2$. The numerical approximations $\widetilde{V}(t_0, y^j)$ of the optimal results are reconstructed via the backward procedure described in Sect. 3. The results $\widetilde{C}_\Gamma\big(t_0, y^j; u_*^0(\cdot)\big)$ of realizations $u_*^0(\cdot)$ of the grid feedback $u^0(t_i, \widehat{x}_i^j)$ are obtained via the direct time procedure described in Sect. 4, for all initial states (t_0, y^j), $y^j \in Q_0$.

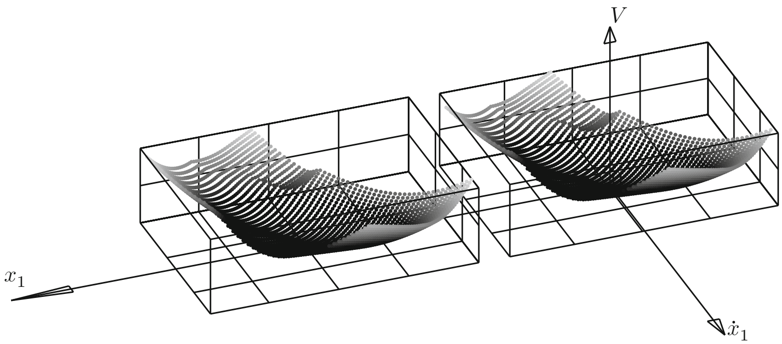


Fig. 1 Graphs of $\widetilde{V}(t,x)$ (*right graph*) and $\widetilde{C}_{\Gamma}\big(t,x;u_*^0(\cdot)\big)$ (*left graph*) at $t=t_0=0$ and $x=y^j$

5.2 Example 2

Nonlinear dynamics

$$\begin{aligned}\dot{x}_1 &= x_2,\\ \dot{x}_2 &= -\sin x_1 + u,\\ |u| &\le 1;\ t\in[1,2].\end{aligned}$$

Payoff functional

$$I\big(t_0,\ x_0;u(\cdot)\big) = \min_{\theta\in[t_0,2]}\left\{-\frac{(x_1(\theta)+\theta-2)^2+x_2^2(\theta)}{2} - \int_{t_0}^{\theta}\sqrt{1-u^2(t)}\mathrm{d}t\right\}.$$

Hamiltonian

$$H(x,p) = p_1x_2 - p_2\sin x_1 - \sqrt{1+p_2^2}.$$

Characteristic system

$$\begin{cases}\dfrac{\mathrm{d}\widehat{x}_1}{\mathrm{d}t} = \widehat{x}_2, & \dfrac{\mathrm{d}\widehat{x}_2}{\mathrm{d}t} = -\sin\widehat{x}_1 - \dfrac{\widehat{p}_2}{\sqrt{\widehat{p}_2^{\ 2}+1}},\\[2ex] \dfrac{\mathrm{d}\widehat{p}_1}{\mathrm{d}t} = \widehat{p}_2\cos\widehat{x}_1, & \dfrac{\mathrm{d}\widehat{p}_2}{\mathrm{d}t} = -\widehat{p}_1, \quad \dfrac{\mathrm{d}\widehat{z}}{\mathrm{d}t} = \dfrac{1}{\sqrt{1+\widehat{p}_2^2}}.\end{cases}$$

Boundary conditions

$$\widehat{x}(T,y) = y,\ \widehat{p}(T,y) = y,\ \widehat{z}(T,y) = -\frac{y_1^2+y_2^2}{2},$$

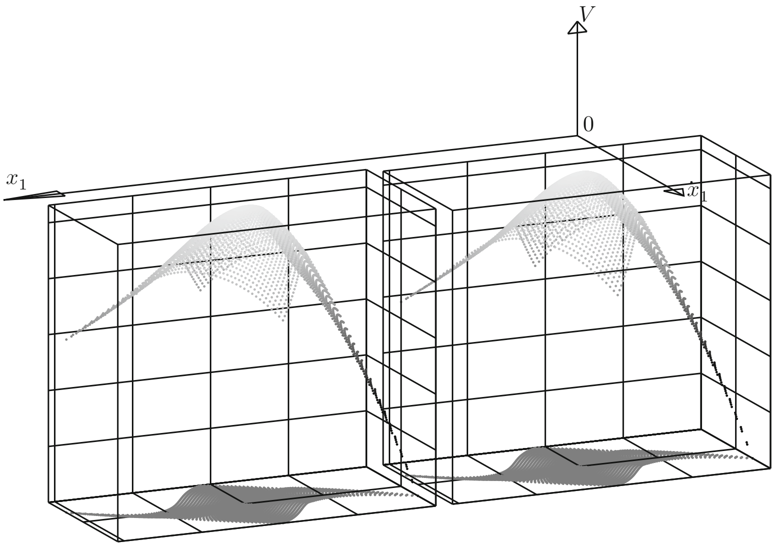


Fig. 2 Graphs of $\widetilde{V}(t,x)$ (*right graph*) and $\widetilde{C}_\Gamma\big(t,x;u_*^0(\cdot)\big)$ (*left graph*) at $t = t_0 = 1$ and $x = y^j$

$$\widehat{z}(t,y) = -\frac{(\widehat{x}_1(t,y)+t-2)^2}{2} - \frac{\widehat{x}_2^2(t,y)}{2}$$
$$\Rightarrow \begin{cases} \widehat{p}_1(t,y) = -(\widehat{x}_1(t,y)+t-2), \\ \widehat{p}_2(t,y) = -\widehat{x}_2(t,y). \end{cases}$$

The approximation parameters are: $\Delta t = 0.02$, $\rho_x = 0.0004$, $D = [-2,2] \times [-2,2]$, $\Delta x = 0.16$. The results of simulations are presented in Fig. 2.

The gray shadows on the plane are nodes y^j of the same adaptive grid Q_0. According to estimate (11), realizations $u_*^0(\cdot)$ of the grid feedback $u^0(t_i,\widehat{x}_i^j)$ provides results $\widetilde{C}_\Gamma\big(t_0,y^j;u_*^0(\cdot)\big)$ close to the numerical approximation $\widetilde{V}(t_0,y^j)$ of optimal results for all initial states (t_0,y^j), $y^j \in Q_0$.

5.3 Example 3

Dynamics

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = u,$$
$$|u| \le 1;\ t \in [0, 0.5].$$

Payoff functional with a nonsmooth terminal function

$$I\big(t_0, x_0; u(\cdot)\big) = \min_{\theta \in [t_0, 0.5]} \left\{ -|x_1(\theta)| + \int_{t_0}^{\theta} u^2(t)\mathrm{d}t \right\}.$$

Hamiltonian

$$H(x, p) = p_1 x_2 - \frac{p_2^2}{4}.$$

Characteristic system

$$\begin{cases} \dfrac{\mathrm{d}\widehat{x}_1}{\mathrm{d}t} = \widehat{x}_2, \ \dfrac{\mathrm{d}\widehat{x}_2}{\mathrm{d}t} = -\frac{\widehat{p}_2}{2}, \\ \dfrac{\mathrm{d}\widehat{p}_1}{\mathrm{d}t} = 0, \quad \dfrac{\mathrm{d}\widehat{p}_2}{\mathrm{d}t} = -\widehat{p}_1, \ \dfrac{d\widehat{z}}{\mathrm{d}t} = -\frac{\widehat{p}_2{}^2}{4}. \end{cases}$$

Boundary conditions

$$\widehat{x}_1(0.5, y) = y_1, \ \widehat{x}_2(0.5, y) = y_2,$$
$$\widehat{p}_1(0.5, y) \in -\partial|y_1| = [-1, 1], \ \widehat{p}_2(0.5, y) = 0, \ \widehat{z}(0.5, y) = -|y_1|,$$
$$\widehat{z}(t, y) = -|\widehat{x}_1(t, y)| \Rightarrow \begin{cases} \widehat{p}_1(t, y) \in [-1, 1], \\ \widehat{p}_2(t, y) = 0. \end{cases}$$

The approximation parameters are: $\Delta t = 0.01$, $\rho_x = 0.0001$‘, $D = [-1.5, 1.5] \times [-1.5, 1.5]$, $\Delta x = 0.06$. The results of simulations are presented in Fig. 3.

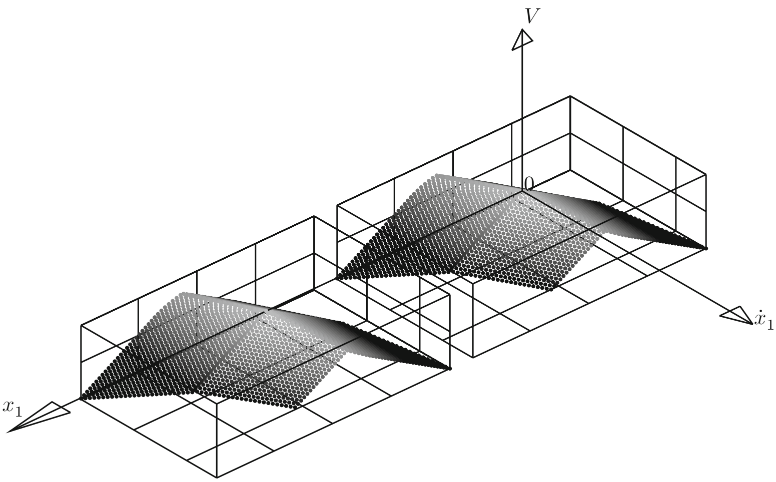


Fig. 3 Graphs of $\widetilde{V}(t, x)$ (*right graph*) and $\widetilde{C}_\Gamma\big(t, x; u_*^0(\cdot)\big)$ (*left graph*) at $t = t_0 = 0$

Acknowledgements Researches are supported by the Russian Foundation for Basic Research (grant 08-01-00410) and by the Program of the Government RF on the State Support for Leading Scientific Schools (grant NSh-2640.2008.1).

References

1. Bardi, M., Capuzzo-Dolcetta, I.: Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations. Birkhäuser, Boston (1997)
2. Bellman, R.: Dynamic Programming. University Press, Princeton (1957)
3. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, New York (1983)
4. Demyanov, V.F., Vasiliev, L.V.: Nondifferentiable Optimization. Springer-Optimization Software, New York (1985)
5. Fleming, W.H., Soner, H.M.: Controlled Markov Processes and Viscosity Solutions. Springer, New York (1993)
6. Krasovskii, N.N.: The Theory of Control of Motion. Nauka, Moscow (1968) (in Russian)
7. Krasovskii, N.N., Subbotin, A.I.: Game-Theoretical Control Problems. Springer, New York (1988)
8. Kurzhanski, A.B., Valyi, I.: Ellipsoidal Calculus for Estimation and Control. Birkhäuser, Boston (1996)
9. Melikyan, A.A.: Generalized Characteristics of First Order PDEs: Applications in Optimal Control and Differential Games. Birkhäuser, Boston (1998)
10. Osipov, Yu.S., Kryazhimskii, A.V.: Inverse Problems for Ordinary Differential Equations: Dynamical Solutions. Gordon and Breach, Boston (1995)
11. Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V., Mishchenko, E.F.: Mathematical Theory of Optimal Processes. Interscience, New York (1962)
12. Subbotin, A.I.: Generalized Solutions of First Order of PDEs:The Dynamical Optimization Perspectives. Birkhäuser, Boston (1995)
13. Subbotin, A.I., Chentsov, A.G.: Optimization of Guaranty in Control Problems. Nauka, Moscow (1981) (in Russian)
14. Subbotina, N.N.: The Method of Characteristics for Hamilton–Jacobi Equations and Applications to Dynamical Optimization. Journal of Mathematical Sciences **135**(3), 2955–3091, Springer, New York (2006)
15. Blagodatskikh, V.I.: The maximum principle for differential inclusions. Proceedings of the Steklov Insitute of Mathematics **166**, 23–43 (1984)
16. Crandall, M.G. and Lions, P.-L.: Viscosity Solutions of Hamilton–Jacobi Equations. Transactions of the American Mathematical Society. **277**(1), 1–42 (1983)
17. 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 UrB RAS, Ekaterinburg (2005) (in Russian)
18. Guseinov, Kh.G., Ushakov, V.N.: The construction of differential inclusions with prescribed properties. Differential equations. **36**(4), 488–496 (2000)
19. Kleimenov, A.F.: On the theory of hierarchical differential two-person games. Preprint (Institute of Mathematics and Mechanics USC, Academy of Sciences USSR, Sverdlovsk, 1985) (in Russian)
20. Rockafellar, R.T., Wets, R.J-B.: Variational Analysis, Springer (1997)
21. Souganidis, P.E.: Approximation schemes for viscosity solutions of Hamilton–Jacobi equations. Journal of Differential Equations **59**, 1–43 (1985)
22. Subbotina, N.N., Kolpakova, E.A.: On a Linear -Quadratic Optimal Control Problem with Geometric Restriction on Controls. International Journal of Tomography and Statistics **5**(W07), 62–67 (2007)

23. Subbotina, N.N, Subbotin, A.I., and Tret'jakov, V.E.: Stochastic and deterministic control. Differential inequalities. Lect. Notes in Control and Information Science, Stochastic Optimization **81**, 728–737, Springer, Berlin (1986)
24. Subbotina, N.N., Tokmantsev, T.B.: A numerical method for the minimax solution of the Bellman equation in the Cauchy problem with additional restrictions. Proceedings of the Steklov Institute of Mathematics, Supp.l, S222–S229 (2006)
25. Subbotina, N.N., Tokmantsev, T.B.: On the Efficiency of Optimal Grid Synthesis in Optimal Control Problems with Fixed Terminal Time. Journal of Differential Equations **45**(11), 1686–1698 (2009)
26. Tarasyev, A.M.: Approximation schemes for constructing minimax solutions of Hamilton–Jacobi equations. Journal of Applied Mathematics and Mechanics **58**(2), 207–221 (1985)

Sub- and Super-optimality Principles and Construction of Almost Optimal Strategies for Differential Games in Hilbert Spaces

Andrzej Święch

Abstract We prove sub- and superoptimality principles of dynamic programming and show how to use the theory of viscosity solutions to construct almost optimal strategies for two-player, zero-sum differential games driven by abstract evolution equations in Hilbert spaces.

1 Introduction

The main purpose of this paper is to show how Isaacs equations and the theory of viscosity solutions can be used to construct nearly optimal strategies for two-player, zero-sum differential games driven by abstract evolution equations in a Hilbert space. Let us briefly describe the game we have in mind. Let H be a real, separable Hilbert space and $T > 0$ be a fixed time horizon. For an initial time $t \in [0, T]$ and $x \in H$, the dynamics of the game is given by an evolution equation

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}s} X(s) = AX(s) + b(s, X(s), W(s), Z(s)) \\ X(t) = x, \end{cases} \tag{1}$$

where A is a linear, densely defined maximal dissipative operator in H, $b : [0, T] \times H \times \mathscr{W} \times \mathscr{Z} \to \mathbb{R}$ for some metric spaces $\mathscr{W}, \mathscr{Z}$, and

$$W \in M(t) := \{W : [t, T] \to \mathscr{W}, \text{ strongly measurable}\},$$

$$Y \in N(t) := \{Z : [t, T] \to \mathscr{Z}, \text{ strongly measurable}\}.$$

We will call $M(t)$ the set of controls for player I, and $N(t)$ the set of controls for player II.

A. Święch (✉)
School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA
e-mail: swiech@math.gatech.edu

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_8,

This setting is very general. The state equation can be, for instance, a single equation or a system of controlled semilinear PDE. We recall that an operator A is maximal dissipative if and only if it generates a semigroup of contractions e^{tA}, i.e., $\|\mathrm{e}^{tA}\| \le 1$.

The pay-off functional associated with the trajectory of the system given by (1) is

$$J(t, x; W(\cdot), Z(\cdot)) = \int_t^T L(s, X(s), W(s), Z(s))\mathrm{d}s + h(X(T)).$$

Player I controls W and wants to maximize J over all choices of Z. Player II controls Z and wants to minimize J over all choices of W. The game is played in continuous time. Following Elliot and Kalton [8], we define admissible strategies for both players and consider the upper and lower versions of the game. In the lower game, player II chooses $Z(s)$ knowing $W(s)$, and in the upper game player I chooses $W(s)$ knowing $Z(s)$.

Admissible strategies of player I:

$$\Gamma(t) := \{\alpha : N(t) \to M(t), \text{ nonanticipating}\}$$

Admissible strategies of player II:

$$\Delta(t) := \{\beta : M(t) \to N(t), \text{ nonanticipating}\}.$$

Strategy α (resp., β) is nonanticipating if whenever $Z_1(r) = Z_2(r)$ (resp., $W_1(r) = W_2(r)$) a.e. on $[t, s]$ then $\alpha[Z_1](r) = \alpha[Z_2](r)$ (resp., $\beta[W_1](r) = \beta[W_2](r)$) a.e. on $[t, s]$ for every $s \in [t, T]$.

The values of the lower and upper games , called the lower and upper values are defined as follows:

Lower value of the game:

$$V(t, x) = \inf_{\beta \in \Delta(t)} \sup_{W \in M(t)} J(t, x; W(\cdot), \beta[W](\cdot))$$

Upper value of the game:

$$U(t, x) = \sup_{\alpha \in \Gamma(t)} \inf_{Z \in N(t)} J(t, x; \alpha[Z](\cdot), Z(\cdot))$$

The lower and upper value Isaacs equations associated with the two value functions are

$$\begin{cases} u_t + \langle Ax, Du\rangle + F^{\mp}(t, x, Du) = 0 \\ u(T, x) = h(x), \end{cases} \tag{2}$$

where the lower value Hamiltonian F^- is defined by

$$F^-(t, x, p) = \sup_{w \in \mathscr{W}} \inf_{z \in \mathscr{Z}} \{\langle b(t, x, w, z), p\rangle + L(t, x, w, z)\}$$

and the upper value Hamiltonian F^+ is defined by

$$F^+(t,x,p) = \inf_{z\in\mathscr{Z}} \sup_{w\in\mathscr{W}} \{\langle b(t,x,w,z), p\rangle + L(t,x,w,z)\}.$$

Above $\langle\cdot,\cdot\rangle$ is the inner product in H, and Du is the Fréchet derivative of a function u.

The lower value function V should be the unique viscosity solution of the lower Isaacs equation (2)$^-$ (with F^-) and the upper value function U should be the unique viscosity solution of the upper Isaacs equation (2)$^+$ (with F^+). The definition of viscosity solution will be given in the next section. Equality $F^- = F^+$ is known as the Isaacs condition. If this happens, by uniqueness of viscosity solutions of (2), we obtain that $V = U$ and we then say that the game has value.

Notice that, since sup inf $\leq$ inf sup, we always have $F^- \leq F^+$. In particular if V is a viscosity solution of (2)$^-$ then it is a viscosity subsolution of (2)$^+$ so if comparison holds for (2)$^+$ (a subsolution is less than or equal to a supersolution) we get $V \leq U$.

There are several papers in which viscosity solutions and Isaacs equations have been used to construct a saddle (and an approximate saddle) point strategies for infinite dimensional differential games in Hilbert spaces [12, 17, 19]. In these papers, Berkovitz's notion of strategy and pay-off [1] is employed. We show how to construct nearly optimal strategies in the Elliott–Kalton sense. Our approach is based on the proofs of sub- and superoptimality principles of dynamic programming which are interesting on their own and in fact are the main results of this paper. The proofs use the method of regularization by sup- and inf-convolutions and integration along trajectories, together with some techniques of [9] and [11]. This method was used before to show sub- and superoptimality principles for finite dimensional control problems [21] and finite dimensional stochastic differential games [22], and was recently generalized to infinite dimensional control problems [10]. Similar method was employed in [3] to construct stabilizing feedbacks for nonlinear systems and in [20] to study feedback stabilization of finite dimensional nonlinear H_∞ control systems.

2 Notation, Definitions, and Background

Throughout this chapter, H is a real separable Hilbert space equipped with the inner product $\langle\cdot,\cdot\rangle$ and the norm $\|\cdot\|$.

Let B be a bounded, linear, positive, self-adjoint operator on H such that A^*B is bounded on H and

$$\langle (A^*B + c_0 B)x, x\rangle \leq 0 \qquad \text{for all } x \in H$$

for some $c_0 \leq 0$. Such an operator always exists, for instance $B = ((-A+I)(-A^*+I))^{-1/2}$ (see [18]). We refer to [5] for various examples of B. Using the operator B we define for $\gamma > 0$ the space $H_{-\gamma}$ to be the completion of H under the norm

$$\|x\|_{-\gamma} = \|B^{\frac{\gamma}{2}} x\|.$$

Let $\Omega \subset [0,T]\times H$. We say that $u{:}\Omega \to \mathbb{R}$ is B-upper-semicontinuous (respectively, B-lower-semicontinuous) on Ω if whenever $t_n \to t$, $x_n \rightharpoonup x$, $Bx_n \to Bx$, $(t,x) \in \Omega$, then $\limsup_{n\to+\infty} u(t_n,x_n) \le u(t,x)$ (respectively, $\liminf_{n\to+\infty} u(t_n,x_n) \ge u(t,x)$). The function u is B-continuous on Ω if it is B-upper-semicontinuous and B-lower-semicontinuous on Ω.

We will say that a function v is B-semiconvex (respectively, B-semiconcave) if there exists a constant $C \ge 0$ such that $v(t,x) + C(\|x\|_{-1}^2 + t^2)$ is convex (respectively, $v(t,x) - C(\|x\|_{-1}^2 + t^2)$ is concave).

We will denote by B_R the open ball of radius R centered at 0 in H.

The following conditions will be assumed throughout the paper.

Hypothesis 1

- *A is a linear, densely defined, maximal dissipative operator in H.*
- *$\mathscr{W}$ and $\mathscr{Z}$ are complete, separable metric spaces.*
- *$b : [0,T] \times H \times \mathscr{W} \times \mathscr{Z} \to H$ is continuous,*

$$\sup_{(t,w,z)\in[0,T]\times\mathscr{W}\times\mathscr{Z}} \|b(t,0,w,z)\| \le +\infty,$$

and there exist a constant $M > 0$ and a local modulus of continuity ω such that

$$\|b(t,x,w,z) - b(s,y,w,z)\| \le M\|x_1 - x_2\|_{-1} + \omega(|t-s|;\|x\| \wedge \|y\|)$$

for all $t,s \in [0,T], x,y \in H, w \in \mathscr{W}, z \in \mathscr{Z}$.
- *$L : [0,T] \times H \times \mathscr{W} \times \mathscr{Z} \to H$ is continuous,*

$$|L(t,x,w,z)| \le C(1 + \|x\|^k) \quad \textit{for some } k \ge 0,$$
$$|L(t,x,w,z) - L(s,y,w,z)| \le +\omega(|t-s| + \|x-y\|_{-1}; \|x\| \vee \|y\|)$$

for all $t,s \in [0,T], x,y \in H, w \in \mathscr{W}, z \in \mathscr{Z}$ and some local modulus ω.

The solution of the Isaacs equation is understood in a modified viscosity sense of Crandall and Lions [5,6]. We consider two sets of tests functions:

$$\begin{aligned} test1 = \{\varphi \in C^1((0,T)\times H) \ &: \varphi \text{ is } B\text{-lower semicontinuous and} \\ &\quad A^* D\varphi \in C((0,T)\times H)\} \end{aligned}$$

and

$$\begin{aligned} test2 = \{g \in C^1((0,T)\times H) \ &: \exists g_0, : [0,+\infty) \to [0,+\infty), \\ &\quad \text{and } \eta \in C^1((0,T)) \text{ positive } \ s.t. \\ &\quad g_0 \in C^1([0,+\infty)),\ g_0'(r) \ge 0\ \forall r \ge 0, \\ &\quad g_0'(0) = 0 \text{ and } g(t,x) = \eta(t) g_0(\|x\|)\} \end{aligned}$$

Sometimes it is necessary to take the finite linear combinations of functions $\eta(t)g_0(\|x\|)$ above as test 2 functions; however, this will not be needed here.

Below we present the definition of viscosity solution. We point out that this definition applies to terminal value problems.

Definition 1. A B-upper semicontinuous function u on $(0, T) \times H$ is a viscosity subsolution of

$$u_t + \langle Ax, Du\rangle + F(t, x, Du) = 0 \quad \text{in } (0, T) \times H \tag{3}$$

if whenever $v - \varphi - g$ has a local maximum at $(\bar{t}, \bar{x}) \in (0, T) \times H$ for $\varphi \in$ *test1* and $g \in$ *test2* then

$$\varphi_t(\bar{t}, \bar{x}) + g_t(\bar{t}, \bar{x}) + \langle A^* D\varphi(\bar{t}, \bar{x}), \bar{x}\rangle + F(\bar{t}, \bar{x}, D\varphi(\bar{t}, \bar{x}) + Dg(\bar{t}, \bar{x})) \geq 0.$$

A B-lower semicontinuous function u on $(0, T) \times H$ is a viscosity supersolution of (3) if whenever $v + \varphi + g$ has a local minimum at $(\bar{t}, \bar{x}) \in (0, T) \times H$ for $\varphi \in test1$ and $g \in test2$ then

$$-\varphi_t(\bar{t}, \bar{x}) - g_t(\bar{t}, \bar{x}) - \langle A^* D\varphi(\bar{t}, \bar{x}), \bar{x}\rangle + F(\bar{t}, \bar{x}, -D\varphi(\bar{t}, \bar{x}) - Dg(\bar{t}, \bar{x})) \leq 0.$$

A function v is a viscosity solution of (3) if it is both a viscosity subsolution and a viscosity supersolution of (3).

For a function v we will denote by $D^+v(t, x)$ the superdifferential of v at (t, x), i.e., the set of all pairs $(a, p) \in \mathbb{R} \times H$ such that

$$v(s, y) - v(t, x) - \langle p, y - x\rangle - a(s - t) \leq o(\|x - y\| + |t - s|).$$

The subdifferential $D^-v(t, x)$is defined the same by reversing the inequality above. It is well known (see for instance [2]) that for a semiconvex (resp., semiconcave) function v, its subdifferential (resp., superdifferential) at a point (s, z) is equal to

$$\overline{\mathrm{conv}}\{(a, p) : v_t(s_n, z_n) \to a, Dv(s_n, z_n) \rightharpoonup p, s_n \to s, z_n \to z\}$$

i.e., it is equal to the closure of the convex hull of the set of weak limits of derivatives of v nearby.

Solution of (1) is understood in the mild sense (see for instance [16]), which means that

$$X(s) = \mathrm{e}^{(s-t)A}x + \int_t^s \mathrm{e}^{(s-\tau)A} b(\tau, X(\tau), W(\tau), Z(\tau))\mathrm{d}\tau.$$

It is well known [16] that under the assumptions of Hypothesis 1 for every $t \in [0, T], x \in H$ there exists a unique mild solution $X \in C(t, T; H)$ of (1). Moreover, there exists a constant $C = C(T, \|x\|)$, independent of $W \in M(t)$ and $Z \in N(t)$, such that

$$\max_{t \le s \le T} \|X(s)\| \le C. \tag{4}$$

Also we have that if φ is a test1 function, $t < \tau < T$, and X is the solution of (1), then

$$\begin{aligned}\varphi(\tau, X(\tau)) - \varphi(t, x) = \int_t^\tau & (\varphi_t(r, X(r)) + \langle X(r), A^* D\varphi(r, X(r))\rangle \\ & + \langle b(r, X(r), W(r), Z(r)), D\varphi(r, X(r))\rangle)\mathrm{d}r \end{aligned} \tag{5}$$

(see [16], Proposition 5.5, page 67).

3 Regularization by B-sup- and B-inf-convolutions

We first recall the results about sup- and inf-convolutions from [10]. Following the functions introduced in [6] and their modifications from [10], for a function $u : (0, T) \times H \to \mathbb{R}$ and $\varepsilon, \beta, \lambda > 0, m \ge 2, K \ge 0$, we define the B-sup-convolution of u by

$$u^{\lambda,\varepsilon,\beta}(t, x) = \sup_{(s,y)} \left\{ u(s, y) - \frac{\|x - y\|_{-1}^2}{2\varepsilon} - \frac{(t - s)^2}{2\beta} - \lambda \mathrm{e}^{2mK(T-s)} \|y\|^m \right\}$$

and the B-inf-convolution of u by

$$u_{\lambda,\varepsilon,\beta}(t, x) = \inf_{(s,y)} \left\{ u(s, y) + \frac{\|x - y\|_{-1}^2}{2\varepsilon} + \frac{(t - s)^2}{2\beta} + \lambda \mathrm{e}^{2mK(T-s)} \|y\|^m \right\}.$$

The following properties of B-sup- and inf-convolutions have been proved in [10], Lemma 4.2.

Lemma 1. *Let w be such that*

$$w(t, x) \le C(1 + \|x\|^k) \quad (\textit{respectively, } \; w(t, x) \ge -C(1 + \|x\|^k)) \tag{6}$$

on $(0, T) \times H$ for some $k \ge 0$. Let $m > \max(k, 2)$. Then:

(i) For every $R > 0$ there exists $M_{R,\varepsilon,\beta}$ such that if $v = w^{\lambda,\varepsilon,\beta}$ (respectively, $v = w_{\lambda,\varepsilon,\beta}$) then

$$|v(t, x) - v(s, y)| \le M_{R,\varepsilon,\beta}(|t - s| + \|x - y\|_{-2}) \quad \textit{on} \;\; (0, T) \times B_R$$

(ii) The function

$$w^{\lambda,\varepsilon,\beta}(t, x) + \frac{\|x\|_{-1}^2}{2\varepsilon} + \frac{t^2}{2\beta}$$

is convex (respectively,

$$w_{\lambda,\varepsilon,\beta}(t,x) - \frac{\|x\|_{-1}^2}{2\varepsilon} - \frac{t^2}{2\beta}$$

is concave). In particular $w^{\lambda,\varepsilon,\beta}$ *(respectively,* $w_{\lambda,\varepsilon,\beta}$*) is* B*-semiconvex (respectively,* B*-semiconcave).*

(iii) If $v = w^{\lambda,\varepsilon,\beta}$ *(respectively,* $v = w_{\lambda,\varepsilon,\beta}$*) and* v *is differentiable at* $(t,x) \in (0,T) \times B_R$ *then* $|v_t(t,x)| \le M_{R,\varepsilon,\beta}$*, and* $Dv(t,x) = Bq$*, where* $\|q\| \le M_{R,\varepsilon,\beta}$

The next lemma is just a restatement of Lemma 4.3 of [10] for Isaacs equations. Its proof is the same as the proof of Lemma 4.3 of [10]. In fact, this result can be shown for general equations not necessarily related to control or game problems.

Lemma 2. *Let Hypothesis 1 be satisfied. Let* w *satisfy (6) and be a* B*-upper-semicontinuous viscosity subsolution (respectively, a* B*-lower-semicontinuous viscosity supersolution) of*

$$w_t + \langle Ax, Dw\rangle + F(t,x,Dw) = 0,$$

where $F = F^-$ *or* $F = F^+$*. Let* $m > \max(k,2)$*. If* $K > 0$ *is big enough, then for every* $R, \delta > 0$ *there exists a nonnegative function* $\gamma_{R,\delta}(\lambda,\varepsilon,\beta)$*, where*

$$\lim_{\lambda\to 0}\limsup_{\varepsilon\to 0}\limsup_{\beta\to 0}\gamma_{R,\delta}(\lambda,\varepsilon,\beta) = 0, \tag{7}$$

such that $w^{\lambda,\varepsilon,\beta}$ *(respectively,* $w_{\lambda,\varepsilon,\beta}$*) is a viscosity subsolution (respectively, supersolution) of*

$$v_t + \langle Ax, Dv\rangle + F(t,x,Dv) = -\gamma_{R,\delta}(\lambda,\varepsilon,\beta) \quad \text{in } (\delta, T-\delta)\times B_R$$

(respectively,

$$v_t + \langle Ax, Dv\rangle + F(t,x,Dv) = \gamma_{R,\delta}(\lambda,\varepsilon,\beta) \quad \text{in } (\delta, T-\delta)\times B_R)$$

for β *sufficiently small (depending on* δ*), in the sense that if* $v - \psi$ *has a local maximum (respectively,* $v + \psi$ *has a local minimum) at* (t,x) *for a test function* $\psi = \varphi + g$ *then*

$$\psi_t(t,x) + \langle x, A^* D\psi(t,x)\rangle + F(t,x,D\psi(t,x)) \ge -\gamma_{R,\delta}(\lambda,\varepsilon,\beta)$$

(respectively,

$$-\psi_t(t,x) - \langle x, A^* D\psi(t,x)\rangle + F(t,x,-D\psi(t,x)) \le \gamma_{R,\delta}(\lambda,\varepsilon,\beta)).$$

The above lemma leaves us with a small problem. It is not even clear if $u^{\lambda,\varepsilon,\beta}$ and $u_{\lambda,\varepsilon,\beta}$ satisfy pointwise inequalities at all points of their differentiability. This is indeed true. In fact under some assumptions on the Hamiltonian F, $w^{\lambda,\varepsilon,\beta}$ satisfies the Isaacs inequality at every point for some elements of its subdifferential and $w_{\lambda,\varepsilon,\beta}$ satisfies the Isaacs inequality at every point for some elements of its superdifferential. The right elements of the sub-/superdifferentials are the weak limits of derivatives. In Lemma 3, we use the notation from Lemma 2. The proof of Lemma 3 is the same as the proof of Lemma 4.5 of [10].

Lemma 3. *Let the assumptions of Lemma 2 be satisfied and let $F(t,x,\cdot)$ be weakly sequentially continuous for every $(t,x) \in (0,T) \times H$. Let $(a,p) \in D^- w^{\lambda,\varepsilon,\beta}(t,x)$ (respectively, $(a,p) \in D^+ w_{\lambda,\varepsilon,\beta}(t,x)$) be such that $(w_t^{\lambda,\varepsilon,\beta}(t_n,x_n), Dw^{\lambda,\varepsilon,\beta}(t_n,x_n)) \rightharpoonup (a,p)$ for some $(t_n,x_n) \to (t,x)$ (respectively, $((w_{\lambda,\varepsilon,\beta})_t(t_n,x_n), Dw_{\lambda,\varepsilon,\beta}(t_n,x_n)) \rightharpoonup (a,p)$). Then*

$$a + \langle x, A^* p\rangle + F(t,x,p) \geq -\gamma(\lambda,\varepsilon,\beta)$$

(respectively,

$$a + \langle x, A^* p\rangle + F(t,x,p) \leq \gamma(\lambda,\varepsilon,\beta)).$$

The assumption that $F^{\mp}(t,x,\cdot)$ be weakly sequentially continuous is a bit restrictive but it is satisfied, for instance, when $\mathscr{W}$ and $\mathscr{Z}$ are compact metric spaces.

Lemma 4. *If $\mathscr{W}$, $\mathscr{Z}$ are compact metric spaces then $F^-(t,x,\cdot)$ and $F^+(t,x,\cdot)$ are weakly sequentially continuous for every $(t,x) \in (0,T) \times H$.*

Proof. We will only show the weak sequential continuity of F^-. Let $p_n \rightharpoonup p$. Recall that $F^- = \sup_w \inf_z$. Therefore for every w there is z_n such that

$$F^-(t,x,p_n) \geq \langle p_n, b(t,x,w,z_n)\rangle + L(t,x,w,z_n) - \frac{1}{n}.$$

By compactness of $\mathscr{Z}$ we can assume that $z_n \to \bar{z}$ for some $\bar{z}$ so passing to the liminf above we get

$$\liminf_{n\to\infty} F^-(t,x,p_n) \geq \langle p, b(t,x,w,\bar{z})\rangle + L(t,x,w,\bar{z}).$$

It is now enough to take the $\inf_z$ and then $\sup_w$ in the above inequality to obtain the weak sequential lower-semicontinuity of F^-. The weak sequential upper-semicontinuity of F^- is proved similarly. $\square$

4 Sub- and Superoptimality Principles and Construction of Almost Optimal Strategies

The proofs of the following sub- and superoptimality inequalities of dynamic programming are the main result of this paper.

Theorem 1. *Let Hypothesis 1 be satisfied and let $F^-(t,x,\cdot)$ and $F^+(t,x,\cdot)$ be weakly sequentially continuous for every $(t,x) \in (0,T) \times H$. Suppose that for every (t,x) there exists a modulus $\omega_{t,x}$ such that if X is the solution of (1) then*

$$\|X(s_2) - X(s_1)\| \le \omega_{t,x}(s_2 - s_1) \tag{8}$$

for all $t \le s_1 \le s_2 \le T$ and all $W \in M(t)$, $Z \in N(t)$. Let u be a function such that

$$|u(t,x)| \le C(1 + \|x\|^k) \quad \textit{for all } x \in H \tag{9}$$

for some $k \ge 0$ and such that for every $R > 0$ there exists a modulus σ_R such that

$$|u(t,x) - u(s,y)| \le \sigma_R(|t-s| + \|x-y\|_{-1}) \tag{10}$$

for all $t, s \in [0,T]$, $\|x\|, \|y\| \le R$. Then:

(i) If u is a viscosity supersolution of $(2)^-$, then for every $0 < t < s < T$, $x \in H$

$$u(t,x) \ge \inf_{\beta \in \Delta(t)} \sup_{W \in M(t)} \left\{ \int_t^s L(\tau, X(\tau), W(\tau), \beta[W](\tau))\mathrm{d}\tau + u(s, X(s)) \right\}. \tag{11}$$

(ii) If u is a viscosity subsolution of $(2)^-$, then for every $0 < t < s < T$, $x \in H$

$$u(t,x) \le \inf_{\beta \in \Delta(t)} \sup_{W \in M(t)} \left\{ \int_t^s L(\tau, X(\tau), W(\tau), \beta[W](\tau))\mathrm{d}\tau + u(s, X(s)) \right\}.$$

(iii) If u is a viscosity supersolution of $(2)^+$, then for every $0 < t < s < T$, $x \in H$

$$u(t,x) \ge \sup_{\alpha \in \Gamma(t)} \inf_{Z \in N(t)} \left\{ \int_t^s L(\tau, X(\tau), \alpha[Z](\tau), Z(\tau))\mathrm{d}\tau + u(s, X(s)) \right\}.$$

(iv) If u is a viscosity subsolution of $(2)^+$, then for every $0 < t < s < T$, $x \in H$

$$u(t,x) \le \sup_{\alpha \in \Gamma(t)} \inf_{Z \in N(t)} \left\{ \int_t^s L(\tau, X(\tau), \alpha[Z](\tau), Z(\tau))\mathrm{d}\tau + u(s, X(s)) \right\}.$$

Remark 1. Condition (8) says that trajectories starting at a fixed point x at a fixed time t are uniformly continuous on $[t,T]$, uniformly in $W \in M(t)$, $Z \in N(t)$. It is a little restrictive, however it seems necessary to obtain uniform estimates on the error terms in the proof. In general, one may expect it to hold when the semigroup e^{tA} has some regularizing properties. It was shown in [10] that it holds for example if $A = A^*$, it generates a differentiable semigroup, and

$$\|A\mathrm{e}^{tA}\| \le \frac{C}{t^\delta}$$

for some $\delta < 2$. One can also check that (8) is satisfied if the semigroup e^{tA} is compact.

Proof. We will only show (i) and (ii) as the proofs of (iii) and (iv) are analogous.

(i) *Step 1.* (Reduction to the B-semiconcave case): Let (t, x) be fixed and let δ be such that $0 < \delta < t < t + h < T - \delta$. We choose m and K as in Lemma 2. We notice that it follows from (9) and (10) that

$$|u_{\lambda,\varepsilon,\beta}(\tau, y) - u(\tau, y)| \le \tilde{\sigma}_R(\lambda + \varepsilon + \beta) \quad \text{for } \tau \in (0, T), \|y\| \le R, \quad (12)$$

where the modulus $\tilde{\sigma}_R$ can be explicitly calculated from σ_R. Since by (4) all trajectories of (1) stay in some ball B_R on time interval $[t, T]$, setting $L := L - \gamma_{R,\delta}(\lambda, \varepsilon, \beta)$, using (7), (12) and Lemma 3, it is enough to show (ii) when u is B-semiconcave and such that for every $(\tau, y) \in (\delta, T-\delta) \times B_R$ there exists $(a, p) \in D^+u(\tau, y)$ such that

$$a + \langle y, A^* p\rangle + F^-(\tau, y, p) \le 0. \quad (13)$$

By Lemma 1 we may assume that $u(\tau, y) - \frac{\|y\|_{-1}^2}{2\varepsilon} - \frac{\tau^2}{2\beta}$ is concave.
Step 2. (Construction of nearly optimal control): Let $W \in M(t)$ be a fixed control. For $n \ge 1$ set $h = (s - t)/n$. Take $(a, p) \in D^+u(t, x)$ satisfying (13). (Recall $F^- = \sup_w \inf_z$.) It then follows from the separability of $\mathscr{W}$ and the continuity of b and L that there exist points $z_1, z_2, \ldots \in \mathscr{Z}, w_1, w_2, \ldots \in \mathscr{W}$ and balls $B_{r_i}(w_i), i = 1, 2, \ldots$ such that $\cup_{i=1}^{\infty} B_{r_i}(w_i) = \mathscr{W}$ and

$$a + \langle x, A^* p\rangle + \langle b(t, x, w, z_i), p\rangle + L(t, x, w, z_i) \le \frac{1}{n}$$

for $w \in B_{r_i}(w_i)$. Define a map $\Psi : \mathscr{W} \to \mathscr{Z}$ by

$$\Psi(w) = z_i \quad \text{if } w \in B_{r_i}(w_i) \Big\backslash \bigcup_{j=1}^{i-1} B_{r_j}(w_j), i = 1, 2, \ldots.$$

Define a control $Z^n \in N(t)$ by

$$Z^n(\tau) = \Psi(W(\tau)) \quad \text{for } \tau \in [t, t + h).$$

Then

$$a + \langle x, A^* p\rangle + \langle b(t, x, W(\tau), Z^n(\tau)), p\rangle + L(t, x, W(\tau), Z^n(\tau)) \le \frac{1}{n} \quad (14)$$

for $\tau \in [t, t+h)$. Denote by $X(\tau)$ the trajectory corresponding to controls $W(\tau)$ and $Z^n(\tau)$.

Step 3. (Integration along trajectories): It follows from the B-semiconcavity of u that

$$u(\tau, y) \leq u(t, x) + a(\tau - t) + \langle p, y - x\rangle + \frac{\|y - x\|_{-1}^2}{2\varepsilon} + \frac{(\tau - t)^2}{2\beta}.$$

Therefore, it follows from (5), (8), and (14) that

$$\begin{aligned} u(t+h, X(t+h)) &\leq u(t,x) + \int_t^{t+h} \left[a + \left\langle X(\tau), A^*\left(p + \frac{B(X(\tau)-x)}{\varepsilon}\right)\right\rangle \right.\\ &\quad \left. + \left\langle b(\tau, X(\tau), W(\tau), Z^n(\tau)), p + \frac{B(X(\tau)-x)}{\varepsilon}\right\rangle\right] \mathrm{d}\tau + \frac{h^2}{2\beta} \\ &\leq u(t,x) + \int_t^{t+h} [a + \langle X(\tau), A^* p\rangle \\ &\quad + \langle b(\tau, X(\tau), W(\tau), Z^n(\tau)), p\rangle] \mathrm{d}\tau + o\left(\frac{1}{n}\right) \\ &\leq u(t,x) + \int_t^{t+h} [a + \langle x, A^* p\rangle + \langle b(t, x, W(\tau), Z^n(\tau)), p\rangle] \mathrm{d}\tau + o\left(\frac{1}{n}\right) \\ &\leq u(t,x) - \int_t^{t+h} L(t, x, W(\tau), Z^n(\tau)) \mathrm{d}\tau + o\left(\frac{1}{n}\right) \\ &\leq u(t,x) - \int_t^{t+h} L(\tau, X(\tau), W(\tau), Z^n(\tau)) \mathrm{d}\tau + o\left(\frac{1}{n}\right), \end{aligned}$$

where $o\left(\frac{1}{n}\right)$ is independent of W and Ψ. We now repeat the above process starting at $(t+h, X(t+h))$ to obtain a control Z_n on $[t+h, t+2h)$ such that

$$\begin{aligned} u(t+2h, X(t+2h)) &\leq u(t+h, X(t+h)) - \int_{t+h}^{t+2h} L(\tau, X(\tau), \\ &\quad W(\tau), Z^n(\tau)) \mathrm{d}\tau + o\left(\frac{1}{n}\right). \end{aligned}$$

After n iterations we produce a control Z_n on $[t, s)$ which we can extend to the interval $[t, T]$ so that $Z_n \in N(t)$ for which

$$u(s, X(s)) \leq u(t,x) - \int_t^s L(\tau, X(\tau), W(\tau), Z^n(\tau)) \mathrm{d}\tau + n o\left(\frac{1}{n}\right),$$

where $o(\frac{1}{n})$ is independent of W and Z^n.
Step 4. (Construction of the strategy): We now define a strategy

$$\beta^n[W](\tau) := Z^n(\tau).$$

By construction of Z_n, it is clear that β^n is nonanticipating, i.e., $\beta^n \in \Delta(t)$. Moreover, we have

$$u(t,x) \geq \int_t^s L(\tau, X(\tau), W(\tau), \beta^n[W](\tau))\mathrm{d}\tau + u(s, X(s)) + no\left(\frac{1}{n}\right),$$

where $o(\frac{1}{n})$ is independent of W and β^n. This gives the superoptimality principle (11) after we take the sup over $W \in M(t)$, then inf over $\beta \in \Delta(t)$, and then let $n \to \infty$.

(*ii*) Let (t,x) be fixed and let δ be such that $0 < \delta < t < t+h < T-\delta$. By the same argument as in Step 1 of (i) it is enough to show (ii) when u is B-semiconcave and is such that for every $(\tau, y) \in (\delta, T-\delta) \times B_R$ (for R big enough) there exists $(a,p) \in D^-u(\tau, y)$ such that

$$a + \langle y, A^* p\rangle + F^-(\tau, y, p) \geq 0. \tag{15}$$

Let Z be any control in $N(t)$. For $n \geq 1$ we set $h = (s-t)/n$. We take $(a,p) \in D^-u(t,x)$ satisfying (15) and we choose $w_1 \in \mathscr{W}$ such that

$$a + \langle x, A^* p\rangle + \inf_{z\in\mathscr{Z}}\{\langle b(t,x,w_1,z), p\rangle + L(t,x,w_1,z)\} \geq -\frac{1}{n},$$

and then we define a control $W^n(\tau) = w_1$ for $\tau \in [t, t+h)$ Arguing similarly as in Step 2 of (i) we then obtain

$$u(t+h, X(t+h)) \geq u(t,x) - \int_t^{t+h} L(\tau, X(\tau), W^n(\tau), Z(\tau))\mathrm{d}\tau + o\left(\frac{1}{n}\right),$$

where X is the solution of (1) with the controls W^n and Z, and the term $o(\frac{1}{n})$ is independent of Z and w_1. After n iterations of this process, we produce points $w_1, ..., w_n \in \mathscr{W}$ and a piecewise constant control $W^n \in M(t)$ such that $w^n(\tau) = w_i$ for $\tau \in [t+(t-1)h, t+ih)$, $i = 1, ..., n$, and such that

$$u(s, X(s)) \geq u(t,x) - \int_t^s L(\tau, X(\tau), W^n(\tau), Z(\tau))\mathrm{d}\tau + no\left(\frac{1}{n}\right),$$

where $o(\frac{1}{n})$ is independent of Z and W^n by (8). Therefore, we can define a nonanticipative strategy $\alpha^n \in \Gamma(t)$ by setting $\alpha^n[Z](\tau) = W^n(\tau)$. This strategy satisfies

$$u(s, X(s)) \geq u(t,x) - \int_t^s L(\tau, X(\tau), \alpha^n[Z](\tau), Z(\tau))\mathrm{d}\tau + no\left(\frac{1}{n}\right), \tag{16}$$

where $o(\frac{1}{n})$ is independent of Z and the definition of α^n. Moreover, $\alpha^n[Z]\big|_{[t+(i-1)h, t+ih)}$ depends only on $Z\big|_{[t, t+(i-1)h)}$ for $i = 1, ..., n$.

It is now standard to notice (see [22]) that for every $\beta \in \Delta(t)$ there exist $\tilde{W} \in M(t)$ and $\tilde{Z} \in N(t)$ such that $\alpha^n[\tilde{Z}] = \tilde{W}$ and $\beta[\tilde{W}] = \tilde{Z}$ on $[t, s]$. We recall how to do this for reader's convenience. We first set $\tilde{W}_{|[t,t+h)} = w_1$ as w_i only depended on (t, x), and then define $\tilde{Z}_{|[t,t+h)} = \beta[\tilde{W}]_{|[t,t+h)}$. (This means that $\tilde{W}$ has to be extended to $[t, T]$ but since $\beta[\tilde{W}]_{|[t,t+h)}$ only depends on $\tilde{W}_{|[t,t+h)}$ we will not worry about this.) We proceed inductively. Suppose we know $\tilde{W}$ and $\tilde{Z}$ on $[t, t+ih)$. By the construction of α^n we now set $\tilde{W}_{|[t,t+(i+1)h)} = \alpha^n[\tilde{Z}]_{|[t,t+(i+1)h)}$ (since $\tilde{W}_{|[t,t+(i+1)h)}$ only depends on $\tilde{Z}_{|[t,t+ih)}$) and then define $\tilde{Z}_{|[t,t+(i+1)h)} = \beta[\tilde{W}]_{|[t,t+(i+1)h)}$. It is clear from this construction that after n iterations we obtain $\alpha^n[\tilde{Z}] = \tilde{W}$ and $\beta[\tilde{W}] = \tilde{Z}$ on $[t, s]$.

Therefore, (16) applied to $\tilde{Z}$ gives us that for every $\beta \in \Delta(t)$ there exists $\tilde{W} \in M(t)$ such that

$$u(t, x) \leq \int_t^s L(\tau, X(\tau), \tilde{W}(\tau), \beta[\tilde{W}](\tau))\mathrm{d}\tau + u(s, X(s)) + no\left(\frac{1}{n}\right),$$

where $o(\frac{1}{n})$ is independent of β and $\tilde{W}$. (We remind that X above is the solution of (1) with $W := \tilde{W}$ and $Z := \beta[\tilde{W}]$.) It is now enough to take the sup over $W \in M(t)$, then inf over $\beta \in \Delta(t)$, and then let $n \to \infty$. $\square$

An immediate corollary of Theorem 1 is that viscosity solutions of $(2)^-$ and $(2)^+$ must be the lower and upper value functions respectively.

Corollary 1. *Let the assumptions of Theorem 1 be satisfied and let $u(T, x) = g(x)$ for $x \in H$. If u is a viscosity solution of $(2)^-$ (resp., $(2)^+$) then $u = V$ (resp., $u = U$). In particular, V and U satisfy the Dynamic Programming Principle.*

Unfortunately, the general proof of existence of viscosity solutions of $(2)^-$ and $(2)^+$ is based on the Dynamic Programming Principle and the direct verification that the value function is a viscosity solution. Proofs by Perron's method with different definitions of solutions are in [13, 24] and in [14] for B-continuous viscosity solutions when A is more coercive. When B is compact existence of solutions of general equations like (1) was shown in [5] by finite dimensional approximations. Proofs using the value functions of the games are in [15,23] and in [4] when $A = 0$. Unfortunately, none of these papers provides an exact reference to the fact that lower and upper value functions considered in this paper are viscosity solutions of $(2)^-$ and $(2)^+$ in the sense of Definition 1. However, the proof of this basically follows the arguments of the finite dimensional proof of [9] with necessary modifications using continuous dependence estimates for (1) and other techniques that can be found in [5, 6, 16]. The reader can also consult [15] for a complete proof in the infinite horizon case, even though it uses a different definition of viscosity solution. Here, we just state the result.

Theorem 2. *Let Hypothesis 1 hold and let h be a function such that*

$$|h(x)| \le C(1 + \|x\|^k) \quad \textit{for all } x \in H$$

for some $k \ge 0$, and such that for every $R > 0$ there exists a modulus ω_R such that

$$|h(x) - h(y)| \le \omega_R(\|x - y\|_{-1}) \quad \textit{if } \|x\|, \|y\| \le R.$$

Then U and V are the unique viscosity solutions of (1) satisfying (9) and (10).

The proofs of the superoptimality principles (i) and (iii) provide an explicit method of construction of almost optimal strategies using the Isaacs equations. Let us, for instance, explain how to construct for $\nu > 0$ a ν-optimal strategy on an interval $[t, s]$ for player II in the lower game. We apply the proof of superoptimality principle (i) to the lower value function V. We first choose $R > \sup_{t \le \tau \le s} \|X(\tau)\|$ and take $\lambda, \varepsilon, \beta$ small such that

$$|V(\tau, y) - V_{\lambda,\varepsilon,\beta}(\tau, y)| \le \frac{\nu}{4} \quad \text{for } \tau \in (0, T), \|y\| \le R$$

and $(s - t)\gamma_{R,\delta}(\lambda, \varepsilon, \beta) < \nu/4$. We then take n big enough and proceed as in Steps 2–4 of the proof of (i) to obtain a strategy $\beta^n \in \Delta(t)$ such that for every $W \in M(t)$

$$\begin{aligned} \frac{\nu}{4} + V_{\lambda,\varepsilon,\beta}(t, x) \ge & \int_t^s (L(\tau, X(\tau), W(\tau), \beta^n[W](\tau)) \\ & -\gamma_{R,\delta}(\lambda, \varepsilon, \beta))\mathrm{d}\tau + V_{\lambda,\varepsilon,\beta}(s, X(s)). \end{aligned}$$

This implies that

$$\nu + V(t, x) \ge \int_t^s L(\tau, X(\tau), W(\tau), \beta^n[W](\tau))\mathrm{d}\tau + V(s, X(s)),$$

i.e., β^n is ν-optimal on the interval $[t, s]$. It is a feedback strategy in a sense that $\beta^n[W](\tau)$ for $t + ih \le \tau < t + (i + 1)h, i = 0, 1, ..., n - 1$, depends on W and on $X(t + ih)$, where $h = (s - t)/n$.

Remark 2. Under a different set of assumptions which guarantee that value functions for the games with A replaced by bounded approximating operators (for instance the Yosida approximations or operators considered in [5]) converge to the original value functions, the same procedure can be used to construct nearly optimal feedback strategies for the approximating games with bounded operators and consequently for the original game. When the operator A is bounded, the assumption (8) about equicontinuity of the trajectories is automatically satisfied.

Acknowledgements A. Święch was supported by NSF grant DMS 0500270.

References

1. Berkovitz, L.: The existence of value and saddle point in games of fixed duration. SIAM J. Control Optim. **23**(2), 172–196 (1985). Erratum and addendum: SIAM J. Control Optim. **26**(3), 740–742 (1988)
2. Borwein, J.M., Preiss, D.: A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions. Trans. Am. Math. Soc. **303**(2), 517–527 (1987)
3. Clarke, F.H., Ledyaev, Yu.S., Sontag, E.D., Subbotin, A.I.: Asymptotic controllability implies feedback stabilization. IEEE Trans. Automat. Control **42**(10), 1394–1407 (1997)
4. Crandall, M.G., Lions, P.L.: Hamilton-Jacobi equations in infinite dimensions. II. Existence of viscosity solutions. J. Funct. Anal. **65**(3), 368–405 (1986)
5. Crandall, M.G., Lions, P.L.: Hamilton-Jacobi equations in infinite dimensions. IV. Hamiltonians with unbounded linear terms. J. Funct. Anal. **90**, 237–283 (1990)
6. Crandall, M.G., Lions, P.L.: Hamilton-Jacobi equations in infinite dimensions. V. Unbounded linear terms and B-continuous solutions. J. Funct. Anal. **97**, 417–465 (1991)
7. Crandall, M.G., Lions, P.L.: Hamilton-Jacobi equations in infinite dimensions. VI. Nonlinear A and Tataru's method refined. In: Evolution equations, control theory, and biomathematics (Han sur Lesse, 1991), pp. 51–89, Lecture Notes in Pure and Appl. Math. **155**, Dekker, New York (1994)
8. Elliott, R.J., Kalton, N.J.: The existence of value in differential games. Memoirs of the American Mathematical Society, **126**. American Mathematical Society, Providence (1972)
9. Evans, L.C., Souganidis, P.E.: Differential games and representation formulas for solutions of Hamilton-Jacobi equations. Indiana Univ. Math. J. **33**, 773–797 (1984)
10. Fabbri, G., Gozzi, F., Święch, A.: Verification theorem and construction of ε-optimal controls for control of abstract evolution equations. J. Convex Anal. **17**(2), 611–642 (2010)
11. Fleming, W.H., Souganidis, P.E.: On the existence of value functions of two-player, zero-sum stochastic differential games. Indiana Univ. Math. J. **38**(2), 293–314 (1989)
12. Ghosh, M.K., Shaiju, A.J.: Existence of value and saddle point in infinite-dimensional differential games. J. Optim. Theory Appl. **121**(2), 301–325 (2004)
13. Ishii, H.: Viscosity solutions for a class of Hamilton–Jacobi equations in Hilbert spaces. J. Funct. Anal. **105**, 301–341 (1992)
14. Kelome, D., Święch, A.: Perron's method and the method of relaxed limits for "unbounded" PDE in Hilbert spaces. Studia Math. **176**(3), 249–277 (2006)
15. Kocan, M., Soravia, P., Święch, A.: On differential games for infinite-dimensional systems with nonlinear, unbounded operators. J. Math. Anal. Appl. **211**(2), 395–423 (1997)
16. Li, X.Y., Yong, J.M.: Optimal control theory for infinite-dimensional systems. Birkhauser, Boston, Cambridge, (1995)
17. Ramaswamy, M., Shaiju, A.J.: Construction of approximate saddle point strategies for differential games in a Hilbert space. J. Optim. Theory Appl. **141**(2), 349–370 (2009)
18. Renardy, M.: Polar decomposition of positive operators and problem od Crandall and Lions. Appl. Anal. **57**(3), 383–385 (1995)
19. Shaiju, A.J.: Infinite horizon differential games for abstract evolution equations. Comput. Appl. Math. **22**(3), 335–357 (2003)
20. Soravia, P.: Feedback stabilization and H^∞ control of nonlinear systems affected by disturbances: The differential games approach. Dynamics, bifurcations, and control (Kloster Irsee, 2001), 173–190, Lecture Notes in Control and Inform. Sci. **273**, Springer, Berlin (2002)
21. Święch, A.: Sub- and superoptimality principles of dynamic programming revisited. Nonlinear Anal. **26**(8), 1429–1436 (1996)
22. Święch, A.: Another approach to the existence of value functions of stochastic differential games. J. Math. Anal. Appl. **204**(3), 884–897 (1996)
23. Tataru, D.: Viscosity solutions of Hamilton–Jacobi equations with unbounded linear terms. J. Math. Anal. Appl. **163**, 345–392 (1992)
24. Tataru, D.: Viscosity solutions for Hamilton-Jacobi equations with unbounded nonlinear term: A simplified approach. J. Differ. Equ. **111**(1), 123–146 (1994)

Part II
Pursuit Evasion Games

1. *Sergei A. Ganebnyi, Sergey S. Kumkov, Valerii S. Patsko and Sergei G. Pyatko*, Extremal Aiming in Problems with Unknown Level of Dynamic Disturbance
2. *Ryusuke Hohzaki*, A Generalization of the Multi-Stage Search Allocation Game
3. *Stéphane Le Ménec*, Linear Differential Game with Two Pursuers and One Evader
4. *Valerii S. Patsko and Varvara L. Turova,* Homicidal Chauffeur Game: History and Modern Studies
5. *Sriram Shankaran, Dušan M. Stipanović and Claire J. Tomlin*, Collision Avoidance Strategies for a Three-Player Game

Extremal Aiming in Problems with Unknown Level of Dynamic Disturbance

Sergei A. Ganebny, Sergey S. Kumkov, Valerii S. Patsko, and Sergei G. Pyatko

Abstract The procedure of extremal aiming (shift), which is well known in the theory of differential games, is applied to problems where no constraint for dynamic disturbance is given *a priori*. Problems are considered with linear dynamics, fixed terminal time, and geometric constraint for the useful control. The objective of the useful control is to guide the system to a given terminal set at the termination instant. A method for constructing a feedback control is suggested, which is adaptive with respect to the disturbance level. The method guarantees guiding the system to the terminal set if the disturbance is not larger than some critical level. With that, a "weak" disturbance is parried by a "weak" useful control. A theorem about the guarantee is formulated and proved. The method is applied to the problem of aircraft landing under wind disturbances.

Keywords Adaptive control · Aircraft landing problem · Differential games · Wind disturbance.

1 Introduction

In the theory of antagonistic differential games [1, 12, 17, 18], there are well-developed methods for problems, which formulation includes geometric constraints for both players. But in many practical situations, a geometric constraint is given only for the useful control (for the first player's control), whereas, defining any strict constraint for the dynamic disturbance (for the second player's control) is not natural. Furthermore, an optimal feedback control for the first player obtained within the

S.A. Ganebny, S.S. Kumkov, and V.S. Patsko (✉)
Institute of Mathematics and Mechanics, Ural Branch of Russian Academy of Sciences, S. Kovalevskaya str. 16, Ekaterinburg, Russia
e-mail: ganebny@imm.uran.ru; patsko@imm.uran.ru

S.G. Pyatko
Academy of Civil Aviation, Pilotov str., 38, St. Petersburg, 196210, Russia
e-mail: boss@nita.ru; nita@nita.ru

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_9,

framework of standard formalization of an antagonistic differential game supposes that the opponent acts in the worst way. But in practical situations, the disturbance in most cases acts nonoptimally.

So, one wants to have a feedback control method, which works successfully for a wide range of the disturbances. With that, the "weaker" or "less optimal" the disturbance is, the "weaker" the parrying control should be. The aim of the paper is to suggest such a method based on the existing theory of differential games.

The central concept used in this paper is the concept of stable bridge [17,18] from the theory of differential games. A stable bridge is a set in the space *time* $\times$ *phase vector*, in which the first player using his control and discriminating the opponent can keep the motion of the system until the termination instant. With that, it is assumed that geometric constraints are given both for the first and for the second players' controls. A concept equivalent to "stable bridge" is the "discriminating kernel" [5].

Imagine a family of differential games having the same dynamics and parametrized by a scalar nonnegative parameter. This parameter defines three families of sets increasing with the growth of the parameter: constraints for the first player's control, constraints for the second player's control, and terminal sets. Assume that the family of the stable bridges corresponding to the dynamics and a triple of pair of constraints and terminal set is also increasing by inclusion. The first player guarantees keeping the system inside a bridge from this family (and, therefore, guiding the system to the corresponding terminal set) if the disturbance is bounded by the corresponding constraint. With that, the parrying useful control obeys its corresponding constraint. So, such a family of stable bridges allows to construct a feedback control for the first player and describe the guarantee provided by this control.

We explain how the adaptation is realized. Let the system be affected by a disturbance, which is bounded by a constraint from the corresponding family. Then the motion of the system under the disturbance action and the mentioned useful control will cross bridges from the built family until it reaches the boundary of the bridge (from above or below), which corresponds to the constraint for the disturbance. After this, the motion will stay within this bridge. Thus, the self-tuning (self-adaptation) of the useful control to the unknown *a priori* level of the disturbance takes place.

The idea of the ordered family of stable bridges is quite generic. Its realization depends on capability of analytic or numerical construction of stable bridges. In the theory of differential games, there are many works [3, 5, 8, 11, 24, 26–30] dealing with numerical algorithms for constructing maximal stable bridges and level sets (Lebesgue sets) of the value function. These methods can be used for constructing such a family of stable bridges.

In this chapter, we consider problems with linear dynamics, fixed terminal time, and convex compact terminal set, to which the first player tries to guide the system. The vector useful control is bounded by a convex compact geometric constraint. These features of the problem allow us to easily construct such a family of stable bridges and corresponding adaptive control. Namely, one should construct in advance and store two maximal stable bridges. On their basis during the motion process, the t-sections of an appropriate stable bridge from the family are computed. The first player's control is generated by extremal shift [17, 18] to the computed section. Efficiency of the algorithm is provided by the fact that all t-sections of the bridges from the family are convex. We formulate and prove a theorem about the guarantee provided by the suggested method to the first player.

A corresponding algorithm is realized [9, 10, 14] for the case when the terminal set is defined by two or three components of the phase vector at the termination instant.

The final part of the paper is devoted to a simulation of the problem of control by an aircraft during landing under wind disturbances.

This paper is close to the previous work [9, 10] of the authors, where the case of scalar useful control was considered.

2 Problem Formulation

Consider a linear differential game with fixed terminal time:

$$\begin{aligned} &\dot{z} = A(t)z + B(t)u + C(t)v, \\ &z \in R^m,\ t \in T,\ u \in P \subset R^P,\ v \in R^q. \end{aligned} \tag{1}$$

Here, u and v are vector controls of the first and second players, respectively; P is a convex compact constraint for the first player's control; $T = [\vartheta_0, \vartheta]$ is the time interval of the game. Assume that the set P contains the origin of the space R^p. The matrix functions A and C are continuous on t. The matrix function B is Lipschitzian in the time interval T. There are no constraints for the control v.

The first player tries to guide n chosen components of the phase vector of system (1) to the terminal set M at the termination instant. The set M is supposed to be convex compactum in the space of these n components of the phase vector z. Assume that the set M contains some neighborhood of the origin of this space. The origin is called the center of the set M. The first player is interested in transferring the chosen components of the vector z as close as possible to this center.

A method for constructing an adaptive control for system (1) is required.

Let us consider a system which right-hand side does not include the phase vector:

$$\begin{aligned} &\dot{x} = D(t)u + E(t)v, \\ &x \in R^n,\ t \in T,\ u \in P \subset R^P,\ v \in R^q. \end{aligned} \tag{2}$$

The passage is provided ([17, p. 160], [18, pp. 89–91]) by means of the relations

$$x(t) = Z_{n,m}(\vartheta, t)z(t),\quad D(t) = Z_{n,m}(\vartheta, t)B(t),\quad E(t) = Z_{n,m}(\vartheta, t)C(t),$$

where $Z_{n,m}(\vartheta, t)$ is a matrix combining n rows of the fundamental Cauchy matrix for the system $\dot{z} = A(t)z$ that correspond to the components of the vector z defining the terminal set. In system (2), the first player tries to guide the phase vector to the set M at the termination instant ϑ.

All the following computations will be made for system (2). The constructed adaptive control $U(t,x)$ is applied to the original system (1) as $U\big(t, Z_{n,m}(\vartheta,t)z\big)$.

3 Family of Stable Bridges

Below, the symbol $S(t) = \{x \in R^m : (t,x) \in S\}$ denotes a section of a set $S \subset T \times R^n$ at instant $t \in T$. Let $O(\varepsilon) = \{x \in R^n : |x| \le \varepsilon\}$ be a ball with radius ε and center at the origin in the space R^n.

In the interval $[\vartheta_0, \vartheta]$, consider an antagonistic differential game with terminal set $\mathscr{M}$ and geometric constraints $\mathscr{P}$, $\mathscr{Q}$ for the players' controls:

$$\begin{aligned} &\dot{x} = D(t)u + E(t)v, \\ &x \in R^n,\ t \in T,\ \mathscr{M},\ u \in \mathscr{P},\ v \in \mathscr{Q}. \end{aligned} \tag{3}$$

Here, matrices $D(t)$ and $E(t)$ are the same as in system (2). The sets $\mathscr{M}$, $\mathscr{P}$, $\mathscr{Q}$ are assumed to be convex compacta. They are considered as parameters of the game.

Below, $u(\cdot)$ and $v(\cdot)$ denote measurable functions of time with values in the sets $\mathscr{P}$ and $\mathscr{Q}$, respectively. A motion of system (3) (and, therefore, of system (2)) emanating from point x_* at instant t_* under controls $u(\cdot)$ and $v(\cdot)$ is denoted by $x\big(\cdot; t_*, x_*, u(\cdot), v(\cdot)\big)$.

Following [17, 18], define the concepts of stable and maximal stable bridges.

A set $W \subset T \times R^n$ is called the *stable bridge* for system (3) for some chosen sets $\mathscr{P}$, $\mathscr{Q}$, and $\mathscr{M}$ if $W(\vartheta) = \mathscr{M}$ and the following *stability* property is satisfied: for any position $(t_*, x_*) \in W$ and any control $v(\cdot)$ of the second player, the first player can choose his control $u(\cdot)$ such that the position $\big(t, x(t)\big) = \big(t, x(t; t_*, x_*, u(\cdot), v(\cdot))\big)$ stays in the set W for any instant $t \in (t_*, \vartheta]$. The set $W \subset T \times R^n$, $W(\vartheta) = \mathscr{M}$, possessing the stability property, and maximal by inclusion is called the *maximal stable bridge*.

The maximal stable bridge is a closed set [17, 18]. Its t-sections are convex due to the linearity of system (3) and convexity of the terminal set $\mathscr{M}$ ([17, p. 87]).

1°. Choose some set $Q_{\max} \subset R^q$ treated as "maximal" or "critical" constraint for the second player's control, which can be considered by the first player as "reasonable" when guiding system (2) to the set M. Assume that the set $Q_{\max}$ contains the origin of its space. Such an assumption is not too burdensome, because under absence of disturbance, guiding should be possible. Denote by W_{main} the maximal stable bridge for system (3) corresponding to the parameters $\mathscr{P} = P$, $\mathscr{Q} = Q_{\max}$, and $\mathscr{M} = M$.

Also assume the set $Q_{\max}$ is chosen such that for some $\varepsilon > 0$ for any $t \in T$ the following inclusion is true:

$$O(\varepsilon) \subset W_{\text{main}}(t). \tag{4}$$

Below, the parameter ε is supposed fixed.

Thus, W_{main} is a closed tube in the space $T \times R^n$ reaching the set M at instant ϑ. Any of its t-section is convex and contains ε-neighborhood of the origin of the space R^n.

2°. Introduce another closed tube $W_{\text{add}} \subset T \times R^n$, whose sections $W_{\text{add}}(t)$ coincide with an attainability set of system (3) at instant t from the initial set $O(\varepsilon)$ taken at instant ϑ_0. When constructing the tube W_{add}, assume that the first player is absent ($u \equiv 0$) and the second player's control is constrained by the set $Q_{\max}$. One can see that W_{add} is the maximal stable bridge for system (3) with $\mathscr{P} = \{0\}$, $\mathscr{Q} = Q_{\max}$, $\mathscr{M} = W_{\text{add}}(\vartheta)$. For any $t \in T$, the section $W_{\text{add}}(t)$ is convex and satisfies the inclusion

$$O(\varepsilon) \subset W_{\text{add}}(t). \tag{5}$$

3°. Consider a family of tubes $W_k \subset T \times R^n$ whose sections $W_k(t)$ are defined as

$$W_k(t) = \begin{cases} k W_{\text{main}}(t), & 0 \le k \le 1, \\ W_{\text{main}}(t) + (k-1) W_{\text{add}}(t), & k > 1. \end{cases} \tag{6}$$

The sets $W_k(t)$ are compact and convex. For any numbers $0 \le k_1 < k_2 \le 1 < k_3 < k_4$ due to relations (4), (5) the following strict inclusions are true:

$$W_{k_1}(t) \subset W_{k_2}(t) \subset W_{k_3}(t) \subset W_{k_4}(t).$$

In [10] and [9], it is shown that the tube W_k when $0 \le k \le 1$ is the maximal stable bridge for system (3) corresponding to the constraint kP for the first player's control, the constraint $kQ_{\max}$ for the second player's control, and the terminal set kM. When $k > 1$, the set W_k is a stable bridge (but generally not a maximal one) for the parameters $\mathscr{P} = P$, $\mathscr{Q} = kQ_{\max}$, $\mathscr{M} = M + (k-1)W_{\text{add}}(\vartheta)$.

So, one has a growing family of stable bridges, where each larger bridge corresponds to a larger constraint for the second player's control. This family is generated by two bridges W_{main} and W_{add} by means of algebraic summation and multiplication by a nonnegative numeric parameter according to relations (6).

Let

$$P_k = \min\{k, 1\} \cdot P, \quad k \ge 0.$$

Define a function $V : T \times R^n \to R$ as $V(t, x) = \min\{k \ge 0 : (t, x) \in W_k\}$. Relations (4) and (5) guarantee that the function $x \mapsto V(t, x)$ for any $t \in T$ is Lipschitzian with a constant $\lambda = 1/\varepsilon$.

4 Adaptive Feedback Control

The adaptive control $(t, x) \mapsto U(t, x)$ is constructed as follows.

Fix a number $\xi > 0$. Consider an arbitrary position (t, x). If $|x| \leq \xi$, let $U(t, x) = 0$. If $|x| > \xi$, find a positive number k^* defining a bridge W_{k^*}, whose section $W_{k^*}(t)$ is away from the point x at the distance ξ. On the boundary of the set $W_{k^*}(t)$ compute the point x^* closest to x. We have $|x - x^*| = \xi$. Define a vector $u^* \in P_{k^*}$ on the basis of the extremum condition:

$$(x^* - x)' D(t) u^* = \max\bigl\{(x^* - x)' D(t) u : u \in P_{k^*}\bigr\}. \tag{7}$$

Let $U(t, x) = u^*$.

This rule for generating the control U can be written in terms of the function V in the following way. In the closed ξ-neighborhood of the point x, find a point x^*, where the function $V(t, \cdot)$ takes its local minimum. Let $k^* = V(t, x^*)$. The control $U(t, x)$ is chosen from the set P_{k^*} by relation (7) along the vector $x^* - x$.

The control U is applied in a discrete scheme [17, 18] with time-step Δ. The choice of the control is implemented at the initial instant of the next time-step Δ. The control is maintained constant until the end of this step.

So, the control (strategy) U is built by the extremal shift rule [17, 18], which is well known in the theory of differential games. We modify this rule for a problem where no geometric constraint is given *a priori* for the second player's control.

When using the strategy U, the level of the first player's control adjusts to the actual level of the second player's control (of the dynamic disturbance). Indeed, when the disturbance is "weak", the motion comes inside the family of the stable bridges, and, therefore, the current index k^* decreases. Vice versa, if the disturbance is strong, this index increases. The index k^* defines the bridge for extremal shift and the level of the constraint for the useful control, where the control is chosen by strategy U.

5 Guarantee Theorem

We formulate the theorem describing the guarantee of the first player under the use of the suggested adaptive control.

Let

$$d = \max_{t \in [\vartheta_0, \vartheta]} \max_{u \in P} \bigl|D(t)u\bigr|.$$

Below, we need a Lipschitz condition for the map $t \mapsto D(t)$. Take its Lipschitz constant β as the maximum of Lipschitz constants of the functions $t \mapsto D_j(t)$, where $D_j(t)$ is the jth column of the matrix $D(t)$, $j = \overline{1, p}$.

Let κ be the maximum of coordinate deviations of the set P from the origin:

$$\kappa = \max_j \max\bigl\{|u_j| : u \in P\bigr\}.$$

Theorem 1. *Let $\xi > 0$ and U be the adaptive strategy of the first player, which shifts extremally with the distance ξ. Choose arbitrary $t_0 \in T$, $x_0 \in R^n$, and $\Delta > 0$.*

Suppose that a control $v(\cdot)$ of the second player is bounded by a set $c^* Q_{\max}$, $c^* \geq 0$, in the interval $[t_0, \vartheta]$. Denote

$$s^* = \max\{c^*, V(t_0, x_0)\}.$$

Let $x(\cdot)$ be a motion of system (2) emanating at instant t_0 from point x_0 under the strategy U in the discrete scheme with step Δ and the control $v(\cdot)$ of the second player. Then, the realization $u(t) = U(t, x(t))$ of the first player's control satisfies the inclusion

$$u(t) \in \min\{s^* + E(t, t_0, \Delta, \xi), 1\} \cdot P, \quad t \in [t_0, \vartheta]. \tag{8}$$

With that, the value $V(t, x(t))$ of the function V obeys the inequality

$$V(t, x(t)) \leq s^* + E(t, t_0, \Delta, \xi) + \lambda\xi, \quad t \in [t_0, \vartheta]. \tag{9}$$

Here,

$$E(t, t_0, \Delta, \xi) = \lambda\Delta(t - t_0)\left(p\beta\kappa + \frac{4d^2 + (p\beta\kappa\Delta)^2}{2\xi}\right) + 2\lambda d\Delta.$$

Proof. Let $x^*(t)$ be the point of extremal shift from the point $x(t)$. To get estimations (8) and (9), it is enough to prove the estimation

$$V(t, x^*(t)) \leq s^* + E(t, t_0, \Delta, \xi). \tag{10}$$

Inclusion (8) follows from (10) by taking into account the extremal shift rule and the fact that the mapping $t \to E(t, t_0, \Delta, \xi)$ is monotone increasing. Inequality (9) follows from the fact that the distance between points $x(t)$ and $x^*(t)$ is not greater than ξ.

(1) In some interval $[t_*, t_* + \delta)$, let the first player use some constant control u^*. This is generated at instant t_* from the extremal shift to the section $W_{k^*}(t_*)$ of a stable bridge corresponding to the index k^*. Let $x^*(t_*)$ be the point of the set $W_{k^*}(t_*)$ closest to $x(t_*)$. We have that $|x(t_*) - x^*(t_*)| \leq \xi$. The unit vector $l(t_*)$ of the extremal shift is directed from the point $x(t_*)$ to the point $x^*(t_*)$ (Fig. 1a).
Assume that the disturbance (the second player's control) has level not greater than $k^* Q_{\max}$, which corresponds to the bridge W_{k^*}. Let $v(\cdot)$ be the realization of the disturbance in the interval $[t_*, t_* + \delta)$. From the stability property of the set W_{k^*}, choose a measurable function $u(\cdot)$ with values from the set P_{k^*} so that the motion from point $x^*(t_*)$ goes in sections $W_{k^*}(t)$ in the interval $[t_*, t_* + \delta)$. Let b be the position reached by this motion at instant $t_* + \delta$.

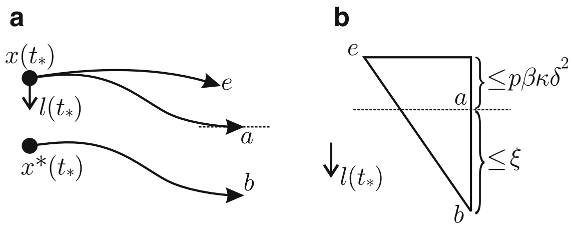


Fig. 1 Proof of Theorem 1

The symbol e denotes the position at instant $t_* + \delta$ of the motion from the point $x(t_*)$ under the constant control u^* and the disturbance $v(\cdot)$. Estimate the distance $r(\delta)$ between points e and b.

Consider an auxiliary motion from point $x(t_*)$, which is the copy of the motion emanating from point $x^*(t_*)$ and generated under the stability property. Denote by a its position at instant $t_* + \delta$.

There are three cases:

$$l'(t_*)e > l'(t_*)b, \qquad l'(t_*)b \ge l'(t_*)e \ge l'(t_*)a, \qquad l'(t_*)e < l'(t_*)a.$$

In the first case, one has $r(\delta) = |e - b| \le |e - a| \le 2d\delta$. This estimation takes into account the fact that points e and a are generated by motions with the same initial position and under the same control $v(\cdot)$.

In the second case, $r(\delta) \le (\xi^2 + (2d\delta)^2)^{1/2}$.

For the third case, we write

$$l'(t_*)(a - e) = l'(t_*)\left(\int_{t_*}^{t_*+\delta} D(t)u(t)\mathrm{d}t - \int_{t_*}^{t_*+\delta} D(t)u^*\mathrm{d}t\right)$$
$$= l'(t_*)\int_{t_*}^{t_*+\delta} D(t_*)\big(u(t) - u^*\big)\mathrm{d}t + l'(t_*)\int_{t_*}^{t_*+\delta} \big(D(t) - D(t_*)\big)\big(u(t) - u^*\big)\mathrm{d}t.$$

Since the control value u^* is chosen accordingly to the extremal shift rule along the vector $l(t_*)$, then

$$l'(t_*)\int_{t_*}^{t_*+\delta} D(t_*)\big(u(t) - u^*\big)\mathrm{d}t \le 0.$$

Therefore,

$$l'(t_*)(a - e) \le l'(t_*)\int_{t_*}^{t_*+\delta} \big(D(t) - D(t_*)\big)\big(u(t) - u^*\big)\mathrm{d}t.$$

Further, one has

$$l'(t_*) \int\limits_{t_*}^{t_*+\delta} \big(D(t) - D(t_*)\big)\big(u(t) - u^*\big)\mathrm{d}t$$

$$= l'(t_*) \int\limits_{t_*}^{t_*+\delta} \sum_{j=1}^{p} \big(D_j(t) - D_j(t_*)\big)\big(u_j(t) - u_j^*\big)\mathrm{d}t.$$

The right-hand side can be estimated from above by the number $p\beta\delta^2\kappa$. Finally,

$$l'(t_*)(a - e) \leq p\beta\delta^2\kappa.$$

Consequently, in the third case, the distance $r(\delta)$ between points e and b satisfies the inequality (Fig. 1b)

$$r(\delta) \leq \sqrt{(\xi + p\beta\delta^2\kappa)^2 + (2d\delta)^2}. \tag{11}$$

Comparing the estimations for all three cases, we conclude that estimation (11) can be taken as the universal one.

(2) Let us change the estimation (11) with square root to linear one, which will be more convenient (Fig. 1c). Namely, for all $\widetilde{y} > 0$ and $y \geq 0$, it is true that

$$\sqrt{y} \leq \sqrt{\widetilde{y}} + \frac{1}{2\sqrt{\widetilde{y}}}\,(y - \widetilde{y}). \tag{12}$$

In our case, let us take $\widetilde{y} = \xi^2$. Due to (11) and (12), it follows that

$$r(\delta) \leq \xi + \frac{1}{2\xi}\Big((\xi + p\beta\delta^2\kappa)^2 + (2d\delta)^2 - \xi^2\Big) = \xi + \delta^2\left(p\beta\kappa + \frac{(p\beta\kappa\delta)^2}{2\xi} + \frac{2d^2}{\xi}\right).$$

Denote

$$\eta(\delta) = p\beta\kappa + \frac{(p\beta\kappa\delta)^2}{2\xi} + \frac{2d^2}{\xi}.$$

Thus, the distance from the point $x(t_*+\delta)$ to the section $W_{k^*}(t_*+\delta)$ is estimated by the inequality

$$r(\delta) \leq \xi + \delta^2\eta(\delta). \tag{13}$$

(3) Compute the minimum of the function $V(t_* + \delta, \cdot)$ in a ball with radius ξ and center at the point $x(t_* + \delta) = e$. Let f be the point where the minimum is reached.

Suppose that point $b \in W_{k^*}(t_* + \delta)$ mentioned in item (1) is located outside the ball with radius ξ and center e. Let h be a point of intersection the boundary of this ball and the segment $[be]$. Using (13), we conclude that

$$|h - b| = r(\delta) - \xi \leq \delta^2 \eta(\delta).$$

One has

$$V(t_* + \delta, f) \leq V(t_* + \delta, h) \leq V(t_* + \delta, b) + \lambda \delta^2 \eta(\delta).$$

Since $V(t_* + \delta, b) \leq V(t_*, x^*(t_*))$, then

$$V(t_* + \delta, f) \leq V(t_*, x^*(t_*)) + \lambda \delta^2 \eta(\delta). \tag{14}$$

Let point b belong to the ball with radius ξ and center e. Then,

$$V(t_* + \delta, f) \leq V(t_* + \delta, b) \leq V(t_*, x^*(t_*)).$$

So, an increment in V for the aiming point is estimated by inequality (14).

(4) Let t be an arbitrary instant from the interval $[t_0, \vartheta]$. To prove inequality (10), let us estimate the variation of the function $\tau \to V(\tau, x^*(\tau))$ in the interval $[t_0, t]$. If $V(t, x^*(t)) \leq s^*$ at instant t, then inequality (10) is obviously true. Assume $V(t, x^*(t)) > s^*$. Let $\tilde{t}$ be the instant of the last reaching the level s^* of the function V by the point $x^*(\tau)$ in the interval $[t_0, t)$.
The choice of the first player's control is realized in discrete instants. Suppose that there is at least one discrete instant in the interval $[\tilde{t}, t]$. The discrete instant closest to $\tilde{t}$ from the right (closest to t from the left) is denoted by $\bar{t}$ (respectively, $\hat{t}$).
In $[\bar{t}, \hat{t}]$, there are entire discrete intervals only. Their number is $(\hat{t} - \bar{t})/\Delta$. Taking into account (14), we get the estimation

$$V(\hat{t}, x^*(\hat{t})) \leq V(\bar{t}, x^*(\bar{t})) + \lambda(\hat{t} - \bar{t})\Delta\eta(\Delta). \tag{15}$$

In $[\hat{t}, t]$, again with (14), one gets

$$V(t, x^*(t)) \leq V(\hat{t}, x^*(\hat{t})) + \lambda(t - \hat{t})^2\eta(t - \hat{t}) \leq V(\hat{t}, x^*(\hat{t})) + \lambda(t - \hat{t})\Delta\eta(\Delta). \tag{16}$$

It remains to estimate the increment $V(\tau, x^*(\tau))$ in $[\tilde{t}, \bar{t}]$.
At instant $\tilde{t}$, one has $V(\tilde{t}, x^*(\tilde{t})) = s^*$. By the condition Theorem 1, the level of the disturbance $v(\cdot)$ is not greater than $c^* \leq s^*$. For the disturbance $v(\cdot)$, find a parrying control $u_{\mathrm{st}}(\cdot)$ such that the motion $x_{\mathrm{st}}(\cdot)$ emanated from the point $x^*(\tilde{t})$ under the controls $u_{\mathrm{st}}(\cdot)$ and $v(\cdot)$ goes in the bridge W_{s^*} in the interval $[\tilde{t}, \bar{t}]$.

The difference of two motions is

$$x(\bar{t}) - x_{\mathrm{st}}(\bar{t}) = x(\tilde{t}) - x^*(\tilde{t}) + \int_{\tilde{t}}^{\bar{t}} D(\tau)\big(u(\tau) - u_{\mathrm{st}}(\tau)\big)\mathrm{d}\tau.$$

One has

$$\left|\int_{\tilde{t}}^{\bar{t}} D(\tau)\big(u(\tau) - u_{\mathrm{st}}(\tau)\big)\mathrm{d}\tau\right| \le 2d(\bar{t} - \tilde{t}).$$

Therefore,

$$|x(\bar{t}) - x_{\mathrm{st}}(\bar{t})| \le |x(\tilde{t}) - x^*(\tilde{t})| + 2d(\bar{t} - \tilde{t}) \le \xi + 2d(\bar{t} - \tilde{t}). \tag{17}$$

Let the point $x_{\mathrm{st}}(\bar{t})$ be located outside the ball with radius ξ and center $x(\bar{t})$. Consider a point h, which is the intersection of the boundary of this ball and the segment connecting the points $x(\bar{t})$ and $x_{\mathrm{st}}(\bar{t})$. Taking into account (17), one gets

$$V(\bar{t}, x^*(\bar{t})) \le V(\bar{t}, h) \le V(\bar{t}, x_{\mathrm{st}}(\bar{t})) + 2\lambda d(\bar{t} - \tilde{t}) \le s^* + 2\lambda \mathrm{d}\Delta.$$

Let point $x_{\mathrm{st}}(\bar{t})$ belong to the ball with radius ξ and center $x(\bar{t})$. Then, $V(\bar{t}, x^*(\bar{t})) \le V(\bar{t}, x_{\mathrm{st}}(\bar{t})) \le s^*$.
The following estimation is then true:

$$V(\bar{t}, x^*(\bar{t})) \le s^* + 2\lambda d\Delta. \tag{18}$$

Gathering estimations (15), (16), and (18), we obtain

$$V(t, x^*(t)) \le s^* + 2\lambda d\Delta + \lambda(t - \bar{t})\Delta\eta(\Delta).$$

Taking into account that $E(t, t_0, \Delta, \xi) = \lambda(t - t_0)\Delta\eta(\Delta) + 2\lambda d\Delta$, we get estimation (10).
Now let there be no discrete instants in the interval $[\tilde{t}, t]$. Then the following estimation holds for the instant t: $V(t, x^*(t)) \le s^* + 2\lambda\Delta$, analogous to estimation (18). Estimation (10) holds obviously.

□

6 Aircraft Landing under Wind Disturbances

Aircraft taking-off and landing stages are the most sensitive to wind disturbances.

Application of modern methods of the mathematical control theory and differential game theory to the taking-off, landing, and abort landing problems in the

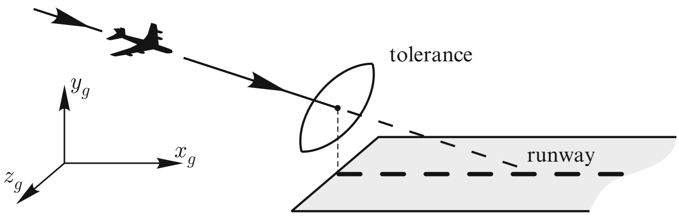


Fig. 2 Aircraft landing problem

presence of wind disturbances was stimulated by papers [20–22] of A. Miele and his collaborators, and also by works [15, 16] of V.M. Kein. Among the investigations on this topic, let us note the articles [2, 4, 6, 19, 23, 25].

In this section, we apply the method of adaptive control described above to a problem of aircraft landing in the presence of wind disturbances.

The landing process is investigated to the time when the runway threshold is reached (Fig. 2). At the instant of passing the threshold, it is required that the main characteristics of the lateral and longitudinal (vertical) channels to be in some tolerances. These characteristics for the lateral channel are the lateral deviation and its velocity, for the vertical channel, they are the vertical deviation and its velocity. About wind, we know the nominal (average) values of its velocity components. The exact constraints for the deviations of the velocity components from these values are not given.

We use the usual notation in the Russian aviation literature.

6.1 Nonlinear System and Problem Formulation

Consider an aircraft motion during landing until passing the threshold of the runway. The nominal motion of the aircraft is uniform descending along the rectilinear glide path. The height of the glide path above the runway threshold is 15 m.

The motion of the aircraft is described by a differential system of the 12th order:

$$
\begin{aligned}
\dot{x}_g &= V_{xg},\\
\dot{V}_{xg} &= [(P\cos\sigma - qSc_x)\cos\psi\cos\vartheta + (P\sin\sigma + qSc_y)\\
&\qquad \times(\sin\psi\sin\gamma - \cos\gamma\cos\psi\sin\vartheta)\\
&\qquad + qSc_z(\sin\psi\cos\gamma + \cos\psi\sin\vartheta\sin\gamma)]/m,\\
\dot{y}_g &= V_{yg},\\
\dot{V}_{yg} &= [(P\cos\sigma - qSc_x)\sin\vartheta + (P\sin\sigma + qSc_y)\cos\vartheta\cos\gamma\\
&\qquad - qSc_z\cos\vartheta\sin\gamma]/m - g,\\
\dot{z}_g &= V_{zg},
\end{aligned}
$$

$$
\begin{aligned}
\dot{V}_{zg} &= [(P\cos\sigma - qSc_x)(-\sin\psi\cos\vartheta) + (P\sin\sigma + qSc_y) \\
&\quad \times(\cos\psi\sin\gamma + \sin\psi\sin\vartheta\cos\gamma) \\
&\quad + qSc_z(\cos\psi\cos\gamma - \sin\psi\sin\vartheta\sin\gamma)]/m, \\
\dot{\vartheta} &= \omega_z\cos\gamma + \omega_y\sin\gamma, \\
\dot{\omega}_z &= [I_{xy}(\omega_x^2 - \omega_y^2) - (I_y - I_x)\omega_x\omega_y + M_z]/I_z, \\
\dot{\psi} &= (\omega_y\cos\gamma - \omega_z\sin\gamma)/\cos\vartheta, \\
\dot{\omega}_y &= [(I_y - I_z)I_{xy}\omega_y\omega_z + (I_z - I_x)I_x\omega_x\omega_z + I_xM_y \\
&\quad + I_{xy}M_x + I_{xy}\omega_z(I_x\omega_y - I_{xy}\omega_x)]/J, \\
\dot{\gamma} &= \omega_x - (\omega_y\cos\gamma - \omega_z\sin\gamma)\tan\vartheta, \\
\dot{\omega}_x &= [(I_y - I_z)I_y\omega_y\omega_z + (I_z - I_x)I_{xy}\omega_x\omega_z + I_yM_x + I_{xy}M_y \\
&\quad + I_{xy}\omega_z(I_{xy}\omega_y - I_y\omega_x)]/J.
\end{aligned} \tag{19}
$$

Here,

- x_g, y_g, z_g = longitudinal, vertical, and lateral coordinates of the aircraft mass center, m, ground-fixed system.
- V_{xg}, V_{yg}, V_{zg} = longitudinal, vertical, and lateral absolute velocity components, m/s.
- ϑ, ψ, γ = pitch, yaw, and bank angles, rad.
- $\omega_x, \omega_y, \omega_z$ = angular velocities, rad/s, body-axes system.
- $q = \rho\hat{V}^2/2$ = dynamic pressure, kg m^{-1}s^{-2}.
- ρ = air density, kg m^{-3}.
- $\hat{V}$ = relative velocity, m/s.
- S = reference surface, m^2.
- I_x, I_y, I_z, I_{xy} = inertia moments, kg m^2.
- $J = I_xI_y - I_{xy}^2$.
- $M_x = qSlm_x$, $M_y = qSlm_y$, $M_z = qSbm_z$ = aerodynamic moments, N m.
- l = wing span, m.
- b = mean aerodynamic chord, m.
- c_x, c_y, c_z = aerodynamic force coefficients.
- m_x, m_y, m_z = aerodynamic moment coefficients, body-axes system.
- m = aircraft mass, kg.
- g = acceleration of gravity, m/s^2.
- σ = thrust inclination, grad.
- P = thrust force, N.

The aircraft controls are the thrust force P, the elevator δ_e, rudder δ_r, and aileron δ_a deflections. The quantities δ_a, δ_r, and δ_e enter to the definition of aerodynamic coefficients. The latter also depend on the attack and sideslip angles. The components of the wind disturbance along the axes x_g, y_g, z_g affect the components $\hat{V}_{xg}$, $\hat{V}_{yg}$, $\hat{V}_{zg}$ of the aerial (relative) velocity: $\hat{V}_{xg} = V_{xg} - W_{xg}$, $\hat{V}_{yg} = V_{yg} - W_{yg}$, $\hat{V}_{zg} = V_{zg} - W_{zg}$.

We describe the simplest way of accounting for the inertiality of the servomechanisms. Assume that the change of the thrust force obeys the equation

$$\begin{aligned} \dot{P} &= -k_p P + \bar{k}_p(\delta_{ps} + \bar{\delta}_p), \\ \bar{\delta}_p &= -41.3^\circ, \quad 47^\circ \le \delta_{ps} \le 112^\circ. \end{aligned} \tag{20}$$

Here, δ_{ps} is the engine control level setting. The servomechanism dynamics for the control surfaces is

$$\begin{aligned} &\dot{\delta}_e = k_e(\delta_{es} - \delta_e), \quad \dot{\delta}_r = k_r(\delta_{rs} - \delta_r), \quad \dot{\delta}_a = k_a(\delta_{as} - \delta_a), \\ &\qquad |\delta_{es}| \le 10^\circ, \quad |\delta_{rs}| \le 10^\circ, \quad |\delta_{as}| \le 10^\circ. \end{aligned} \tag{21}$$

Numerical data used in systems (19), (20), and (21) corresponds to Tupolev Tu-154. Corresponding constants and the formulas for aerodynamic coefficients are taken from [23].

It is required to construct a feedback control such that at the instant of passing the runway threshold, deviations of Δz_g and ΔV_{zg} belong to a terminal set M_L. Respectively, the deviations of Δy_g and ΔV_{yg} belong to some set M_V. These sets are taken from the technical tolerances providing further successful landing after passing the runway threshold.

6.2 Linear Systems for Lateral and Vertical Channels

The original nonlinear system after linearization near the nominal motion along the descent glide path is disjoined into two subsystems of lateral motion (LM) and vertical motion (VM).

For numerical simulations, the glide path inclination is taken equal to -2.66°. The nominal aerial velocity is 72.2 m/s. The nominal parameters of the wind are: $W_{xg0} = -5$ m/s, $W_{yg0} = W_{zg0} = 0$.

The linear LM system is

$$\dot{\mathbf{x}} = A\mathbf{x} + Bu + Cw,$$

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.0769 & -5.5553 & 0 & 9.2719 & 0 & -1.4853 & 0 \\ 0 & 0 & 0 & 1.0013 & 0 & 0 & 0 & 0 \\ 0 & -0.0129 & -0.9339 & -0.2588 & -0.0883 & -0.0303 & -0.2456 & -0.0460 \\ 0 & 0 & 0 & -0.0514 & 0 & 1 & 0 & 0 \\ 0 & -0.0331 & -2.3865 & -0.9534 & -0.2256 & -1.4592 & -0.2327 & -0.6894 \\ 0 & 0 & 0 & 0 & 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -4 \end{bmatrix}$$

$$B = \begin{bmatrix} 0\,0\,0\,0\,0\,0\,4\,0 \\ 0\,0\,0\,0\,0\,0\,0\,4 \end{bmatrix}',$$

$$C = \left[0,\ 0.0769,\ 0,\ 0.0129,\ 0,\ 0.0331,\ 0,\ 0\right]'.$$

Here, $\mathbf{x}_1 = \Delta z_g$ is a deviation in z_g from the nominal motion; $\mathbf{x}_2 = \Delta V_{zg}$ is a deviation in V_{zg}; $\mathbf{x}_3 = \Delta\psi$, $\mathbf{x}_5 = \Delta\gamma$ are deviations of the yaw and bank angles. The coordinates $\mathbf{x}_7$, $\mathbf{x}_8$ are deviations of the rudder and ailerons. The controls of the first player are the scheduled deviations $u_1 = \Delta\delta_{rs}$ of the rudder and $u_2 = \Delta\delta_{as}$ of the ailerons. Constraints are

$$|\Delta\delta_{rs}| \le 10^\circ, \qquad |\Delta\delta_{as}| \le 10^\circ.$$

The variable $w = \Delta W_{zg}$ denotes the lateral component of the wind velocity. In real life, the wind changes inertially. To describe this, we use the following wind dynamics:

$$\Delta\dot{W}_{zg} = 0.5(v_{zg} - \Delta W_{zg}).$$

The terminal set in the plane $\mathbf{x}_1 = \Delta z_g$, $\mathbf{x}_2 = \Delta V_{zg}$, which is a simplification of the set M_L, is taken as the hexagon with vertices $(-6, 0)$, $(-6, 1.5)$, $(0, 1.5)$, $(6, 0)$, $(6, -1.5)$, $(0, -1.5)$.

The linear VM system is

$$\dot{\mathbf{x}} = A\mathbf{x} + Bu + Cw,$$

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.0501 & 0 & -0.0973 & -2.6422 & 0 & 0.0628 & 0.9971 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0.2409 & 0 & -0.6387 & 45.2782 & 0 & 1.4479 & 0.0813 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0.0003 & 0 & 0.0069 & -0.5008 & -0.5263 & -0.3830 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2.7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 \end{bmatrix}',$$

$$C = \begin{bmatrix} 0 & 0.0501 & 0 & -0.2409 & 0 & -0.0003 & 0 & 0 \\ 0 & 0.0973 & 0 & 0.6387 & 0 & -0.0069 & 0 & 0 \end{bmatrix}'.$$

Here, $\mathbf{x}_1 = \Delta x_g$ and $\mathbf{x}_3 = \Delta y_g$ are deviations in x_g and y_g from the nominal motion; $\mathbf{x}_2 = \Delta V_{xg}$ and $\mathbf{x}_4 = \Delta V_{yg}$ are deviations in V_{xg} and V_{yg}; $\mathbf{x}_5 = \Delta\vartheta$ is a deviation of the pitch angle. The coordinates $\mathbf{x}_7$, $\mathbf{x}_8$ are deviations of the elevator and thrust force. The controls of the first player are the scheduled deviations $u_1 = \Delta\delta_{ps}$ of the thrust force and $u_2 = \Delta\delta_{es}$ of the elevator. Constraints are

$$|\Delta\delta_{ps}| \le 27^\circ, \qquad |\Delta\delta_{es}| \le 10^\circ.$$

The variables $w_1 = \Delta W_{xg}$ and $w_2 = \Delta W_{yg}$ denote the longitudinal and vertical components of the wind velocity. Taking into account inertia of the wind, we write the law of change of these variables as follows:

$$\Delta \dot{W}_{xg} = 0.5(v_{xg} - \Delta W_{yg}), \qquad \Delta \dot{W}_{yg} = 0.5(v_{yg} - \Delta W_{yg}).$$

The terminal set in the plane $\mathbf{x}_3 = \Delta y_g$, $\mathbf{x}_4 = \Delta V_{yg}$, which is a simplification of the set M_V, is taken as the hexagon with vertices $(-3, 0)$, $(-3, 1)$, $(0, 1)$, $(3, 0)$, $(3, -1)$, $(0, -1)$.

To construct the adaptive control for these linear systems, one has to define the constraint $Q_{\max}$ for the second player's control. For the lateral channel, the set is taken as $|v_{zg}| \leq 10$ m/s. For the vertical channel, the set $Q_{\max}$ is equal to $|v_{xg}| \leq 6$ m/s, $|v_{yg}| \leq 4$ m/s.

6.3 Wind Microburst Model

We shall suppose that during landing, the aircraft crosses a wind microburst zone.

The wind microburst is a natural phenomenon appearing when an air down-flow strikes the ground and fluxes horizontally, with creating the whirls [7]. When reaching the microburst zone, at first, the aircraft meets headwind, which changes quite fast (within ten seconds) to a down-flow and further to a tail wind. Such a change is a very complex phenomenon from the point of view of the aircraft aerodynamics. The headwind increases the aircraft aerial velocity and, therefore, the lifting force; vice versa, the tail wind and the down-flow decrease it. So, such a change of the wind direction leads to a quick decrease of the lifting force.

For the simulation, we use the wind microburst model from the paper [13]. In the space, we determine a torus (Fig. 3). Out of the torus, the turbulence is created, but inside it the proportional decrease of the wind velocity takes place as closing

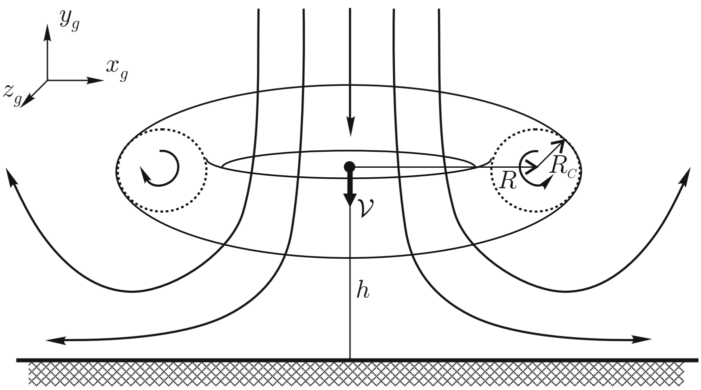


Fig. 3 Wind microburst model

to the torus main ring. The microburst parameters are the following: $\mathscr{V}$ is the wind velocity at the central point (this value is not maximal, the maximal value can be up to twice larger in the torus boundary layers); h is the height of the torus main ring over the ground; R is the radius of the central line of the torus ring; $R_C = 0.8h$ is the torus main ring radius; $(\tilde{x}_0, \tilde{z}_0)$ is the torus center position in the ground plane.

The microburst acts additionally to a constant nominal wind with the components W_{xg0}, W_{yg0}, and W_{zg0}.

6.4 Simulation Results

To simulate the landing process (Figs. 4–6), the original nonlinear system is used. When constructing control during the process, we compute the prognosis backward time until passing the runway threshold at each instant of the discrete scheme.

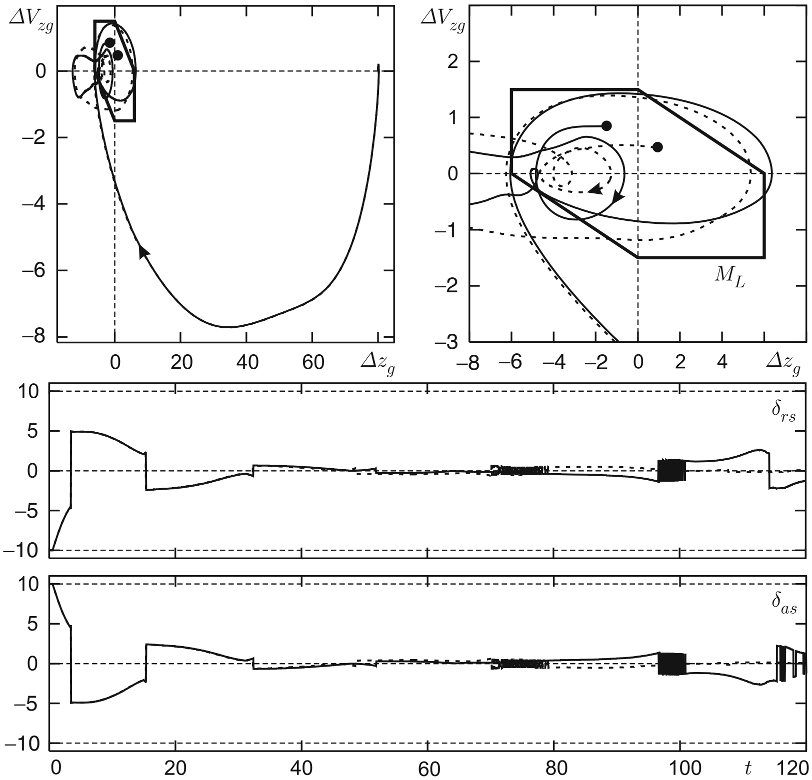


Fig. 4 *First row*: at the left, trajectories for the lateral channel in the plane $\Delta z_g \times \Delta V_{zg}$ are shown; at the right, there is an enlarged fragment of the left picture. *In the bottom*: graphs of command deviations of the rudder δ_{rs} (deg) and ailerons δ_{as} (deg)

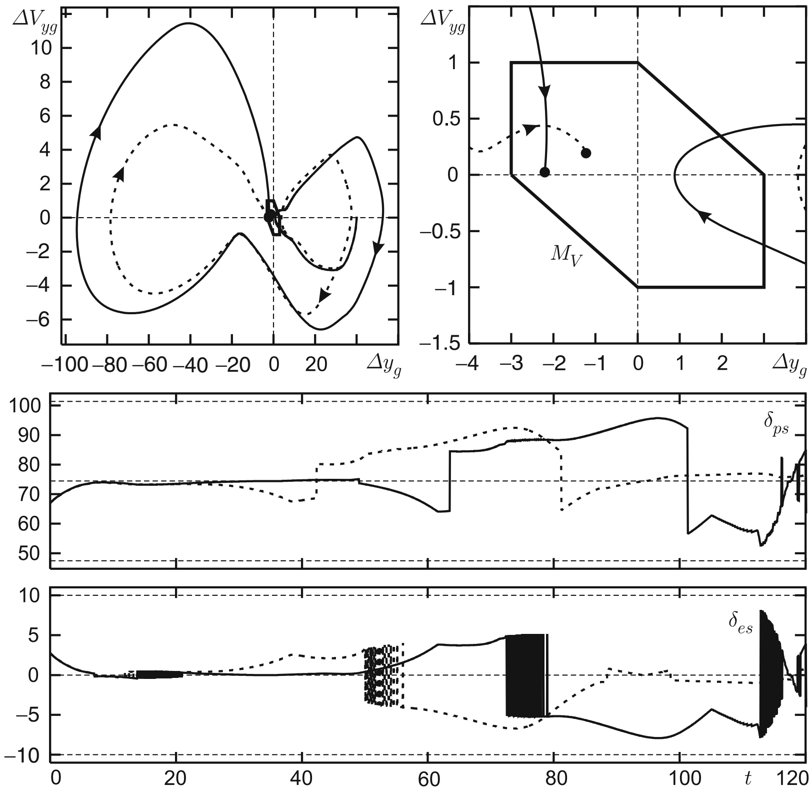


Fig. 5 *First row*: at the left, trajectories for the vertical channel in the plane $\Delta y_g \times \Delta V_{yg}$ are shown; at the right, there is an enlarged fragment of the left picture. *In the bottom*: graphs of the command deviations of the thrust force control δ_{ps} (deg) and elevator δ_{es} (deg)

Then, the controls δ_{rs}, δ_{as} in the lateral channel and δ_{ps}, δ_{es} in the vertical one are generated for some small time interval by means of adaptive control laws obtained for the linear systems.

The results of the simulation are presented for two variants of the wind microburst. Parameters of the first variant are: $\mathscr{V} = 10$ m/s, $h = 600$ m, $R = 1{,}200$ m, the longitudinal distance of the central point from the runway threshold is 4,000 m, the lateral distance from the central point to the glide path is 500 m. In the second variant, the velocity $\mathscr{V}$ at the central point is stronger, 15 m/s, it is located closer to the runway, and the longitudinal distance is only 2,500 m.

Let the initial position of the aircraft be 8,000 m from the runway threshold along the axis x_g and deviate from the nominal position on the glide path for 40 m up and 80 m sideward.

Each of Figs. 4–6 contains results for both variants. The solid lines denote the graphs corresponding to the second variant (with a stronger microburst), the dashed lines correspond to the variant of the weaker wind. The terminal sets are shown by

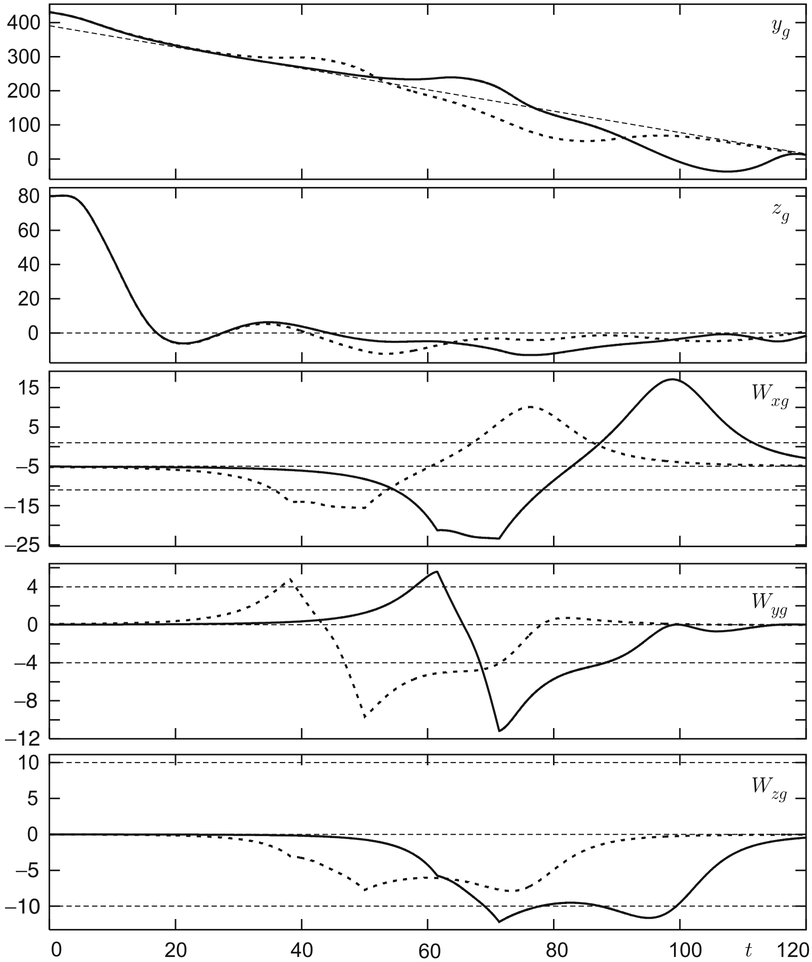


Fig. 6 Graphs of altitude y_g, m; lateral deviation z_g, m; components W_{xg}, W_{yg}, and W_{zg} of wind velocity, m/s

thick solid lines. In the graphs of useful controls and disturbances, the thin dashed lines denote the nominal and maximal (expected extremal for the disturbances) values. In the graphs of phase coordinates, the dashed lines show the nominal values.

In Figs. 4 and 5, one can see the results of the simulation for the lateral and vertical channels. In Fig. 6, the graphs of the height y_g and lateral deviation z_g are given. Also, there are graphs of the components W_{xg}, W_{yg}, and W_{zg} of the wind disturbance.

For the first microburst variant, it is seen that despite the wind velocity for some time intervals exceeded the expected value, the control successfully parried the disturbance, and trajectories reached their terminal sets. Formally, the condition of passing the tolerance at the instant of reaching the runway threshold is fulfilled for the second variant of the microburst too. But the altitude graph shows dissatisfactory result. Approximately in 20 s before passing the runway threshold, the aircraft collides with the ground. The explanation can be the following: a significant change of the longitudinal and vertical components of the wind velocity happened at quite small altitude. Also, our method cannot take into account the state constraint on the altitude of the aircraft.

Conclusion

The method of adaptive control suggested in this paper is based on approaches of theory of antagonistic differential games and is oriented to problems where a dynamic disturbance is bounded, but the level of its constraint is unknown *a priori*. Specific properties of stable bridges in linear differential games with fixed terminal time allow to create an algorithm, whose ideology is very simple. Its efficiency was checked in the framework of the control problem by an aircraft during landing under wind disturbances.

Acknowledgements This work was supported by Program "Mathematical control theory" of the Presidium of the RAS, the Russian Foundation for Basic Research grants Nos. 09-01-00436 and 10-01-96006).

References

1. Bardi, M., Capuzzo-Dolcetta, I.: Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations. Birkhauser, Boston (1997)
2. Botkin, N.D., Kein, V.M., Patsko, V.S.: The model problem of controlling the lateral motion of an aircraft during landing. J. Appl. Math. Mech. **48**(4), 395–400 (1984)
3. Botkin, N.D., Ryazantseva, E.A.: Algorithms for constructing the solvability set in a linear differential game of high dimension. Trudy Instituta Matematiki i Mekhaniki, Ekaterinburg, **2**, 128–134 (1992) (in Russian)
4. Bulirsch, R., Montrone, F., Pesch, H.J.: Abort landing in the presence of windshear as a minimax optimal control problem; part 1: Necessary conditions; part 2: Multiple shooting and homotopy. J. Opt. Th. Appl. **70**(1) 1–23; **70**(2), 223–254 (1991)
5. Cardaliaguet, P., Quincampoix, M., Saint-Pierre, P.: Set-valued numerical analysis for optimal control and differential games. In: Bardi, M., Raghavan, T.E., Parthasarathy, T. (eds.) Annals of the ISDG, Vol. 4, pp. 177–247 (1999)
6. Chen, Y.H., Pandey, S.: Robust control strategy for take-off performance in windshear. Opt. Control Appl. Meth. **10**(1), 65–79 (1989)
7. Dole, C.E.: Flight Theory and Aerodynamics. Wiley, New York (1981)
8. Falcone, M.: Numerical methods for differential games based on partial differential equations. Int. Game Theory Rev. **8**(2), 231–272 (2006)
9. Ganebny, S.A., Kumkov, S.S., Patsko, V.S.: Control design in problems with an unknown level of dynamic disturbance. J. Appl. Math. Mech. **70**(5), 680–695 (2006)
10. Ganebny, S.A., Kumkov, S.S., Patsko, V.S., Pyatko, S.G.: Constructing robust control in differential games: application to aircraft control during landing. In: Jørgensen, S., Quincampoix,

M., and Vincent, T.L. (eds.) Annals of the ISDG, Vol. 9: Adv. in Dynamic Games and Appl., pp. 69–92. Birkhauser, Boston (2007)
11. Grigorenko, N.L., Kiselev, Yu.N., Lagunova, N.V., Silin, D.B., Trin'ko, N.G.: Methods for solving differential games. Math. Modelling, Moscow State University, Moscow, 296–316 (1993) (in Russian)
12. Isaacs, R.: Differential Games. Wiley, New York (1965)
13. Ivan, M.: A ring-vortex downburst model for real-time flight simulation of severe windshear. AIAA Flight Simulation Technologies Conf., St. Louis, Miss., 57–61 (1985)
14. Ivanov, A.G., Ushakov, A.V.: Usage of maximal stable bridges for constructing adaptive control in three-dimensional linear systems. Problems of Theor. and Appl. Math., Proc. of the 39th Regional Youth Conf., Ekaterinburg, 250–254 (2007) (in Russian)
15. Kein, V.M.: Optimization of Control Systems with Minimax Criterion. Nauka, Moscow (1985) (in Russian)
16. Kein, V.M., Parikov, A.N., Smurov, M.Yu.: On means of optimal control by the extremal aiming method. J. Appl. Math. Mech. **44**(3), 306–310 (1980)
17. Krasovskii, N.N, Subbotin, A.I.: Positional Differential Games. Nauka, Moscow (1974) (in Russian)
18. Krasovskii, N.N., Subbotin, A.I.: Game–Theoretical Control Problems. Springer, New York (1988)
19. Leitmann, G., Pandey, S.: Aircraft control for flight in an uncertain environment: Take-off in windshear. J. Opt. Th. Appl. **70**(1), 25–55 (1991)
20. Miele, A., Wang, T., Melvin, W.W.: Optimal take-off trajectories in the presence of windshear. J. Opt. Th. Appl. **49**(1), 1–45 (1986)
21. Miele, A., Wang, T., Tzeng, C.Y., Melvin, W.W.: Optimal abort landing trajectories in the presence of windshear. J. Opt. Th. Appl. **55**(2), 165–202 (1987)
22. Miele, A., Wang, T., Wang, H., Melvin, W.W.: Optimal penetration landing trajectories in the presence of windshear. J. Opt. Th. Appl. **57**(1), 1–40 (1988)
23. Patsko, V.S., Botkin, N.D., Kein, V.M., Turova, V.L., Zarkh, M.A.: Control of an aircraft landing in windshear. J. Opt. Th. Appl. **83**(2), 237–267 (1994)
24. 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. Sb. Math. **192**(10), 1515–1542 (2001)
25. Suebe, N., Moitie, R., Leitmann, G.: Aircraft taking-off in windshear: a viability approach. Set-Valued Anal. **8**, 163–180 (2000)
26. Subbotin, A.I., Patsko, V.S. (eds.): Algorithms and Programs for Solving Linear Differential Games. Institute of Mathematics and Mechanics, Sverdlovsk (1984) (in Russian)
27. Taras'ev, A.M., Ushakov, V.N., Khripunov, A.P.: Construction of positional absorption sets in game-theoretic control problems. Tr. Inst. Matematiki i Mekhaniki, Ekaterinburg, **1**, 160–177 (1992) (in Russian)
28. Ushakov, V.N.: Construction of solutions in differential games of pursuit-evasion. Diff. Inclus. and Opt. Control, Lecture Notes in Nonlinear Analysis, Vol. 2, Nicholas Copernicus University, Torun, 269–281 (1998)
29. Zarkh, M.A., Ivanov, A.G.: Construction of the game value in a linear differential game with fixed terminal instant. Tr. Inst. Matematiki i Mehkaniki, Ekaterinburg, **2**, 140–155 (1992) (in Russian)
30. Zarkh, M.A., Patsko, V.S.: Numerical solution of a third-order directed game. Soviet J. Comput. Syst. **26**(4), 92–99 (1988)

A Generalization of the Multi-Stage Search Allocation Game

Ryusuke Hohzaki

Abstract This chapter deals with a multistage two-person zero-sum game called *multistage search allocation game* (MSSAG), in which a searcher and a target participate. The searcher distributes his searching resource in a search space to detect the target and the target moves under an energy constraint to evade the searcher. At the beginning of each stage, the searcher is informed of the target's position and of his moving energy, and the target knows the amount of the searcher's resource used so far. The game ends when the searcher detects the target. The payoff of the game is the probability of detecting the target during the search. There have been few search games on the MSSAG. We started the discussion on the MSSAG in a previous paper, where we assumed an exponential function for the detection probability of the target. In this chapter, we introduce a general function for the detection probability, aiming at more applicability. We first formulate the problem as a dynamic program. Then, by convex analysis, we propose a general method to solve the problem, and elucidate some general properties of optimal solutions.

1 Introduction

Search theory dates back to the scientific analysis of the antisubmarine operations in World War II. In the early history of the search theory, almost all researchers were interested in the one-sided search problem of optimizing the searcher's plan (see for instance [17] for military applications and anti-submarine warfare (ASW) in WWII and [24] for a generalization to optimal search). The search problem was then extended to search games, where not only the searcher's strategies, but also the target's ones are optimized.

R. Hohzaki (✉)
Department of Computer Science, National Defense Academy, Hashirimizu, Yokosuka, 239-8686, Japan
e-mail: hozaki@cc.nda.ac.jp

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_10,

From modeling point of view, a simpler problem is the single-stage game (SSG) with a stationary target, as in [4] and [6]. Many papers dealing with the moving strategy of the searcher in the SSG were published, such as [1] and [5]. The so-called Ruckle problem ([22] and [7]) is a special case of the SSG.

In multistage games (MSGs), players take advantage of some information about their enemies obtained at each stage. In those games, the searcher's strategy is to choose his positions for moving or trapping, such as in [18, 25] and [19].

If the searcher's mobility is greatly superior to the target's one, and if the searcher can move wherever he likes while the target takes a step forward, the moving strategy is not important for the searcher. In that case, the distribution of the searching resource could be the searcher's strategy. Papers considering this strategy mostly deal with single-stage games, see for instance [11–13, 15, 16, 20, 26], and [14]. In [2], another type of SSG is considered, where both players distribute their resource.

We refer to the search game played by the distributor of searching resource and the moving target as search allocation game (SAG), as in [8]. Few papers consider the SAG (see [9]) and the multistage search allocation game (MSSAG). A discussion on the MSSAG can be found in [10], where we introduced an exponential function for the detection probability of the target in a discrete search space. The purpose of this paper is to propose a general methodology to solve the MSSAG with a general function for the detection probability as payoff, and to find the value and equilibrium points of the MSSAG.

In the next section, we describe some assumptions for an MSSAG and formulate it as a dynamic programming problem. In Sect. 3, we find an equilibrium solution for a single-stage game as a preliminary result and apply it to our MSSAG to derive a system of equations for the value of the game and an optimal solution. In Sect. 4, we elucidate some properties of the value of the game and optimal strategies of players. We analyze optimal strategies using some numerical examples in Sect. 5 and conclude our discussion with some remarks in Sect. 6.

2 Description of Assumptions and Formulation

We consider a multistage stochastic search game, in which a searcher and a target participate as two players.

A1. A multistage game is defined in a discrete time space and the stage number is represented as the residual number of time periods until the end of the game $n = 0$.

A2. A search space consists of discrete cells $K = \{1, \ldots, K\}$.

A3. A searcher and a target play the game. The searcher tries to detect the target by distributing his searching resources into cells under a limited searching cost constraint. On the other hand, the target moves in the search space to avoid detection by the searcher. We make the following assumptions about the players' strategies, information sets, payoff, and stage transitions.

(1) At the beginning of each stage n, the searcher is informed of the current cell of the target, say k, and of his residual energy. The target knows how much resource the searcher has distributed in the past.
(2) The target chooses a cell to move to from cell k probabilistically. The movement is constrained geographically, restricting the feasible moves to a set of cells $N(k) \subseteq K$. The target spends an amount of energy $\mu(i, j)$ to move from cell i to j. We assume that $k \in N(k)$, $\mu(i, j) > 0$ for $i \neq j$ and $\mu(k, k) = 0$. The target then inevitably stays forever in a cell when his energy is exhausted. His initial energy is denoted by e_0.
(3) The searcher makes a distribution plan of his searching resource in the search space. It costs $c_i > 0$ for the searcher to distribute one unit of searching resource in cell i. His initial budget is denoted Φ_0.
(4) If the target is at cell i, the searcher can detect the target with probability $1 - q_i(x)$ if he/she has distributed x units of resource to cell i. That is, $q_i(x)$ is the nondetection probability function of cell i for resource x. This function has the following properties:
(Q1) $q_i(x)$ is twice continuously differentiable with respect to x It is a monotone decreasing and positive-valued function, with $q_i(0) = 1$ and a convergence point $q_i^\infty = \lim_{x\to\infty} q_i(x) \geq 0$.
(Q2) $q_i(x)$ is a strictly convex function of x.
Upon detection of the target, the searcher gains one and the target loses the same, and then the game terminates.
(5) If no detection occurs at stage n, the game steps forward to the next stage $n-1$.

A4. The game ends either when detection occurs or when it reaches the last stage $n = 0$. The searcher acts as a maximizer and the target as a minimizer.

Our problem is a multistage stochastic game with the detection probability of the target as payoff. The cells that the target having residual energy e can reach from cell k is given by

$$N(k, e) \equiv \{i \in N(k) \mid \mu(k, i) \leq e\}.$$

We denote a strategy of the target, having energy e and staying at cell k, by $\{p(k, i), i \in \boldsymbol{K}\}$, where $p(k, i) \geq 0$ is the probability that he chooses cell i as his next cell. It is easily seen the feasible region of the target strategy is given by

$$P_k(e) = \left\{ \{p(k,i), i \in \boldsymbol{K}\} \;\middle|\; p(k,i) \geq 0, i \in N(k,e), \right.$$
$$\left. p(k,i) = 0, i \in \boldsymbol{K} - N(k,e), \sum_{i \in N(k,e)} p(k,i) = 1 \right\}.$$

On the other hand, we denote a strategy of the searcher by $\{\varphi(i), i \in \boldsymbol{K}\}$, where $\varphi(i) \geq 0$ is the amount of the resource distributed in cell i. If he/she has a residual budget Φ, which is expendable for his distribution cost, he/she faces the constraint $\sum_i c_i\varphi(i) \le \Phi$. Since he can anticipate that distributing his/her searching resource into cells not belonging to $N(k,e)$ is of no use, a feasible region of his strategy is

$$\Psi(\Phi) = \left\{ \{\varphi(i), i \in \boldsymbol{K}\} \,\middle|\, \varphi(i) \ge 0, i \in N(k,e), \right.$$
$$\left. \varphi(i) = 0, i \in \boldsymbol{K} - N(k,e), \sum_{i \in N(k,e)} c_i\varphi(i) \le \Phi \right\}.$$

Now consider a target with energy e at cell k and a searcher with residual budget Φ at the beginning of stage n. We consider the game played from that point, which is assumed to have a value $v_n(k,e,\Phi)$. The existence of the value will be proved later on. If the target moves to cell i, the searcher can detect the target with probability $1 - q_i(\varphi(i))$ by distributing an amount $\varphi(i)$ of searching resource. Unless detection occurs, the game continues and moves to the next stage. The transition state is (current cell, residual energy) $= (i, e - \mu(k,i))$ for the target and residual budget $\Phi' = \Phi - \sum_{j \in N(k,e)} c_j\varphi(j)$ for the searcher. Therefore, the game in a state (n,k,e,Φ) moves to a state $(n-1, i, e - \mu(k,i), \Phi')$, $i \in N(k,e)$ with probability $p(k,i)q_i(\varphi(i)) \ge 0$ or terminates with probability $\sum_{i \in N(k,e)} p(k,i)(1 - q_i(\varphi(i))) < 1$. It is well known that a stochastic game ends with certainty and has a unique value if it has some positive probability of termination at each stage, as shown in [21] or [23]. Recall that the game also terminates at the last stage if no detection occurred before. The value of the multistage game is calculated recursively and, assuming it exists, the value of the game $v_n(k,e,\Phi)$ satisfies the following recursive equation

$$\begin{aligned} v_n(k,e,\Phi) &= \max_{\varphi \in \Psi(\Phi)} \min_{p \in P_k(e)} \sum_{i \in N(k,e)} p(k,i) \left\{ (1 - q_i(\varphi(i))) \right. \\ &\quad \left. + q_i(\varphi(i)) v_{n-1}(i, e - \mu(k,i), \Phi - \sum_{j \in N(k,e)} c_j\varphi(j)) \right\} \\ &= 1 - \min_{\varphi \in \Psi(\Phi)} \max_{p \in P_k(e)} \sum_{i \in N(k,e)} p(k,i) \\ &\quad \times \left\{ 1 - v_{n-1}(i, e - \mu(k,i), \Phi - \textstyle\sum_j c_j\varphi(j)) \right\} q_i(\varphi(i)). \end{aligned} \tag{1}$$

For $n = 0$, the value is obviously zero.

$$v_0(k,e,\Phi) = 0.$$

When $\Phi = 0$, the game reaches the last stage without any detection of the target, and the value of the game is again zero from Assumption (Q1).

$$v_n(k, e, 0) = 0.$$

When $e = 0$, the target always stays at cell k because of $N(k, e) = \{k\}$ from Assumption A3-(2) and the searcher searches there for the remainder of the game. As a result, (1) becomes:

$$v_n(k, 0, \Phi) = 1 - \min_{0 \le \varphi(k) \le \Phi/c_k} \{1 - v_{n-1}(k, 0, \Phi - c_k\varphi(k))\} q_k(\varphi(k))$$
$$\textit{Initial value}: \ v_1(k, 0, \Phi) = 1 - q_k(\Phi/c_k).$$

Even though $q_k(\cdot)$ is convex, this equation is difficult to solve. However under the following condition, we can obtain a closed-form of solution.

(Q2') $\log\ q_i(x)$ is convex in x.

In this case, we can derive an optimal solution $\varphi(k) = \Phi/c_k/n$ at every stage $k = n, n-1, \cdots, 1$ and the value of the game

$$v_n(k, 0, \Phi) = 1 - \left\{ q_k \left(\frac{\Phi}{n c_k} \right) \right\}^n, \tag{2}$$

is obtained by mathematical induction.

Substitute $h_n(k, e, \Phi) \equiv 1 - v_n(k, e, \Phi)$ to yield an alternative formulation of (1)

$$\begin{aligned} &h_n(k, e, \Phi) \\ &\quad = \min_{\varphi \in \Psi(\Phi)} \max_{p \in P_k(e)} \sum_{i \in N(k,e)} p(k, i) h_{n-1}(i, e - \mu(k, i), \Phi - \textstyle\sum_j c_j \varphi(j)) q_i(\varphi(i)), \end{aligned} \tag{3}$$

which is easier to deal with. The initial conditions are

$$h_0(k, e, \Phi) = 1, \ \ h_n(k, e, 0) = 1.$$

The optimization problem (3) is a formulation of the zero-sum game under the criterion of nondetection probability, where the target is assumed to be rewarded with one only if he avoids the searcher until the last stage. Under Assumption (Q2'), we have the initial condition

$$h_n(k, 0, \Phi) = \left\{ q_k \left(\frac{\Phi}{n c_k} \right) \right\}^n,$$

when $e = 0$.

We now discuss the minimax optimization problem (3). First, we divide the minimization $\min_{\varphi \in \Psi(\Phi)}$ into two parts. One is a minimization given that budget Φ' is used at stage n and the other is a minimization with respect to Φ'. We have a double-layered minimization problem, as follows.

$$
\begin{aligned}
& h_n(k, e, \Phi) \\
& = \min_{0 \le \Phi' \le \Phi} \min_{\varphi \in \Psi_E(\Phi')} \max_{p \in P_k(e)} \sum_{i \in N(k,e)} p(k,i) h_{n-1}(i, e - \mu(k,i), \Phi - \Phi') q_i(\varphi(i)),
\end{aligned}
\tag{4}
$$

where $\Psi_E(\Phi')$ is defined as

$$
\Psi_E(\Phi') = \left\{ \{\varphi(i), i \in \boldsymbol{K}\} \middle| \varphi(i) \ge 0, i \in N(k,e), \varphi(i) = 0, i \in \boldsymbol{K} - N(k,e), \right. \\
\left. \sum_{i \in N(k,e)} c_i \varphi(i) = \Phi' \right\}.
$$

3 Equilibrium for the Game

Here, given Φ and Φ', we discuss the second and the third optimization of problem (4), which begins with $\min_{\varphi \in \Psi_E(\Phi')}$. For simplicity, we denote β_i as a substitute for $h_{n-1}(i, e - \mu(k,i), \Phi - \Phi')$, p_i for $p(k,i)$, φ_i for $\varphi(i)$ and A for $N(k,e)$. We focus on the following minimax optimization to obtain its optimal solution φ_i, p_i just at stage n.

$$
\min_{\{\varphi_i\}} \max_{\{p_i\}} \sum_{i \in A} p_i \beta_i q_i(\varphi_i). \tag{5}
$$

A minimax value of the problem coincides with a maxmin value because the objective function is linear in p_i and convex in φ_i, as shown in [12]. The problem is a kind of the so-called hide-and-search game, where a stationary target chooses once a cell i with probability p_i to hide and a searcher distributes his searching resource of budget Φ' to detect the target. This problem has been studied for specific forms of the nondetection function $q_k(\cdot)$, see [3] and [8] . Here, we are going to solve the general problem.

Recall that feasible regions are $\varphi_i \ge 0,\ i \in A,\ \sum_{i \in A} c_i \varphi_i = \Phi'$ for φ_i and $p_i \ge 0,\ i \in A,\ \sum_{i \in A} p_i = 1$ for p_i.

Using the transformation

$$
\max_{\{p_i\}} \sum_{i \in A} p_i \beta_i q_i(\varphi_i) = \max_i \beta_i q_i(\varphi_i),
$$

problem (5) can be formulated as a convex programming problem as follows.

$$\begin{aligned} (P_0) \quad & \min\ \rho \\ s.t.\ \ & \beta_i q_i(\varphi_i) \le \rho,\ i \in A \\ & \textstyle\sum_{i \in A} c_i \varphi_i = \Phi' \\ & \varphi_i \ge 0,\ i \in A. \end{aligned}$$

Problem (P_0) can be solved analytically by convex analysis or nonlinear optimization methods. We can easily derive the following system Karush–Kuhn–Tucker conditions for the problem:

$$\textstyle\sum_{\{i|\beta_i \ge \rho\}} c_i q_i^{-1}(\rho/\beta_i) = \Phi' \tag{6}$$

$$\varphi_i = \left\{q_i^{-1}(\rho/\beta_i)\ \ if\ \beta_i \ge \rho,\ \ 0\ \ if\ \beta_i < \rho\right\} \tag{7}$$

$$p_i = \begin{cases} c_i/\beta_i/q_i'(\varphi_i) \Big/ \sum_{\{j,\beta_j \ge \rho\}} c_j/\beta_j/q_j'(\varphi_j)\ , & i \in \{j \in A \mid \beta_j \ge \rho\}, \\ 0 & , i \notin \{j \in A \mid \beta_j \ge \rho\}. \end{cases} \tag{8}$$

To derive an optimal solution for problem (P_0), we first obtain a root ρ for (6), which is an optimal value for the single-stage game (5). By substituting ρ into (7), we can derive an optimal strategy $\{\varphi_i\}$. Equation (8) leads us to $\{p_i\}$.

Now let us derive the value of the game $h_1(k, e, \Phi)$ for $n = 1$ by formula (6)–(8). Noting that $h_0(k, e, \Phi) = 1$, an original problem (3) is just a single-stage game (5) with $\beta_i = 1$. If $\rho > 1$, $p(k, i) = 0$ for every i from (8), which is a contradiction. Therefore, we have $\rho \le 1$ and then $\{j \in A|\beta_j \ge \rho\} = N(k, e)$. The value $h_1(k, e, \Phi)$ is given by a root $\rho^* \in (\max_i q_i^\infty\ , 1\,]$ of equation

$$\sum_{i \in N(k,e)} c_i q_i^{-1}(\rho) = \Phi, \tag{9}$$

and the following formulas give optimal solutions:

$$\varphi^*(i) = q_i^{-1}(\rho^*), i \in N(k, e) \tag{10}$$

$$p^*(k, i) = c_i \Big/ q_i'(\varphi^*(i)) \Big/ \textstyle\sum_{j \in N(k,e)} c_j \Big/ q_j'(\varphi^*(j))\ \ , i \in N(k, e). \tag{11}$$

The above results for the single-stage game can be applied to the part beginning with $\min_{\varphi \in \Psi_E(\Phi')} \max_{p \in P_k(e)}$ of problem (4) and an optimal value $h_n(k, e, \Phi)$ at stage n is obtained in general as follows.

First, find a root $\rho = \rho_n(k, e, \Phi, \Phi')$ of equation

$$\sum_{i \in A_n(k,e,\Phi,\Phi';\rho)} c_i q_i^{-1}(\rho/h_{n-1}(i, e - \mu(k, i), \Phi - \Phi')) = \Phi', \tag{12}$$

where $A_n(k, e, \Phi, \Phi'; \rho) \equiv \{i \in N(k, e) \mid \rho \le h_{n-1}(i, e - \mu(k, i), \Phi - \Phi')\}$, and optimize the value with respect to Φ' by

$$h_n(k, e, \Phi) = \rho_n(k, e, \Phi, \Phi'^*) = \min_{\{\Phi', 0 \le \Phi' \le \Phi\}} \rho_n(k, e, \Phi, \Phi') \tag{13}$$

to obtain $h_n(k, e, \Phi)$. Using notation $\rho^* = h_n(k, e, \Phi)$, optimal solutions φ^*, p^* are given by

$$\varphi^*(i) = \begin{cases} q_i^{-1}(\rho^*/h_{n-1}(i, e - \mu(k, i), \Phi - \Phi'^*)), & i \in A_n(k, e, \Phi, \Phi'^*; \rho^*) \\ 0, & i \notin A_n(k, e, \Phi, \Phi'^*; \rho^*) \end{cases} \tag{14}$$

$$p^*(k, i) = \begin{cases} \dfrac{c_i / h_{n-1}(i, e - \mu(k, i), \Phi - \Phi'^*)/q_i'(\varphi^*(i))}{\displaystyle\sum_{j \in A_n(k, e, \Phi, \Phi'^*; \rho^*)} c_j / h_{n-1}(j, e - \mu(k, j), \Phi - \Phi'^*)/q_j'(\varphi^*(j))}, \\ \qquad\qquad i \in A_n(k, e, \Phi, \Phi'^*; \rho^*) \\ 0, \ i \notin A_n(k, e, \Phi, \Phi'^*; \rho^*). \end{cases} \tag{15}$$

We now discuss the characteristics of optimal strategies of the players. Provided that the searcher fails to detect the target at the current stage n, the resultant expected payoff of the nondetection probability may depend on the cell the target chooses to move to. If so, the target would be quite willing to go to the cell with higher expected payoff. That is why the searcher equalizes the expected payoff. A value $-q_i'(\varphi^*(i))h_{n-1}(i, e - \mu(k, i), \Phi - \Phi'^*)$ is a marginal expected payoff after stage n, which the increase of unit searching resource at the stage causes for the detection probability of the target. Therefore, strategy (15) indicates that the target should choose his next cell with probability proportional to the marginal cost, which is the expense needed to earn the marginal expected payoff. The strategy is based on the target's rational guess that his enemy would avoid the search in cells with high marginal costs.

Now we itemize the procedure to solve our multistage search game and obtain an equilibrium point. Assume that the game is at stage T, when a target with energy e_0 is at cell $s \in$K and a searcher has budget Φ_0 in hand. Two players are now going to play a multistage search game with nondetection probability of the target as payoff. We can estimate the value of the game, $h_T(s, e_0, \Phi_0)$, and the optimal strategies of the players at each stage by the following procedure. It is assumed that initial energy e_0 and moving-energy-consumption function $\mu(i, j)$ are integer-valued. We denote a set of energy by $\boldsymbol{E} = \{0, 1, \cdots, e_0\}$.

Step 1. Set $n = 1$. For every $k \in \boldsymbol{K}$, $e \in \boldsymbol{E}$ and $0 \le \Phi \le \Phi_0$, we calculate an equilibrium point in the following manner.
First find a root ρ^* of (9), which is the value of the game $h_1(k, e, \Phi)$ at the initial stage $n = 1$. When we substitute ρ^* into expression (10), we obtain an optimal searcher's strategy $\varphi^*(i)$. The strategy is used to get an optimal target's strategy $\{p^*(k, i), i \in \boldsymbol{K}\}$ by (11).

Step 2. Increase n by one, $n = n + 1$. For every $k \in \boldsymbol{K}$, $e \in \boldsymbol{E}$ and $0 \leq \Phi \leq \Phi_0$, we use the following procedure.
Using $\{h_{n-1}(i, e', \Phi'),\ i \in \boldsymbol{K},\ e' \in \boldsymbol{E},\ \Phi' \leq \Phi_0\}$ which have been obtained, find a root $\rho = \rho_n(k, e, \Phi, \Phi')$ of (12). We solve a minimization problem (13) with respect to Φ' to find the value of the game $\rho^* = h_n(k, e, \Phi)$. By substituting ρ^* into (14) and (15), we obtain optimal strategies of the searcher and the target.
Step 3. If $n = T$, stop the procedure. Otherwise, go to (Step 2).

The above is the basic procedure to solve our search game. However, it does not guarantee that the game is always solved. One approach is to divide Φ into some discrete portions to make the optimization problem (13) solvable in the sense of discrete approximation. If, on the other hand, Φ is kept as a continuous variable, we need to be lucky for function $\rho_n(k, e, \Phi, \Phi')$ to be easy to handle with respect to Φ'. Using a function of nondetection probability $q_i(x) = \exp(-\alpha_i x)$ is an example where this happens. As explained in detail in Hohzaki [10], this function allows to obtain analytical forms for the optimal strategies and the values of the game at every stage.

Before closing this section, let us reconfirm the followings. The value $h_1(k, e, \Phi)$ determined by (9) is the value of the game for $n = 1$. However, $h_n(k, e, \Phi)$ determined by (13) for stage n is the minimax value, but we do not know presently whether this value is equal to the value of the game.

4 Properties of Optimal Solution

Here, we elucidate some properties of optimal solutions for the MSSAG. The following theorem states some monotonicity properties of the value of the game $h_n(k, e, \Phi)$.

Theorem 1. *(i) $h_n(k, e, \Phi)$ is positive-valued and monotone nonincreasing in the number of stages n.*

$$h_n(k, e, \Phi) \geq h_{n+1}(k, e, \Phi) > 0.$$

(ii) $h_n(k, e, \Phi)$ is monotone non-decreasing in energy e.
(iii) $h_n(k, e, \Phi)$ is monotone non-increasing in budget Φ.

Proof. (i) The positivity obvious. We will prove the monotone non-increasingness by induction. In the initial case of $n = 1$, monotonicity is verified because $h_0(k, e, \Phi) = 1 \geq \rho = h_1(k, e, \Phi)$ holds for ρ given by (9). Let us assume that $h_1(k, e, \Phi) \geq h_2(k, e, \Phi) \geq \cdots \geq h_{n-1}(k, e, \Phi) \geq h_n(k, e, \Phi)$. With respect to a set of cells $A_n(\cdot)$, we see that $A_n(k, e, \Phi, \Phi'; \rho) \supseteq A_{n+1}(k, e, \Phi, \Phi'; \rho)$ by its definition and then

$$q_i^{-1}\left(\frac{\rho}{h_{n-1}(k, e - \mu(k, i), \Phi - \Phi')}\right) \geq q_i^{-1}\left(\frac{\rho}{h_n(k, e - \mu(k, i), \Phi - \Phi')}\right)$$

from the monotone decreasing-ness of the inverse function $q_i^{-1}(\cdot)$. It leads us to the inequality $\rho_n(k,e,\Phi,\Phi') \geq \rho_{n+1}(k,e,\Phi,\Phi')$ in terms of ρ given by (12). Therefore, property (i) is valid because $h_n(k,e,\Phi) = \min_{0 \leq \Phi' \leq \Phi} \rho_n(k,e,\Phi,\Phi') \geq \min_{0 \leq \Phi' \leq \Phi} \rho_{n+1}(k,e,\Phi,\Phi') = h_{n+1}(k,e,\Phi)$.

(ii) Since $h_0(\cdot) = 1$, (ii) holds for $n = 0$. For $e \leq e'$, we have $N(k,e) \subseteq N(k,e')$, $P_k(e) \subseteq P_k(e')$. Applying the assumption $h_{n-1}(i, e - \mu(k,i), \Phi - \sum_j c_j \varphi(j)) \leq h_{n-1}(i, e' - \mu(k,i), \Phi - \sum_j c_j \varphi(j))$ to formulation (3) easily leads us to $h_n(k,e,\Phi) \leq h_n(k,e',\Phi)$ for stage n.

(iii) $\Psi(\Phi) \subseteq \Psi(\Phi')$ holds for $\Phi \leq \Phi'$. By applying this, inequality to formulation (3), we can prove property (iii). □

Practically, properties (ii) and (iii) of Theorem 1 is self-explanatory from the inherent characteristics of the problem. Property (i) is important from two points of view. First, it indicates that an increase in the number of stages is advantageous for the searcher and disadvantageous for the target. At the beginning of each stage, the target is informed of the residual budget of the searcher. On the other hand, the searcher can know the current cell of the target as well as his residual energy, which helps the searcher focus his searching resource in a comparatively small area and conduct an effective search for the target. If you assume otherwise that the searcher could not obtain information about the target position and he had to consider the possible area of the target expanding rapidly as time elapses, the difficulty of the search would not be comparable with the original case at all. We can say that the information about the target position gives the searcher a lot of advantage. The second point of property (i) of Theorem 1 implies that the value of the game converges to a certain value as the number of stages becomes $n \to \infty$ and the game has a kind of asymptotic stability.

From Theorem 1, we have the following corollary.

Corollary 1. *There is a positive probability that the current cell of the target be chosen as is destination cell at the next stage.*

Proof. First, recall that the current cell k belongs to $N(k,e)$ from Assumption A3-(2). A set of cells that the target possibly chooses as his next cell is $A_n(k,e,\Phi,\Phi'^*;\rho^*)$, as seen from (15). We can verify the validity of the corollary by showing

$$\rho^* = h_n(k,e,\Phi) \leq h_{n-1}(k, e - \mu(k,k), \Phi - \Phi'^*) = h_{n-1}(k,e,\Phi - \Phi'^*),$$

which holds true because

$$h_n(k,e,\Phi) \leq h_{n-1}(k,e,\Phi) \leq h_{n-1}(k,e,\Phi - \Phi'^*)$$

from Theorem 1(i) and (iii). □

From here, we add Assumption (Q2') and proceed to the discussion further.

Lemma 1. *The value of the game $h_1(k,e,\Phi)$ is convex in Φ at stage $n = 1$. In addition, if* $\log\ q_k(x)$ *is convex,* $\log\ h_1(k,e,\Phi)$ *is also convex.*

Proof. Because $h_1(\cdot)$ is derived as a root ρ of (9), it is a function of Φ. We explicitly denote it by $\rho(\Phi)$. We differentiate both sides of (9) with respect to Φ to obtain

$$\sum_{i\in N(k,e)} c_i \frac{\mathrm{d}q_i^{-1}(\rho)}{\mathrm{d}\rho}\frac{\mathrm{d}\rho}{\mathrm{d}\Phi} = 1.$$

If we differentiate the above once more, the result is

$$\sum_{i\in N(k,e)} c_i \left\{ \frac{\mathrm{d}^2 q_i^{-1}(\rho)}{\mathrm{d}\rho^2}\left(\frac{\mathrm{d}\rho}{\mathrm{d}\Phi}\right)^2 + \frac{\mathrm{d}q_i^{-1}(\rho)}{\mathrm{d}\rho}\frac{\mathrm{d}^2\rho}{\mathrm{d}\Phi^2}\right\} = 0. \tag{16}$$

For $y_i \equiv q_i^{-1}(\rho)$, we have

$$\begin{aligned}
\frac{\mathrm{d}y_i}{\mathrm{d}\rho} &= \frac{1}{\mathrm{d}\rho/\mathrm{d}y_i} = \frac{1}{q_i'(y_i)} = \frac{1}{q_i'(q_i^{-1}(\rho))},\\
\frac{\mathrm{d}^2 y_i}{\mathrm{d}\rho^2} &= \frac{\mathrm{d}}{\mathrm{d}\rho}\left(\frac{\mathrm{d}y_i}{\mathrm{d}\rho}\right) = \frac{\mathrm{d}}{\mathrm{d}y_i}\left(\frac{1}{q_i'(y_i)}\right)\cdot\frac{\mathrm{d}y_i}{\mathrm{d}\rho} = -\frac{q_i''(y_i)}{\{q_i'(y_i)\}^2}\cdot\frac{1}{q_i'(y_i)}\\
&= -\frac{q_i''(q_i^{-1}(\rho))}{\{q_i'(q_i^{-1}(\rho))\}^3}.
\end{aligned}$$

The substitution of the above expressions into (16) brings

$$\sum_{i\in N(k,e)} c_i \left\{ -\frac{q_i''(q_i^{-1}(\rho))}{\{q_i'(q_i^{-1}(\rho))\}^3}\left(\frac{\mathrm{d}\rho}{\mathrm{d}\Phi}\right)^2 + \frac{1}{q_i'(q_i^{-1}(\rho))}\frac{\mathrm{d}^2\rho}{\mathrm{d}\Phi^2}\right\} = 0 \tag{17}$$

and it can be organized into

$$\frac{\mathrm{d}^2\rho}{\mathrm{d}\Phi^2}\sum_{i\in N(k,e)} \frac{c_i}{q_i'(q_i^{-1}(\rho))} = \left(\frac{\mathrm{d}\rho}{\mathrm{d}\Phi}\right)^2 \sum_{i\in N(k,e)} \frac{c_i q_i''(q_i^{-1}(\rho))}{\{q_i'(q_i^{-1}(\rho))\}^3}.$$

From the monotone decreasingness and the strictly convexity of $q_i(x)$, it holds that $q_i'(x) < 0$ and $q_i''(x) > 0$. Therefore, it should be $\mathrm{d}^2\rho/\mathrm{d}\Phi^2 \geq 0$ for the validity of the above equation. Now we have proved the first half of the lemma.

To prove the second half, we use notation $r_i(x) \equiv \log q_i(x)$ and $\sigma(\Phi) \equiv \log\rho(\Phi)$. We substitute the transformed expressions $q_i' = r_i' \exp r_i$, $q_i'' = (r_i')^2 \exp r_i + r_i'' \exp r_i$, $\mathrm{d}\rho/\mathrm{d}\Phi = \exp\sigma\ \mathrm{d}\sigma/\mathrm{d}\Phi$ and $\mathrm{d}^2\rho/\mathrm{d}\Phi^2 = \exp\sigma\ \mathrm{d}^2\sigma/\mathrm{d}\Phi^2 + \exp\sigma\ (\mathrm{d}\sigma/\mathrm{d}\Phi)^2$ into (17) and put the result together with $\exp\{r_i(y_i(\Phi))\} = q_i(y_i(\Phi)) = \rho(\Phi) = \exp\sigma(\Phi)$ to obtain

$$\sum_{i\in N(k,e)} c_i \left\{ -\frac{(r_i')^2\exp r_i + r_i''\exp r_i}{(r_i'\exp r_i)^3}\left(\mathrm{e}^{\sigma}\cdot\frac{\mathrm{d}\sigma}{\mathrm{d}\Phi}\right)^2\right.$$

$$
\left. + \frac{1}{r_i' \exp r_i} \left(e^{\sigma} \cdot \frac{d^2\sigma}{d\Phi^2} + e^{\sigma} \cdot \left(\frac{d\sigma}{d\Phi} \right)^2 \right) \right\}
$$

$$
= \sum_{i \in N(k,e)} c_i \left\{ \frac{(r_i' \exp r_i)^2 - (r_i')^2 \exp r_i \exp \sigma - r_i'' \exp r_i \exp \sigma}{(r_i' \exp r_i)^3} e^{\sigma} \left(\frac{d\sigma}{d\Phi} \right)^2 \right.
$$

$$
\left. + \frac{1}{r_i' \exp r_i} e^{\sigma} \cdot \frac{d^2\sigma}{d\Phi^2} \right\}
$$

$$
= \sum_{i \in N(k,e)} c_i \left\{ - \frac{r_i''}{r_i'^3 \exp r_i} e^{\sigma} \left(\frac{d\sigma}{d\Phi} \right)^2 + \frac{1}{r_i' \exp r_i} e^{\sigma} \cdot \frac{d^2\sigma}{d\Phi^2} \right\} = 0.
$$

Now we see $d^2\sigma/d\Phi^2 \geq 0$ from $r_i' < 0,\ r_i'' \geq 0$. □

Expression (2) is derived from a one-sided optimization problem where the target with no energy can do nothing but stay at cell k and the searcher worries about how he/she distributes the total amount Φ of searching resource into each of the stages. Even for the simple problem, we cannot derive any solution in general but we can have an optimal value (2) by an additional assumption (Q2'). The following theorem states the importance of the assumption.

Theorem 2. *If* $\log\ q_k(x)$ *is convex for x,* $\log\ h_n(k, e, \Phi)$ *is convex for* Φ*, too.*

Proof. From Lemma 1, the theorem is valid for $n = 1$. To prove the theorem by mathematical induction, we assume that it is true for stage $n-1$. For simplicity, let us abbreviate $h_n(k, e, \Phi)$ and $h_{n-1}(i, e - \mu(k, i), \Phi)$ to $h_n(k, \Phi)$ and $h_{n-1}(i, \Phi)$, respectively. The value of the game $h_n(k, \Phi)$ is given by an optimization (13), in which $\rho_n(k, e, \Phi, \Phi')$ is an optimal value for $\min_{\varphi}\ \max_p \sum_{i \in N(k,e)} p(k, i) h_{n-1}(i, \Phi - \Phi') q_i(\varphi(i))$. Therefore, for $0 \leq \alpha \leq 1$, the following transformation is possible:

$$
\alpha \rho_n(k, e, \Phi_1, \Phi_1') + (1-\alpha) \rho_n(k, e, \Phi_2, \Phi_2')
$$

$$
= \alpha \min_{\varphi_1 \in \Psi_E(\Phi_1')} \max_{p_1 \in P_k(e)} \sum_{i \in N(k,e)} p_1(k, i) h_{n-1}(i, \Phi_1 - \Phi_1') q_i(\varphi_1(i))
$$

$$
+ (1-\alpha) \min_{\varphi_2 \in \Psi_E(\Phi_2')} \max_{p_2 \in P_k(e)} \sum_{i \in N(k,e)} p_2(k, i) h_{n-1}(i, \Phi_2 - \Phi_2') q_i(\varphi_2(i)).
$$

We replace p_1 and p_2 with a common variable p and proceed to transform the above further.

$$
\geq \min_{\varphi_1 \in \Psi_E(\Phi_1'), \varphi_2 \in \Psi_E(\Phi_2')} \max_{p \in P_k(e)} \sum_{i \in N(k,e)} p(k, i) \left\{ \alpha h_{n-1}(i, \Phi_1 - \Phi_1') q_i(\varphi_1(i)) \right.
$$

$$
\left. + (1-\alpha) h_{n-1}(i, \Phi_2 - \Phi_2') q_i(\varphi_2(i)) \right\}.
$$

From the convexity of functions $\log h_{n-1}(\cdot)$ and $\log q_i(\cdot)$, $h_{n-1}(i, \Phi') q_i(\varphi(i)) = \exp\{\log h_{n-1}(i, \Phi') + \log q_i(\varphi(i))\}$ is also strictly convex for a variable vector $\{\Phi', \varphi(i)\}$ and then the above expression can be transformed as follows:

$$\geq \min_{\varphi_1 \in \Psi_E(\Phi_1'), \varphi_2 \in \Psi_E(\Phi_2')} \max_{p \in P_k(e)} \sum_{i \in N(k,e)} p(k,i) h_{n-1}(i, \alpha(\Phi_1 - \Phi_1') + (1-\alpha)(\Phi_2 - \Phi_2')) \times q_i(\alpha\varphi_1(i) + (1-\alpha)\varphi_2(i)).$$

$\varphi_1 \in \Psi_E(\Phi_1')$ and $\varphi_2 \in \Psi_E(\Phi_2')$ imply that $\varphi = \alpha\varphi_1 + (1-\alpha)\varphi_2 \in \Psi_E(\alpha\Phi_1' + (1-\alpha)\Phi_2')$. Using $\Phi = \alpha\Phi_1 + (1-\alpha)\Phi_2$ and $\Phi' = \alpha\Phi_1' + (1-\alpha)\Phi_2'$, we have

$$\begin{aligned} &\geq \min_{\varphi \in \Psi_E(\Phi')} \max_{p \in P_k(e)} \sum_{i \in N(k,e)} p(k,i) h_{n-1}(i, \Phi - \Phi') q_i(\varphi(i)) \\ &= \rho_n(k, e, \Phi, \Phi') = \rho_n(k, e, \alpha\Phi_1 + (1-\alpha)\Phi_2, \alpha\Phi_1' + (1-\alpha)\Phi_2'). \end{aligned}$$

Now we understand that $\rho_n(k, e, \Phi, \Phi')$ is convex for (Φ, Φ').

From (7), we can see that the value $\rho_n(k, e, \Phi, \Phi')$ equals $h_{n-1}(i, \Phi - \Phi') q_i(\varphi(i))$ for a certain cell i. That is, $\log \rho_n(k, e, \Phi, \Phi')$ equals $\log\{h_{n-1}(i, \Phi - \Phi') q_i(\varphi(i))\}$ for a certain i, which is a convex expression for φ. As the result, we can say that $\log \rho_n(k, e, \Phi, \Phi')$ is given by a minimax optimization.

$$\begin{aligned} &\log \rho_n(k, e, \Phi, \Phi') \\ &\quad = \min_{\varphi \in \Psi_E(\Phi')} \max_{p \in P_k(e)} \sum_{i \in N(k,e)} p(k,i) \log\{h_{n-1}(i, \Phi - \Phi') q_i(\varphi(i))\}. \end{aligned}$$

After repeating the previous discussion with respect to the minimax optimization problem, it is understood that $\log \rho_n(k, e, \Phi, \Phi')$ is convex for (Φ, Φ').

Problem (13) is equivalent to a problem

$$\log h_n(k, e, \Phi) = \min_{\{\Phi', 0 \leq \Phi' \leq \Phi\}} \log \rho_n(k, e, \Phi, \Phi').$$

We have equations

$$\alpha \log h_n(k, e, \Phi_1) = \min_{0 \leq \Phi_1' \leq \Phi_1} \alpha \log \rho_n(k, e, \Phi_1, \Phi_1')$$

$$(1-\alpha) \log h_n(k, e, \Phi_2) = \min_{0 \leq \Phi_2' \leq \Phi_2} (1-\alpha) \log \rho_n(k, e, \Phi_2, \Phi_2')$$

for $0 \leq \alpha \leq 1$, $\Phi_1 > 0$, $\Phi_2 > 0$ and an inequality $0 \leq \alpha\Phi_1' + (1-\alpha)\Phi_2' \leq \alpha\Phi_1 + (1-\alpha)\Phi_2$ for $0 \leq \Phi_1' \leq \Phi_1$, $0 \leq \Phi_2' \leq \Phi_2$. From the convexity of $\log \rho_n(k, e, \Phi, \Phi')$ with respect to (Φ, Φ'), the following transformation is possible:

$$\begin{aligned} &\alpha \log h_n(k, e, \Phi_1) + (1-\alpha) \log h_n(k, e, \Phi_2) \\ &\quad \geq \min_{\{(\Phi_1', \Phi_2') | 0 \leq \alpha\Phi_1' + (1-\alpha)\Phi_2' \leq \alpha\Phi_1 + (1-\alpha)\Phi_2\}} \\ &\quad \left\{\alpha \log \rho_n(k, e, \Phi_1, \Phi_1') + (1-\alpha) \log \rho_n(k, e, \Phi_2, \Phi_2')\right\} \end{aligned}$$

$$\geq \min_{\{(\Phi_1',\Phi_2')|0\leq\alpha\Phi_1'+(1-\alpha)\Phi_2'\leq\alpha\Phi_1+(1-\alpha)\Phi_2\}} \log\rho_n(k,e,\alpha\Phi_1+(1-\alpha)\Phi_2,\alpha\Phi_1'+(1-\alpha)\Phi_2')$$
$$= \min_{\{\Phi'|0\leq\Phi'\leq\alpha\Phi_1+(1-\alpha)\Phi_2\}} \log\rho_n(k,e,\alpha\Phi_1+(1-\alpha)\Phi_2,\Phi') = \log h_n(k,e,\alpha\Phi_1+(1-\alpha)\Phi_2).$$

Now we finally have completed the proof. □

The convexity of $\log h_n(k,e,\Phi)$ indicates the convexity of the following expression, which appears in formulation (3), for $\{\varphi(i), i\in N(k,e)\}$.

$$\log\ h_{n-1}(i,e-\mu(k,i),\Phi-\sum_j c_j\varphi(j)) + \log\ q_i(\varphi(i)).$$

In practice, we can verify it by

$$\log\ h_{n-1}\left(i,e-\mu(k,i),\Phi-\left(\alpha\sum_j c_j\varphi(j)+(1-\alpha)\sum_j c_j\varphi'(j)\right)\right)$$
$$= \log\ h_{n-1}\left(i,e-\mu(k,i),\alpha\left(\Phi-\sum_j c_j\varphi(j)\right)+(1-\alpha)\right.$$
$$\left.\times\left(\Phi-\sum_j c_j\varphi'(j))\right)\right) \leq \alpha\log\ h_{n-1}\left(i,e-\mu(k,i),\Phi-\sum_j c_j\varphi(j)\right)$$
$$+(1-\alpha)\log\ h_{n-1}\left(i,e-\mu(k,i),\Phi-\sum_j c_j\varphi'(j)\right).$$

Therefore, $h_{n-1}(i,e-\mu(k,i),\Phi-\sum_j c_j\varphi(j))q_i(\varphi(i))$ is strictly convex for $\{\varphi(i), i\in N(k,e)\}$. We already know that for a game with the expected payoff, which is linear for the strategy of one player and convex for the strategy of the other player, its minimax value coincides with its maxmin value, as shown in Hohzaki [12]. Hence, the minimax value given by (13) is truly the value of the game and optimal strategies are given by formulas (14) and (15). The following theorem establishes the existence of the value of the game.

Theorem 3. *If* $\log\ q_k(x)$ *is convex for* x*, there exists an equilibrium point consisting of a target strategy of selecting his next cell probabilistically and a searcher strategy of a distribution plan of his searching resource. The minimax value always coincides with the maxmin value for the expected payoff, which gives us the value of the game.*

5 Numerical Examples

Here, we adopt $q_i(x) = \exp(-\alpha_i x)$ as a function of the nondetection probability of target. As mentioned in the end of Sect. 3, [10] studied this special case in accordance with the basic procedure for solution (Step 1)–(Step 3) in Sect. 3 and obtained the equilibrium in closed-form. Here, we only outline the value of the game.

For stage n, target cell k, residual target energy e and residual budget of the searcher Φ, the value of the game $h_n(k, e, \Phi)$ is given by the following expression:

$$h_n(k, e, \Phi) = \exp(-\Phi/\gamma_n(k, e)),$$

where coefficient $\gamma_n(k, e)$ is independent of budget Φ. By a numerical algorithm, the coefficient is calculated from parameters α_j and c_j of cell j in the reachable region of the target in the future from his present cell. It indicates the inefficiency of unit searching budget for detection of target because the nondetection probability $h_n(k, e, \Phi)$ gets larger as the coefficient becomes larger. Hohzaki also derived the closed-forms of optimal target strategy $\{p^*(k, i)\}$ and searcher's strategy $\{\varphi^*(i)\}$.

Based on the preliminary knowledge above, we consider a multistage search game in a discrete search space, illustrated in Fig. 1. Five cells align in each row and each column so that there are 25 cells in total. Parameter $\alpha_i = 1$ is set for black-colored cells: 1, 7, 8, 12, 13, 14, 18 and 19, and $\alpha_i = 0.1$ for gray-colored cells: 2, 3, 4, 5, 6, 9, 10, 11, 15, 16, 17, 20, 21, 22, 23 and 24, while Cell 25 has a $\alpha_i = 0.01$. All cells have the same parameter $c_i = 1$. The parameter α_i indicates the effectiveness of searching resource in the relevant cell i. For larger α_i, the searcher expects larger probability of detecting target there. Therefore, the target had better avoid the black cells, but he/she is willing to choose cell 25 as a kind of shelter if he/she can go there. Energy-consumption function $\mu(i, j)$ is set to be one if cells i and j are adjacent to each other and four if i and j are two blocks far away from each other. From cell 7, the target spends 1 energy to move to cells 1 or 2 and spends 4 to go to cells 17, 19 or 4, for example. However, the target cannot move to cells three blocks or more away from the current cell.

For several cases of (n, k, e) as combination (current stage, present cell of target, energy of target in hand), we calculate coefficient $\gamma_n(k, e)$, optimal distribution

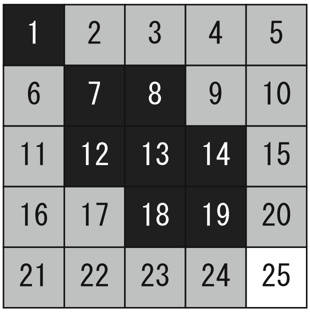


Fig. 1 Search Space

Fig. 2 Case 1: (n,k,e)=(2,7,1)

	. 244 0	. 244 0		
. 244 0	. 024 . 072			
. 244 0				

Fig. 3 Case 2: (n,k,e)=(2,13,2)

	. 037 . 037	. 037 . 037	. 370 . 370	
	. 037 . 037	. 037 . 037	. 037 . 037	
	. 370 . 370	. 037 . 037	. 037 . 037	

Fig. 4 Case 3: (n,k,e)=(2,13,4)

		. 001 0		
			. 001 0	
				. 998 0

of searching resource for the searcher, $\varphi^*(i)$, and optimal moving strategy of the target, $p^*(k,i)$, and illustrate them in Figs. 2–6. A cell enclosed with 4 bold lines indicates the present cell of the target. In each cell, two numbers are written. The upper is $p^*(k,i)$ and the lower is $\varphi^*(i)$. No number means that the cell is not chosen by the target and that the searcher does not scatter any searching resource there. According to Hohzaki [10], optimal distribution $\varphi^*(i)$ is proportional to searching budget Φ. We write only the proportional constant of $\varphi^*(i)$ as the optimal distribution of resource. The number indicates how many percentage of residual resource Φ the searcher should distribute in the cell because budget Φ is equal to the amount of searching resource since $c_k = 1$.

We now analyze optimal strategies of players in some cases of (n,k,e). In the caption of each case below, we show (n,k,e), $\gamma_n(k,e)$ and how many percentage

Fig. 5 (n,k,e)=(1,13,4)

.001 .001	.009 .009	.009 .009	.009 .009	.009 .009
.009 .009	.001 .001	.001 .001	.009 .009	.009 .009
.009 .009	.001 .001	.001 .001	.001 .001	.009 .009
.009 .009	.009 .009	.001 .001	.001 .001	.009 .009
.009 .009	.009 .009	.009 .009	.009 .009	.856 .856

Fig. 6 (n,k,e)=(1,19,3)

		.001 .001	.001 .001	.010 .010
		.001 .001	.001 .001	.010 .010
		.010 .010	.010 .010	.958 .958

points of all resources should be consumed at the current stage, that is, Φ'^*/Φ, where Φ'^* is an optimal value of variable Φ' for optimization problem (13).

Case 1. (Fig. 2): $(n,k,e) = (2,7,1)$, $\gamma_n(k,e) = 10.78$, $\Phi'^*/\Phi = 0.0722$

(1) The target does not go to neighborhood cells with $\alpha_i = 1$. But he/she goes to neighborhood cells with $\alpha_i = 0.1$, that is, cells 2, 3, 6 and 11, or chooses to stay at the present cell 7 without consuming any energy. The probability of staying is one tenth of that of moving.

(2) The searcher distributes just 0.072 of searching resource in the present cell 7 of the target at the current stage $n = 2$. At the beginning of the next stage $n = 1$, the searcher is informed of the target position. If the position is cell 2, 3, 6, or 11, the target must have exhausted his energy so that the searcher will be able to focus his/her resource into the cell. That is why the searcher leaves most of his/her resource, 1-0.072=0.928, for the last stage $n = 1$.

Case 2. (Fig. 3): $(n,k,e) = (2,13,2)$, $\gamma_n(k,e) = 27.0$, $\Phi'^*/\Phi = 1.0$

(1) If the energy is larger, $e = 2$, the target can move to cell 19 expecting to slip into the shelter-cell 25 at the last stage $n = 1$. Therefore, the searcher exhausts his/her resource at the present stage.

(2) Case $(n,k,e) = (2,13,3)$ with more energy has the same optimal strategies as this case.

Case 3. (Fig. 4): $(n,k,e) = (2,13,4)$, $\gamma_n(k,e) = 1000.2$, $\Phi'^*/\Phi = 0.0002$

(1) As the energy gets much larger to $e = 4$, the target can directly move to the shelter-cell 25 at this stage. He/she chooses this option with high probability 0.998. One of other cells he could move to is 19 with comparatively large efficiency of search. But the cell 19 is only the cell via which the target finally goes to cell 25 at stage $n = 1$. The last option of the target is to stay at the current cell 13, from which he can move to cell 25 at the last stage.

(2) The searcher can anticipate that the target is most likely to move to cell 25 with probability 0.998. The situation will be the same at the next stage $n = 1$. The searcher keeps almost all resource to focus his searching resource for the next search at $n = 1$. Practically, a very small fraction of resource 0.0002 is used at the present stage $n = 2$.
If the target goes to cell 25, his/her energy runs on empty so that the searcher will distribute all his resource there. When the target moves to cell 13 or 19, optimal strategies of players are shown in Figs. 5 or 6, respectively. The target could choose to move to every cell the residual energy allows him to go to. In spite of these situations, the probability that the target chooses cell 25 is considerably high.

6 Conclusions

In this chapter, we deal with a multistage stochastic game called a multistage search allocation game (MSSAG) with detection probability as payoff. A searcher distributes his/her searching resource to detect a target and the target moves to evade the searcher. At the beginning of each stage, both players know the current position and the residual energy of the target and the residual budget of the searcher. We formulate the problem without specifying a specific form for the nondetection function which indicates how inefficient the searching resource is on the detection probability of target. Assuming only the convexity and the differentiability of the function, we discuss the existence of the value of the game and some properties of optimal strategies of players. Generally speaking, the game has properties similar to a general convex game, the details of which are stated in Theorems 1 and 3. We prove that the game converges to a steady state as the number of stages becomes infinity but we need to specify the form of the nondetection function to derive equilibrium points, the value of the game and the convergence point.

We show that a larger number of stages has negative effects on the target side. It is caused from the structure of information sets and the payoff. Specifically, the structure is set to be advantageous for the searcher. The feasibility of the assumptions ought to be examined in some practical modeling in the future. In our modeling, we take the detection probability as payoff. However, there are a variety of criteria for which more complicated forms of payoffs could be required.

References

1. Baston, V.J., Bostock, F.A.: A one-dimensional helicopter-submarine game. Naval Research Logistics **36**, 479–490 (1989)
2. Baston, V.J., Garnaev, A.Y.: A search game with a protector. Naval Research Logistics **47**, 85–96 (2000)
3. Danskin, J.M.: The Theory of Max-Min. Springer, N.Y. (1967)
4. Danskin, J.M.: A helicopter versus submarine search game. Operations Research **16**, 509-517 (1968)
5. Eagle, J.N., Washburn, A.R.: Cumulative search-evasion games. Naval Research Logistics **38**, 495–510 (1991)
6. Gal, S.: Search Games. Academic, N.Y. (1980)
7. Garnaev, A.Y.: On a Ruckle problem in discrete games of ambush. Naval Research Logistics **44**, 353–364 (1997)
8. Garnaev, A.Y.: Search Games and Other Applications of Game Theory. Springer, Tokyo (2000)
9. Hohzaki, R.: A search allocation game. European Journal of Operational Research **172**, 101–119 (2006)
10. Hohzaki, R.: A multi-stage searh allocation game with the payoff of detection probability. Journal of the Operations Research Society of Japan **50**, 178–200 (2007)
11. Hohzaki, R., Iida, K.: A search game with reward criterion. Journal of the Operations Research Society of Japan **41**, 629–642 (1998)
12. Hohzaki, R., Iida, K.: A solution for a two-person zero-sum game with a concave payoff function. In: Takahashi, W., Tanaka, T. (eds.) Nonlinear Analysis and Convex Analysis, pp. 157–166. World Science Publishing, London (1999)
13. Hohzaki, R., Iida, K.: A search game when a search path is given. European Journal of Operational Research **124**, 114–124 (2000)
14. Hohzaki, R., Iida, K., Komiya, T.: Discrete search allocation game with energy constraints. Journal of the Operations Research Society of Japan **45**, 93–108 (2002)
15. Iida, K., Hohzaki, R., Furui, S.: A search game for a mobile target with the conditionally deterministic motion defined by paths. Journal of the Operations Research Society of Japan **39**, 501–511 (1996)
16. Iida, K., Hohzaki, R., Sato, K.: Hide-and-search game with the risk criterion. Journal of the Operations Research Society of Japan **37**, 287–296 (1994)
17. Koopman, B.O.: Search and Screening. pp. 221–227. Pergamon, (1980)
18. Meinardi, J.J.: A sequentially compounded search game. In: Theory of Games: Techniquea and Applications, pp. 285–299. The English Universities Press, London (1964)
19. Nakai, T.: A sequential evasion-search game with a goal. Journal of the Operations Research Society of Japan **29**, 113–122 (1986)
20. Nakai, T.: Search models with continuous effort under various criteria. Journal of the Operations Research Society of Japan **31**, 335–351 (1988)
21. Owen, G.: Game Theory. Academic, N.Y. (1995)
22. Ruckle, W.H.: Ambushing random walk II: continuous models. Operations Research **29**, 108–120 (1981)
23. Shapley, L.S.: Stochastic games. Proceedings of the National Academy of Sciences of the U.S.A. **39**, 1095–1100 (1953)
24. Stone, L.D.: Theory of Optimal Search. pp. 136–178. Academic, N.Y. (1975)
25. Washburn, A.R.: Search-evasion game in a fixed region. Operations Research **28**, 1290–1298 (1980)
26. Washburn, A.R., Hohzaki, R.: The diesel submarine flaming datum problem. Military Operations Research **6**, 19–33 (2001)

Linear Differential Game with Two Pursuers and One Evader

Stéphane Le Ménec

1 Introduction

We consider head-on scenarios with two pursuers (P1 and P2) and one evader (E). The purpose of this paper is to compute two-on-one differential game No Escape Zones (NEZ) [1, 8]. The main objective consists in designing suboptimal strategies in many-on-many engagements. 1×1 NEZ and 2×1 NEZ are components involved in suboptimal approaches we are interested in (i.e. "Moving Horizon Hierarchical Decomposition Algorithm" [6]). Several specific "two pursuers one evader" differential games have been already studied [2, 4, 7]. Nevertheless, we propose to compute 2×1 NEZ from 1×1 DGL/1 NEZ because DGL models are games with well-defined analytical solutions [12].

The overview of the paper is as follows. We first recall the one-on-one DGL/1 game model. Then, we define the criterion of the two-on-one game and we give physical hints about the evader behavior with respect to different initial conditions. For the case with same time-to-go we start computing the evader strategy in the particular situation where the evader is able to evade and has to achieve a trade-off between evading the first and the second pursuer. Then, we extend the solution to the other initial conditions and we design the new 2×1 NEZ. We finally solve the case where the times-to-go are not the same.

2 One-on-one DGL/1 Game

The original problem is nonlinear, however under some assumptions a linear formulation can be obtained after linearization of the original dynamics. As a recall, we briefly summarize some results about one-on-one pursuit evasion games using

S. Le Ménec (✉)
EADS/MBDA France, 1 avenue Réaumur, 92358 Le Plessis-Robinson Cedex, France
e-mail: stephane.le-menec@mbda-systems.com

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_11,

DGL/1 models. DGL/1 differential games are coplanar interceptions with constant velocities, bounded controls assuming small motion variations around the collision course triangle. Under this assumption, the kinematics is linear. Each player is represented as a first-order system (time lag constant). The criterion is the terminal miss distance (terminal cost only, absolute value of the terminal miss perpendicular to the initial Line Of Sight, LOS). DGL/1 are fixed time duration differential games, with final time defined by the closing velocity (assumed constant) and the pursuer-evader range. The terminal projection procedure [3] allows to reduce the initial four dimension state vector representation to a scalar representation and to represent the optimal trajectories in the ZEM (Zero Effort Miss), Tgo (Time to Go) coordinate frame (i.e. related coordinates forecast under absence of controls to the terminal instant). According to notations and normalizations defined in [12], the DGL/1 kinematics is as follows (please refer to the Appendix for more explanations about the terminal projection procedure and about the DGL/1 kinematics):

$$\frac{\mathrm{d}z}{\mathrm{d}\theta} = \mu h\left(\theta\right) u - \varepsilon h \left(\frac{\theta}{\varepsilon}\right) v$$

$$h\ (\alpha) = \mathrm{e}^{-\alpha} + \alpha - 1$$

$$\mu = \frac{a_{P\max}}{a_{E\max}}, \quad \varepsilon = \frac{\iota_E}{\iota_P}.$$

In this framework, $u = \frac{a_{Pc}}{a_{P\max}}$ and $v = \frac{a_{Ec}}{a_{E\max}}$ are, respectively, the pursuer and evader controls ($|u| \leq 1$, $|v| \leq 1$). $a_{P\max}$ and $a_{E\max}$ are the maximum accelerations. a_{Pc} and a_{Ec} are the lateral acceleration demands. ι_P and ι_E are the pursuer and evader time lag constants. The independent variable is the normalized time to go:

$$\theta = \frac{\tau}{\iota_P}, \quad \tau = t_f - t,$$

where t_f is the final time. The nondimensional state variable is the normalized Zero Effort Miss:

$$z\left(\theta\right) = \frac{ZEM\left(\tau\right)}{\iota_P^2 a_{E\ \max}}$$

The Zero Effort Miss distance is given below for DGL/1:

$$ZEM(\tau) = y + \dot{y}\tau - \ddot{y}_P \iota_P^2 h\left(\theta\right) + \ddot{y}_E \iota_E^2 h\left(\frac{\theta}{\varepsilon}\right).$$

Here, y is the relative perpendicular miss, $\dot{y}$ the relative perpendicular velocity, $\ddot{y}_P$ the (instantaneous) perpendicular pursuer acceleration, and $\ddot{y}_E$ the perpendicular evader acceleration. The nondimensional cost function is the normalized terminal miss distance subject to minimization by the pursuer and maximization by the evader:

$$J = |z_f| = |z(\theta = 0)|$$

The (ZEM, Tgo) frame is divided into two regions, the regular area and the singular one. For some appropriate differential game parameters (pursuer to evader maximum acceleration ratio μ and evader to pursuer time lag ratio ε), the singular area plays the role of a capture zone also called NEZ (leading to zero terminal miss), while the regular area corresponds to the noncapture zone. The NEZ can be bounded in time (closed) or unbounded in time (open). However, for all the individual games, we restrict the game parameters to the cases corresponding to nonempty, unbounded, and monotone increasing (in backward time) NEZ ($\mu > 1$). The natural optimal strategies are bang–bang controls corresponding to the sign of ZEM (some refinements exist when defining optimal controls inside the NEZ). We start the 2×1 DGL/1 analysis with unbounded 1×1 NEZ as pictured in Fig. 1 (NEZ delimited by the two solid lines, noncapture zone corresponding to the state space filled with optimal trajectories in dashed lines). Moreover, we first assume same Tgo in each DGL/1 game (same initial range, same velocity for each pursuer).

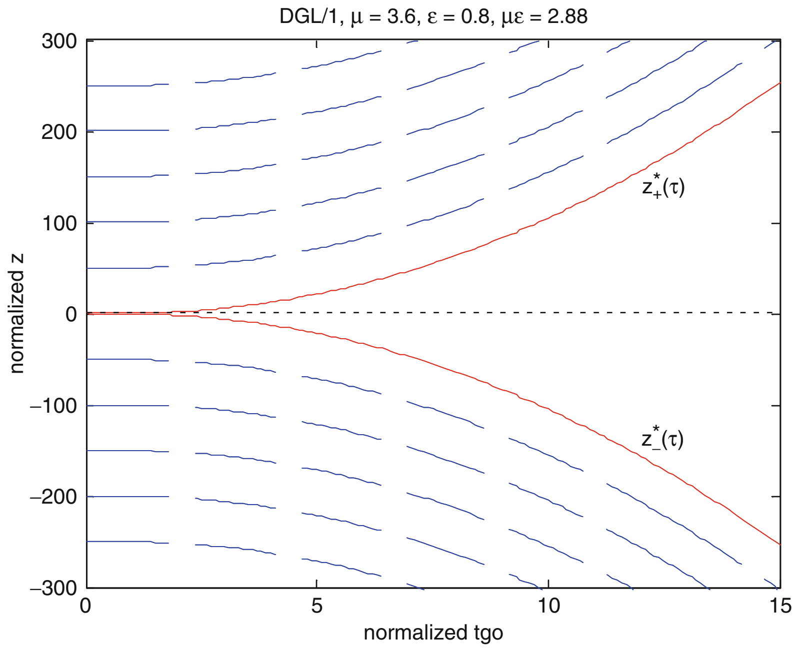


Fig. 1 Unbounded NEZ (ZEM, Tgo frame)

3 Two Pursuers One Evader DGL/1 Game

3.1 Criterion

The outcome we consider in the 2×1 game is the minimum of the two terminal miss distances.

$$J_{2\times1}(u_1, u_2, v) = \min\{J_1(u_1, v), J_2(u_2, v)\}$$

with J the terminal miss, u_i the control of Pi, $i = \{1, 2\}$ and v the evader control.

$$\begin{aligned} J_1(u_1, v) &: \text{ P1 E DGL/1 terminal miss} \\ J_2(u_2, v) &: \text{ P2 E DGL/1 terminal miss} \\ J_{2\times1}(u_1, u_2, v) &: \ 2 \times 1 \text{ DGL/1 outcome.} \end{aligned}$$

In this case (not true in general), the 2×1 game only makes a change in the evader optimal command. The presence of a second pursuer does not change the behavior of the pursuers. Indeed, the optimal controls for each pursuer in the 2×1 game are the same as the ones in their respective 1×1 game.

When considering two pursuers and one evader (in the framework of DGL representations), some cases are easy to solve. If E is "above" (or symmetrically "below") the two pursuers (in ZEM, Tgo frame), the optimal evasion with respect to each pursuer (considered alone) and according to both pursuers together is to turn right (resp. left) with maximum acceleration. In these cases (E "above" or "below"), there are no changes in evasion trajectory (optimal strategies) due to the addition of a pursuer.

For the other initial conditions (see Fig. 2), we have to refine the optimal evasion behavior. If E is between P1 and P2 (as described in Fig. 2), then the optimal 2×1 evasion is a trade-off between the incompatible P1 E and P2 E optimal escapes. This situation is the most interesting one. Even if the evader is outside both

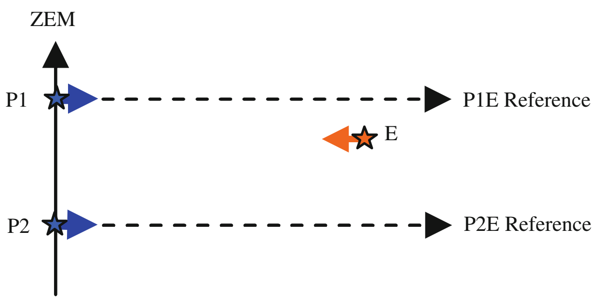


Fig. 2 2×1 game with Evader between the two Pursuers (ZEM, Tgo frame)

NEZes (and between the two pursuers) a capture by collaboration of two pursuers is possible. The optimal behavior of the evader is to go to the center of the segment between the positions of two pursuers at the terminal instant. Please, notice that in Fig. 2 the initial LOS references (in each DGL/1 game) have been chosen parallel.

3.2 Trade-off Evader Control

We first focus on the particular case where the evader is between P1 and P2; when the evader is able to evade P1 and P2 (initial conditions outside the NEZ) and when no control saturation occurs on the optimal trade-off strategy $v^*_{2\times1}$. We compute the 2×1 DGL/1 game solution $\left(J_{2\times1}\left(u^*_1, u^*_2, v^*_{2\times1}\right)\right)$ considering that the optimal evasion control is a constant value during the entire game and considering that the terminal miss distances with respect to P1 and P2 have to be the same. In the case of Fig. 2, $v^*_1 = -1$ (left turn), $v^*_2 = 1$ (right turn) and $-1 \le v^*_{2\times1} \le 1$. In the specific game we encounter (see Appendix about the 1×1 game features for more explanations), we have the following properties:

$$\begin{array}{c} J_1\left(u^*_1, v^*_2\right) \\ J_2\left(u^*_2, v^*_1\right) \\ J_{2\times1}\left(u^*_1, u^*_2, v\right) \end{array} \le J_{2\times1}\left(u^*_1, u^*_2, v^*_{2\times1}\right) = J_1\left(u^*_1, v^*_{2\times1}\right) = J_2\left(u^*_2, v^*_{2\times1}\right) \le \begin{array}{c} J_1\left(u^*_1, v^*_1\right) \\ J_2\left(u^*_2, v^*_2\right) \end{array}$$

According to notations and normalizations detailed in the previous section, we write the final distances equality condition in the following manner (where Z_i are the ZEM in each PE game):

$$Z_1\left(t_f\right) = -Z_2\left(t_f\right).$$

Note that the minus sign is due to the y-axis orientation like in Fig. 2. In normalized variable, we get (with ι_{Pi} the Pursuer time lag constants):

$$z_{1f}\iota^2_{P1} = -z_{2f}\iota^2_{P2}, \quad \left(z_{if} = z_i\left(\tau = 0\right)\right).$$

Looking at Fig. 3, the normalized zero-effort-miss at final time can be written (where Δz are the miss distance diminutions due to no more application of bang–bang evader controls):

$$\begin{aligned} z_{1f} &= z_1(\tau) + z_{1\max}(\tau) + \Delta z_1(\tau = 0) \\ z_{2f} &= z_2(\tau) - z_{2\max}(\tau) + \Delta z_2(\tau = 0). \end{aligned}$$

The computation of the trade-off strategy is based on the following main idea. The evasion strategy $v^*_{2\times1}$ affects the miss distances with respect to P1 and to P2 in accordance with $\Delta z_1(\tau = 0)$ and with $\Delta z_2(\tau = 0)$. Δz_1 increases (z_1 is negative, $\Delta z_1 \ge 0$) as long as $v^*_{2\times1}$ is far from $v^*_{1\times1/P1}$ (strategy in the P1E game). In the same

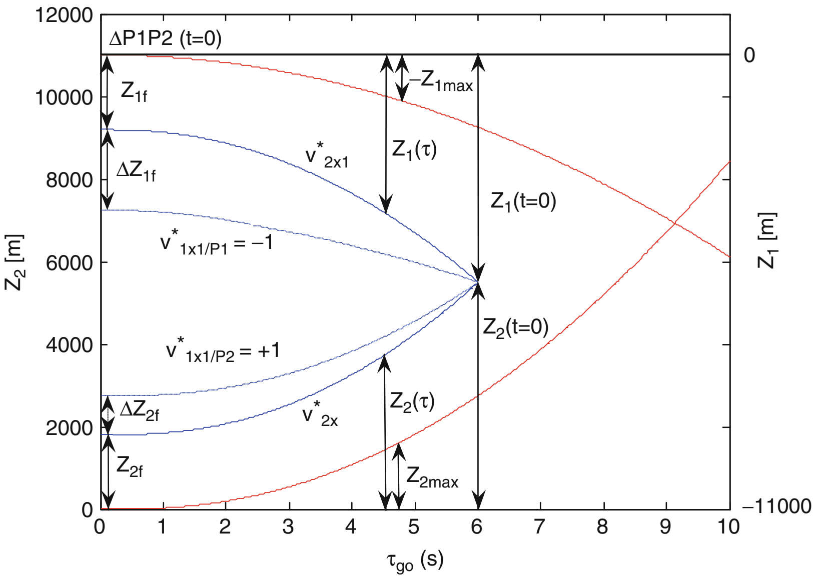


Fig. 3 Calculation of $v^*_{2\times1}$, $[\Delta P1P2\,(\tau) = z_2\,(\tau) - z_1\,(\tau)]$

way, Δz_2 decreases when $v^*_{2\times1}$ is nonequal to $v^*_{1\times1/P2}$ (z_2 is positive, $\Delta z_2 \leq 0$). The trade-off evasion is obtained when the minimum of the two miss distances is enlarged as much as possible (equal in case of no saturation).

In Fig. 3, for simplicity reasons we plot the 2×1 game in the ZEM Tgo frame on one single figure (Z_2 vertical axis on the left side, Z_1 on the right side). It is done by superimposition of the two plots of each pursuer-evader pair with the game initial condition as point of coincidence (P2 located at origin and P1 located at $x = 0$ and $y = 11000 = \Delta P1P2\,(t = 0)$), even if during the games, the Δ P1 P2 range decreases.

The 1×1 DGL/1 no-escape-zone expressions are well known [11]:

$$\begin{aligned} z_{1\,\max} &= \mu_1 H_{11} - \varepsilon_1 H_{21} \\ z_{2\,\max} &= \mu_2 H_{12} - \varepsilon_2 H_{22}, \end{aligned}$$

where:

$$\begin{aligned} H_{1i} &= H_1\,(\theta)\,, \quad i = 1,2 \\ H_1\,(\theta) &= \int_0^\theta h\,(\xi)\,\mathrm{d}\xi = \frac{\theta^2}{2} - h\,(\theta) \end{aligned}$$

and

$$\begin{aligned} H_{2i} &= H_2\,(\theta, \varepsilon_i)\,, \quad i = 1,2 \\ H_2\,(\theta, \varepsilon) &= \int_0^\theta h\left(\frac{\xi}{\varepsilon}\right)\mathrm{d}\xi = \frac{\theta^2}{2\varepsilon} - \varepsilon h\left(\frac{\theta}{\varepsilon}\right). \end{aligned}$$

The third term Δz_i that corresponds to the changes in z_{if} due to $v^*_{2\times 1}$ is:

$$\Delta z_1 = \varepsilon_1 H_{21}\left(v^*_{2\times 1} - v^*_{1\times 1/P1}\right) = \varepsilon_1 H_{21}\left(v^*_{2\times 1} + 1\right); \quad v^*_{1\times 1/P1} = -1 \text{ when considering P1E}$$

$$\Delta z_2 = \varepsilon_2 H_{22}\left(v^*_{2\times 1} - v^*_{1\times 1/P2}\right) = \varepsilon_2 H_{22}\left(v^*_{2\times 1} - 1\right); \quad v^*_{1\times 1/P2} = +1 \text{ when considering P2E.}$$

Substituting these expressions lead to:

$$z_{1f} = z_1 + (\mu_1 H_{11} - \varepsilon_1 H_{21}) + \varepsilon_1 H_{21}\left(v^*_{2\times 1} + 1\right)$$
$$z_{2f} = z_2 - (\mu_2 H_{12} - \varepsilon_2 H_{22}) + \varepsilon_2 H_{22}\left(v^*_{2\times 1} - 1\right)$$

from which, using the initial equality, we can isolate $v^*_{2\times 1}$ (subject to saturation if $v^*_{2\times 1}$ is higher than the evader bounds):

$$v^*_{2\times 1}(z) = \frac{(-\mu_1 H_{11} - z_1)\,\iota^2_{P1} + (\mu_2 H_{12} - z_2)\,\iota^2_{P2}}{\varepsilon_2 H_{22}\iota^2_{P2} + \varepsilon_1 H_{21}\iota^2_{P1}}$$

$v^*_{2\times 1}$ is relevant outside the 2×1 NEZ and is a "reasonable" solution inside the capture zone too. Figure 4 illustrates the optimal trajectories we obtain when the time to go are the same for each pursuer ($\mu_1 = 2$, $\varepsilon_1 = 2$, $\mu_2 = 3$, $\varepsilon_2 = 0.85714$).

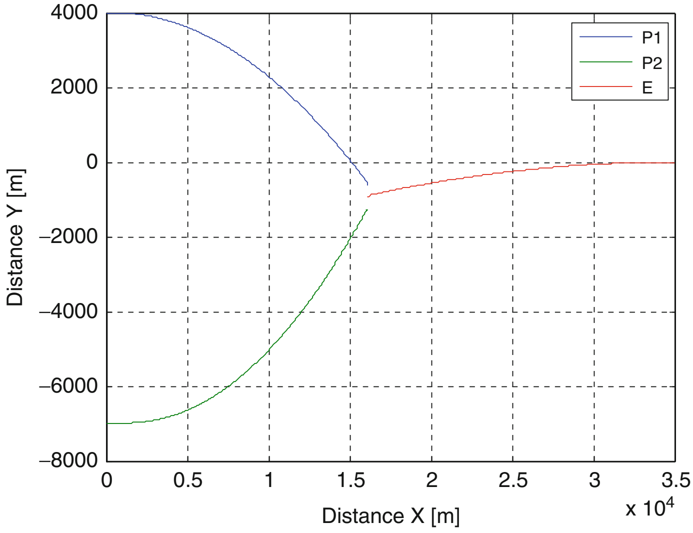


Fig. 4 Example with evader playing $v^*_{2\times 1}$ (E on the right side, P1 top left and P2 bottom left)

3.3 Two-on-one No Escape Zone

An assumption is made for the parameters of the games that NEZ in each game is nonempty, unbounded, and monotonously grows in backward time (general case). The state space of the 2×1 game is divided into five regions. In the first one, P1 is able to intercept any evader maneuver (P1 E NEZ). The second region is the P2 capture zone. For the game parameters, we consider the intersection between P1 E NEZ and P2 E NEZ is nonempty. In the overlapping area, P1 or P2 can capture E alone. Moreover, there exists a capture zone where both P1 and P2 are required to intercept E. The purpose of this section is to compute this "extra part" of the NEZ. All these state space regions belong to the 2×1 NEZ. The remaining state space is the noncapture zone.

The overall 2×1 NEZ is then the collection of P1 E NEZ plus P2 E NEZ plus an extra state space area. If the evader plays in the worst way (from the evader point of view) $\mathrm{v_i} = -\mathrm{sign}\,(\mathrm{z_i})$ then a new limit is under consideration. In addition to $z_{i\,\max}$, we define $z_{i\,\min}$:

$$z_{1\,\min} = \mu_1 H_{11} + \varepsilon_1 H_{21}$$

$$z_{2\,\min} = \mu_2 H_{12} + \varepsilon_2 H_{22}.$$

The new 2×1 frontier is then the line between the two intersections that are solution of the equations:

$$Z_{2\,\max}(\theta) = -Z_{1\,\min}(\theta) + \Delta P_1 P_2\,(t = 0)$$

$$Z_{2\,\min}(\theta) = -Z_{1\,\max}(\theta) + \Delta P_1 P_2\,(t = 0)\,.$$

These two equations give a unique solution $\tau_{2\times1}$, which characterizes the new limit in the 2×1 game.

Figure 5 shows the boundary of the 2×1 game ($\tau_{2\times1} \approx 7.2$ sec .) with the same parameters as in the previous section. The state space area outside the two 1×1 NEZ (left part between the solid curves before intersection) but with $\tau \geq \tau_{2\times1}$ (vertical segment in solid line) now also belongs to the 2×1 NEZ.

The justification of this result is as follows. On the upper limit of the segment, the evader has to play the optimal evasion (-1) against P1 because this point belongs to the limit of the P1/E capture zone. By construction of the segment, this point is the last point allowing to avoid P2 playing -1 (the optimal evasion against P2 is normally $+1$). The situation is symmetric when considering the other extreme point of the segment. Numerical simulations with the evader playing $v^*_{2\times1}$ on several cases confirmed the intuition that this limit is a vertical segment and not a more complex curve.

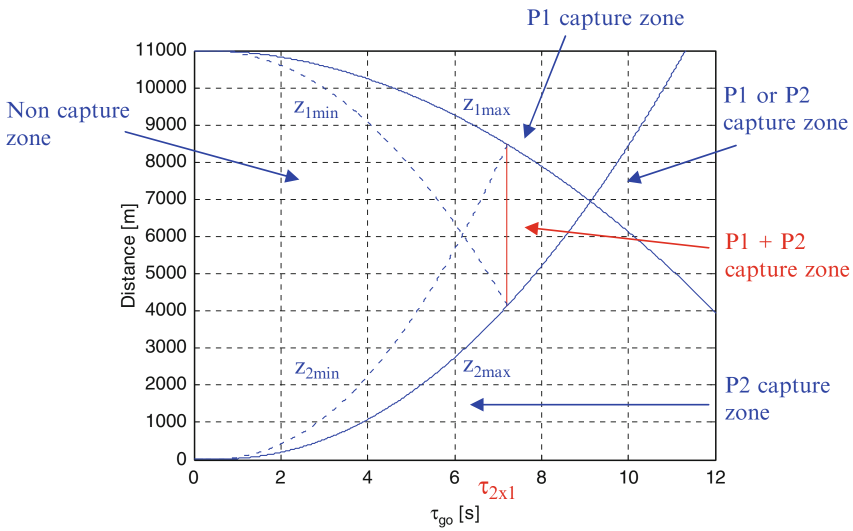


Fig. 5 NEZ extension in 2 × 1 game ($\mu_1 = 2$, $\varepsilon_1 = 2$, $\mu_2 = 3$, $\varepsilon_2 = 0.85714$, $\tau_{2\times 1} \approx 7.2096$)

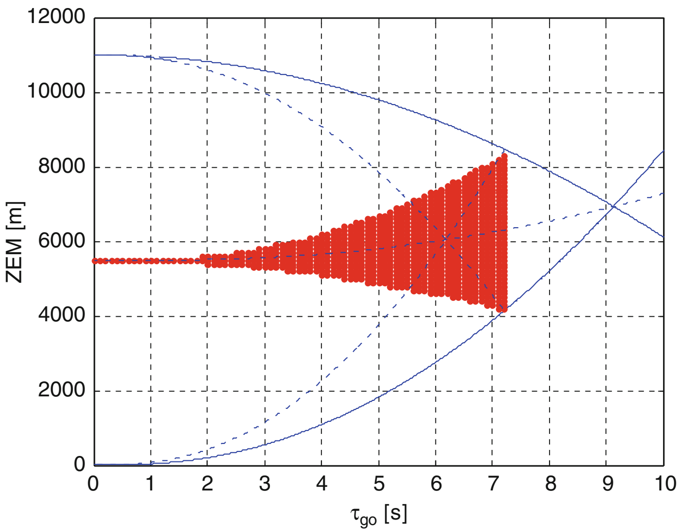


Fig. 6 Evader control saturation; the marks in the center of the figure correspond to initial conditions leading to $z_{if} \neq 0$, $z_{1f} = z_{2f}$ (no saturation of $v^*_{2\times 1}$)

For analysis purpose we defined a grid with potential 2 × 1 initial conditions, ran all the games P1 playing u^*_1, P2 playing u^*_2, E playing sat($v^*_{2\times 1}$) and summarized the results on Fig. 6. Figure 6 characterizes the saturations that apply in the evader optimal controls. Outside the 2×1 capture zone (always still considering E* between the two pursuers), three different end games can happen:

If $z_{1f} = 0$ and $z_{2f} = 0$, then the initial conditions belong to the 2×1 no-escape-zone (limit case $\tau = \tau_{2\times1}$).
If $z_{if} \neq 0$ and $z_{1f} = z_{2f}$, then the initial conditions belong to the 2×1 nonsaturated zone (area filled with marks in the center of Fig. 6).
If $z_{if} \neq 0$ and $z_{1f} \neq z_{2f}$, then the initial conditions belong to the 2×1 saturated zone (case corresponding to the "trade off" evader controls overcoming the control bounds).

This new limit (2×1 boundary), as well as the zone of $v^*_{2\times1}$ saturation, has been investigated in linear and nonlinear simulations to confirm the computation of $\tau_{2\times1}$. Up to now, the 2×1 game trajectories (Figs. 3, 5, 6) have been plotted on a single (ZEM, τ_{go}) representation. Figure 7 represents the barrier of the 2×1 game in the state space $(ZEM, \tau_{go}, \Delta P1P2)$; $ZEM = Z_2$; $\Delta P1P2 = Z_2 - Z_1$. On the left side delimited by the P2/E NEZ limits P2 intercepts E. On the right part (delimited by the surface build up P1/E NEZ limits) E is captured by P1. The capture zone extension due to the presence of two pursuers (P1 + P2 capture zone) is depicted by the set drawn in the closeness of the intersections between the P1/E surface and the P2/E surface. The P1 or P2 capture zone is far away behind the P1 + P2 zone. Moreover, in Fig. 7 we draw an optimal trajectory (black line with cross marks on it) starting (and remaining) on the new 2×1 boundary.

On Fig. 8, projections on the ZEM, τ_{go} plan of the Fig. 7 have been plotted at different $\Delta P1P2$ values.

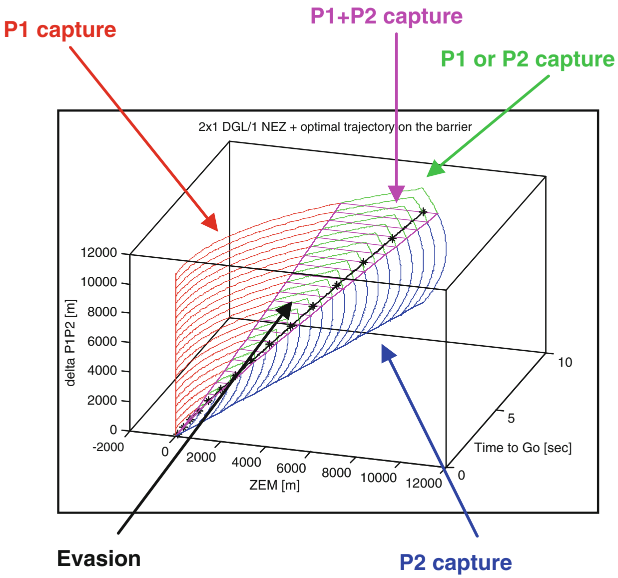


Fig. 7 2×1 NEZ

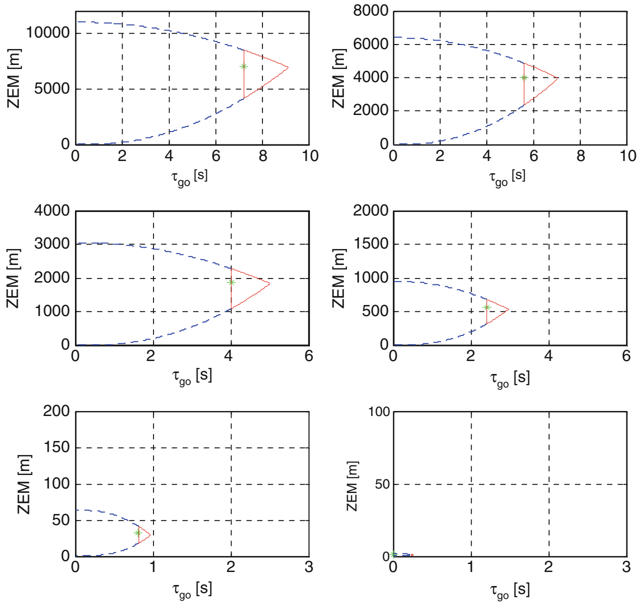


Fig. 8 Trajectory (cross mark) on the 2 × 1 NEZ

3.4 Case of Different Times-to-go

In this configuration, the two pursuers are launched at different times and therefore have different times-to-go. The previous expression of $v^*_{2\times1}$ is no longer valid and thus needs to be generalized to take into account the difference in times-to-go. Note that in this new case, the evader is assumed to switch its command to $v^*_{1\times1}$ (and faces only one pursuer like in a 1×1 game) when it goes beyond the first opponent. The optimal evader control, always called $v^*_{2\times1}$, should still lead to the equality of the final distances (in the absence of evader control saturations):

$$Z_1(\tau_1 = 0) = -Z_2(\tau_2 = 0)$$

$$z_1(0)\,\iota^2_{P1} = -z_2(0)\,\iota^2_{P2}.$$

Looking at Fig. 9, case where $\tau_2 < \tau_1$, $\Delta\tau = \tau_1 - \tau_2$, the normalized zero-effort-miss at final time can be now written:

$$z_1(0) = z_1(\tau_1) + z_{1\,\max}(\tau_1) + \Delta z_1(0)$$

$$z_2(0) = z_2(\tau_2) - z_{2\,\max}(\tau_2) + \Delta z_2(0)$$

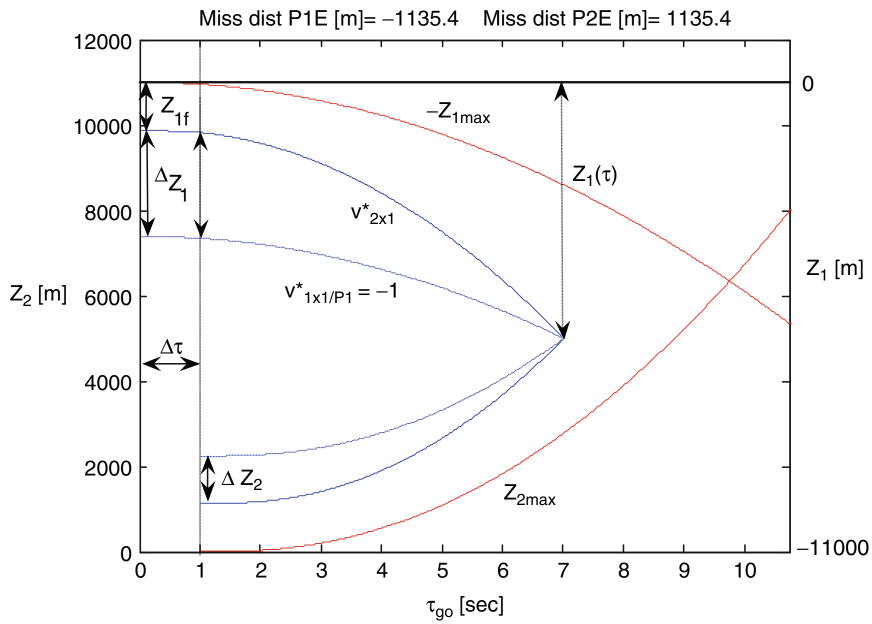


Fig. 9 Calculation of $v^*_{2\times1}$ with different time-to-go ($\tau_2 < \tau_1$)

However, $\Delta z_1(0) = \Delta z_1(\Delta\tau)$ because the evader plays only against P1 ($v^*_{1\times1} = -1$) during $\Delta\tau$ after crossing P2 (see Fig. 9). Moreover, we should notice that in Fig. 9 the x axis corresponds to τ_1 values ($\tau_1 = \tau_2 + \Delta\tau$). The third terms that corresponds to the new change in z_{1f} due to $v^*_{2\times1}$ is:

$$\Delta z_1(\Delta\tau) = \varepsilon_1 H_{21}(\Delta\tau \to \tau_1)\left(v^*_{2\times1} - v^*_{1\times1}\right)$$

and, thanks to the properties of integrals,

$$H_{21}(\Delta\tau \to \tau_1) = \int_{\Delta\tau}^{\theta_1} h\left(\frac{\xi}{\varepsilon}\right) \mathrm{d}\xi = H_{21}(\tau_1) - H_{21}(\Delta\tau)$$

the expression becomes:

$$\Delta z_1(\Delta\tau) = \left[\varepsilon_1 H_{21}(\tau_1) - \varepsilon_1 H_{21}(\Delta\tau)\right]\left(v^*_{2\times1} + 1\right).$$

Substituting these expressions leads to:

$$z_1(0) = z_1(\tau_1) + \mu_1 H_{11}(\tau_1) - \varepsilon_1 H_{21}(\Delta\tau) + \left[\varepsilon_1 H_{21}(\tau_1) - \varepsilon_1 H_{21}(\Delta\tau)\right] v^*_{2\times1}.$$

Moreover, the expression of z_{2f} remains the same as before when time-to-go coincide:

$$z_2(0) = z_2(\tau_2) - [\mu_2 H_{12}(\tau_2) - \varepsilon_2 H_{22}(\tau_2)] + \varepsilon_2 H_{22}(\tau_2)\left(v_{2\times 1}^* - 1\right)$$

from which, using the initial equality, we can isolate $v_{2\times 1}^*$, in the case where $\tau_2 < \tau_1$:

$$v_{2\times 1}^*(z) = \frac{[-\mu_1 H_{11}(\tau_1) - z_1(\tau_1) + \varepsilon_1 H_{21}(\Delta\tau)]\,\iota_{P1}^2 + [\mu_2 H_{12}(\tau_2) - z_2(\tau_2)]\,\iota_{P2}^2}{\varepsilon_2 H_{22}(\tau_2)\,t_{P2}^2 + [\varepsilon_1 H_{21}(\tau_1) - \varepsilon_1 H_{21}(\Delta\tau)]\,t_{P1}^2}$$

Due to symmetry, the same computation can be done switching the role of pursuers 1 and 2. Expression of $v_{2\times 1}^*$ in case where $\tau_2 > \tau_1$ is then as follows:

$$v_{2\times 1}^*(z) = \frac{[-\mu_1 H_{11}(\tau_1) - z_1(\tau_1)]\,\tau_{P1}^2 + [\mu_2 H_{12}(\tau_2) - \varepsilon_2 H_{22}(\Delta\tau) - z_2(\tau_2)]\,\iota_{P2}^2}{[\varepsilon_2 H_{22}(\tau_2) - \varepsilon_2 H_{22}(\Delta\tau)]\,\iota_{P2}^2 + \varepsilon_1 H_{21}(\tau_1)\,\iota_{P1}^2}$$

Note that the particular case where $\Delta\tau = 0$ leads to the expression of $v_{2\times 1}^*$ found in Sect. 3.2 (same time-to-go). By the way, we compute the 2×1 NEZ extension when times-to-go differ. Figure 10 shows $\tau_{2\times 1}$ limit remaining as a straight line (dot dash line in Fig. 10 tagged with an arrow mark).

Figure 11 describes an engagement initiated inside the capture zone.

In Fig. 12, the game starts outside the capture zone.

The crosses in Fig. 13 represent an optimal trajectory on the 2×1 NEZ boundary at different time instants when P1 and P2 times-to-go are not equal.

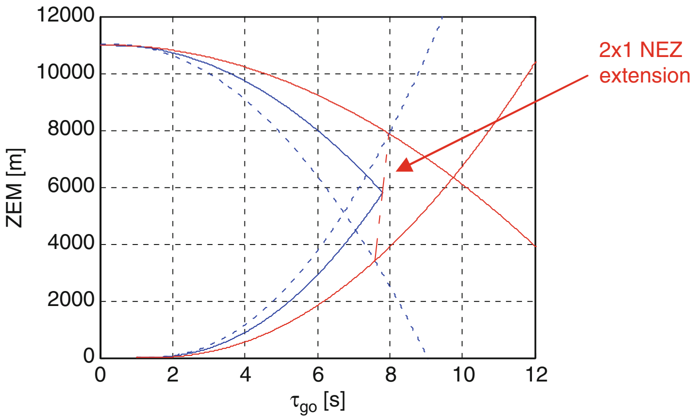


Fig. 10 2×1 NEZ for $\tau_1 - \tau_2 = 1$ sec

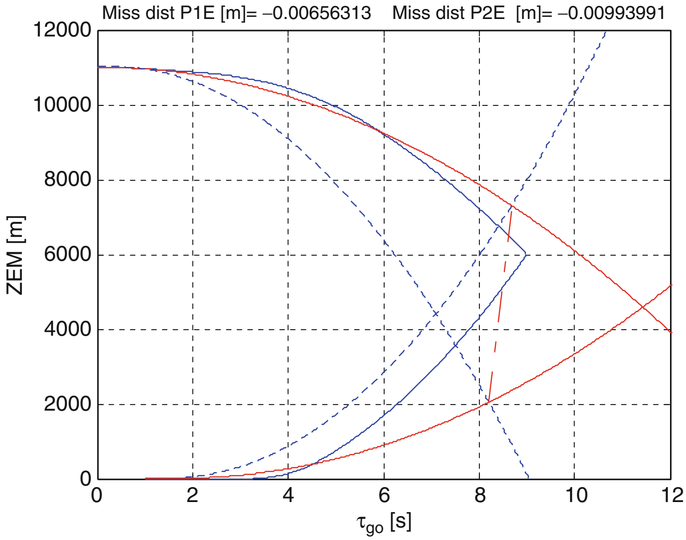


Fig. 11 Initial condition inside the NEZ: $\tau_0 = 9\,\mathrm{s}$, $y_0 = 6{,}000\,\mathrm{m}$

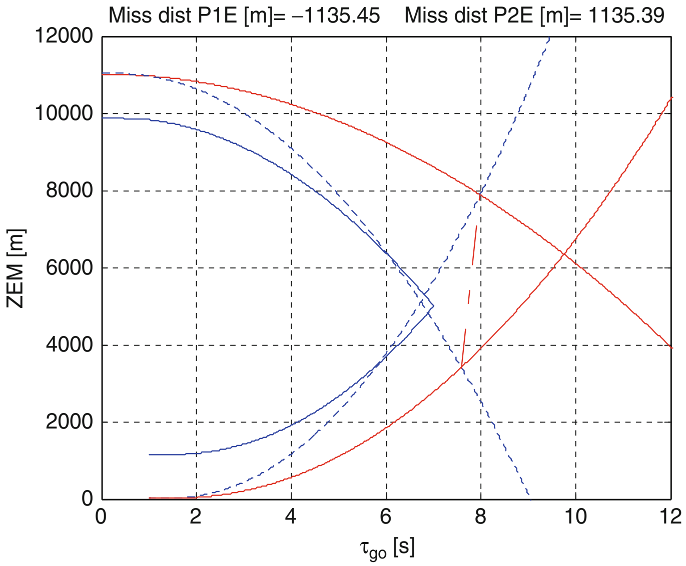


Fig. 12 Initial condition outside the NEZ: $\tau_0 = 9\,\mathrm{s}$, $y_0 = 6{,}000\,\mathrm{m}$

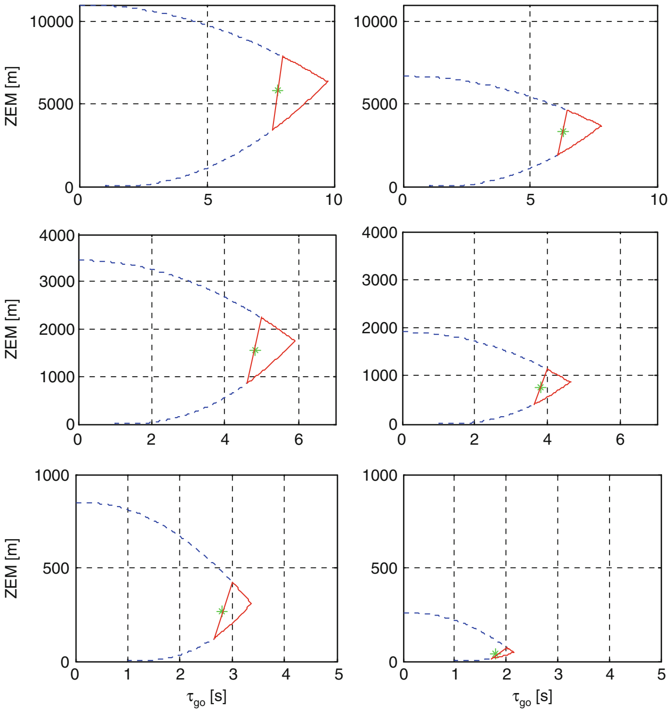


Fig. 13 2×1 NEZ sections at different time instants (*dot lines*: limits of the 1×1 NEZes, *solid lines*: 2×1 NEZ extension, *star mark*: state position)

3.5 Conclusion

We extend DGL/1 to the three players game involving two pursuers and one evader. We solve the game showing that for some initial conditions located between the pursuers the optimal evasion strategy is no longer a maximum turn control. The optimal evasion then consists in a "trade-off" strategy to drive the evader between the two pursuers. Considering 2×1 games we enlarge the interception area compared to solutions only involving two independent 1×1 NEZ. This approach could be applied when DGL/1 NEZ are bounded (closed) and when considering DGL/0 dynamics (linear differential game with first order for the pursuer and zero order for the evader).

The next step consists in transposing these results to 3D engagements with several players and realistic models, and to design (suboptimal) assignment strategies based on NEZ. From the interception point of view, we notice that the design of allocation strategies in N×P engagements require to solve 1×1 NEZ, 2×1 NEZ, but also many-on-one NEZ.

Moreover, it could be interesting to compare this solution approach to other algorithms addressing the problem of computing, approaching or over approximating reachable sets with nonlinear kinematics (level set methods [10], victory domains from viability theory [5]). In a general manner, other approaches dedicated to sub-optimal multiplayer strategies are relevant: reflection of forward reachable sets [9], minimization/maximization of the growth of particular level set functions [13, 14], Multiple Objective Optimization approach [10], LQ approach with evader terminal constrains and specific guidance law (Proportional Navigation) for the pursuers [11].

Appendix

For a time invariant linear system with two independent controls $u(t)$ and $v(t)$:

$$\dot{X}(t) = A \cdot X(t) + B \cdot u(t) + C \cdot v(t)$$

and a terminal cost function:

$$J = f(D.X(t_f)) = f(x_1(t_f)), \quad D = (1, 0, \ldots 0), \quad X = (x_1, \ldots)$$

the terminal projection transformation can be written as follows:

$$Z(t) = D \cdot \Phi(t_f, t) \cdot X(t)$$

with $\Phi(t_f, t)$ the transition matrix of the hereafter homogeneous systems:

$$\dot{\Phi}(t_f, t) = -\Phi(t_f, t) \cdot A; \Phi(t_f, t_f) = I.$$

$Z(t)$ or ZEM ("Zero Effort Miss") is a scalar variable given the terminal miss distance if at time t the controls u and v are turned into 0. Based on the following definitions:

$$\tilde{B}(t_f, t) = D \cdot \Phi(t_f, t) \cdot B; \quad \tilde{C}(t_f, t) = D \cdot \Phi(t_f, t) \cdot C$$

the derivative of Z with respect to time can be written as:

$$\dot{Z}(t) = \tilde{B}(t_f, t) \cdot u + \tilde{C}(t_f, t) \cdot v$$

with the following terminal cost:

$$J = f\left(Z(t_f)\right)$$

When considering the DGL/1 kinematics (first order transfer function for the pursuer and for the evader):

$$X^T = \left[\, y,\ \dot{y},\ \ddot{y}_P,\ \ddot{y}_E \,\right]$$

$$\dot{X}(t) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & \frac{-1}{\iota_P} & 0 \\ 0 & 0 & 0 & \frac{-1}{\iota_E} \end{bmatrix} X(t) + \begin{bmatrix} 0 \\ 0 \\ \frac{a_{P\max}}{\iota_P} \\ 0 \end{bmatrix} u(t) + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{a_{E\max}}{\iota_E} \end{bmatrix} v(t),$$

where y is the relative perpendicular miss, ι_P and ι_E are time lag constants, $a_{P\max}$ and $a_{E\max}$ are, respectively, the Pursuer and the Evader control bounds and $y(t)$is the perpendicular miss with respect to the initial Line Of Sight. According to the terminal projection procedure, the initial linear system is equivalent to the representation:

$$\Phi_{1,1\ldots4}(t_f,t) = \left[1,\ \tau,\ -\iota_P^2 h(\theta),\ \iota_E^2 h\left(\tfrac{\theta}{\varepsilon}\right)\right]$$

$$\varepsilon = \frac{\iota_E}{\iota_P}, \quad \theta = \frac{\tau}{\iota_P}, \quad \tau = t_f - t, \quad h(\alpha) = \mathrm{e}^{-\alpha} + \alpha - 1$$

$$Z(\tau) = y + \dot{y}\tau - \ddot{y}_P \iota_P^2 h(\theta) + \ddot{y}_E \iota_E^2 h\left(\frac{\theta}{\varepsilon}\right).$$

According to the terminal projection procedure, we study the four-dimension DGL/1 model in a one dimension frame.

$$\tilde{B}(t_f,t) = -a_{P\max}\iota_P h(\theta), \qquad \tilde{C}(t_f,t) = a_{E\max}\iota_E h\left(\frac{\theta}{\varepsilon}\right)$$

$$\dot{Z}(\tau) = -\frac{1}{\tau_P}\frac{\mathrm{d}Z}{\mathrm{d}\theta}(\theta) = \iota_P a_{E\max}\left[\varepsilon h\left(\frac{\theta}{\varepsilon}\right) v - \mu h(\theta)\, u\right]$$

$$\mu = \frac{a_{P\max}}{a_{E\max}}, \qquad z = \frac{Z}{\iota_P^2 a_{E\max}}$$

$$\frac{\mathrm{d}z}{\mathrm{d}\theta}(\theta) = \mu h(\theta) u - \varepsilon h\left(\frac{\theta}{\varepsilon}\right) v$$

Moreover, for the sake of simplicity, as we normalized the time to go variable τ into θ, we also normalize the state variable Z into z.

Then, using the Hamilton Jacobi Bellman Isaacs equations it is easy to show that:

$$u^*(t) = v^*(t) = signe\left[Z(t_f)\right], \quad Z(t_f) \neq 0,$$

but also:

$$u^*(t) = v^*(t) = signe\left[Z(t)\right], \quad Z(t_f) \neq 0.$$

The structure of the DGL/1 solutions (in terms of optimal controls and optimal trajectories) explains the specific properties we obtain when studying the 2×1 game.

References

1. Basar, T., Olsder, G.J.: Dynamic Non Cooperative Game Theory. Academic, London and San Diego (1982)
2. Ben-Asher, J., Cliff, E.M., Kelly, H.J.: Optimal Evasion with a Path-Angle Constraint and Against Two Pursuers. Journal of Guidance, Control and Dynamics **11**(4), (1988)
3. Bryson, A.E., Ho, Y.C.: Applied Optimal Control. Hemisphere Publishing Corporation, USA (1975)
4. Cardaliaguet, P.: A differential game with two players and one target. SIAM J. Control Optimization **34**(4), 1441–1460 (1996)
5. Cardaliaguet, P., Quincampoix, M., Saint-Pierre, P.: Set-valued numerical analysis for optimal control and differential games. In: Bardi, M., Raghavan, T.E.S., Parthasarathy, T. (eds) Annals of the International Society of Dynamic Games, Vol. 4, pp. 177–247. Birkhäuser (1999)
6. Ge, J., Tang, L., Reimann, J., Vachtsevanos, G.: Suboptimal Approaches to Multiplayer Pursuit-Evasion Differential Games. AIAA 2006-6786 Guidance, Navigation, and Control Conference, 21–24 August 2006, Keystone, Colorado
7. Hagedorn, P., Breakwell, J.V.: A Differential Game with Two Pursuers and One Evader. Journal of Optimization Theory and Application **18**(2), 15–29 (1976)
8. Isaacs, R.: Differential Games. Wiley, New York (1967)
9. Jang, J.S., Tomlin, C.J.: Control Strategies in Multi-Player Pursuit and Evasion Game. AIAA 2005-6239 Guidance, Navigation, and Control Conference, 15–18 August 2005, San Francisco, California
10. Mitchel, M., Bayen, A., Tomlin, C.J.: A Time Dependent Hamilton-Jacobi Formulation of Reachable Sets for Continuous Dynamic Games. IEEE Transactions on Automatic Control **50**(7) (2005)
11. Rusnak, I.: The Lady, The Bandits, and The Bodyguards – A Two Team Dynamic Game. Proceedings of the 16th World IFAC Congress (2005)
12. Shinar, J., Shima, T.: Nonorthodox Guidance Law Development Approach for Intercepting Maneuvering Targets. Journal of Guidance, Control and Dynamics **25**(4) (2002)
13. Stipanovic, D.M., Melikyan, A., Hovakimyan, N.: Some Sufficient Conditions for Multi-Player Pursuit-Evasion Games with Continuous and Discrete Observations. 12th ISDG Conference, Sophia-Antipolis, France (2006)
14. Stipanovic, D.M., Sriram, S., Tomlin, C.J.: Strategies for Agents in Multi-Player Pursuit-Evasion Games. 11th ISDG Conference, Tucson, Arizona (2004)

Homicidal Chauffeur Game: History and Modern Studies

Valerii S. Patsko and Varvara L. Turova

Abstract "Homicidal chauffeur" game is one of the most well-known model problems in the theory of differential games. "A car" striving as soon as possible to run over "a pedestrian" – this demonstrative model suggested by R. Isaacs turned out to be appropriate for many applied problems. No less remarkable is the fact that the game is a difficult and interesting object for mathematical investigation. This chapter gives a survey of publications on the homicidal chauffeur problem and its modifications.

Keywords Backward procedures · Differential games · Homicidal chauffeur game · Numerical constructions · Value function

1 Introduction

"Homicidal chauffeur" game was suggested and described by Rufus Philip Isaacs in the report [13] for the RAND Corporation in 1951. A detailed description of the problem was given in his book "Differential games" published in 1965. In this problem, a "car" whose radius of turn is bounded from below and the magnitude of the linear velocity is constant pursues a noninertia "pedestrian" whose velocity does not exceed some given value. The names "car", "pedestrian", and "homicidal chauffeur" turned out to be very suitable, even if real objects that R. Isaacs meant [7, p. 543] were a guided torpedo and an evading small ship.

V.S. Patsko (✉)
Institute of Mathematics and Mechanics, Ural Branch of Russian Academy of Sciences,
S. Kovalevskaya str. 16, Ekaterinburg, Russia
e-mail: patsko@imm.uran.ru

V.L. Turova
Technische Universität München, Boltzmannstr. 3, 85747 Garching bei München,
Germany
e-mail: turova@ma.tum.de

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_12,

The attractiveness of the game is connected not only with its clear applied interpretation, but also with the possibility of transition to reference coordinates, which enables to deal with a two-dimensional state vector. In the reference coordinates, we obtain a differential game in the plane. Due to this, the analysis of the geometry of optimal trajectories and singular lines that disperse, join, or refract optimal paths becomes more transparent.

The investigation started by R. Isaacs was continued by John Valentine Breakwell and Antony Willits Merz. They improved Isaacs' method for solving differential games and revealed new types of singular lines for problems in the plane. A systematic description of the solution structure for the homicidal chauffeur game depending on the parameters of the problem is presented in the PhD thesis by A. Merz supervised by J. Breakwell at Stanford University. The work performed by A. Merz seems to be fantastic, and his thesis, to our opinion, is the best research among those devoted to concrete model game problems.

Our chapter is an appreciation of the invaluable contribution made by the three outstanding scientists: R. Isaacs, J. Breakwell, and A. Merz to the differential game theory. Thanks to the help of Ellen Sara Isaacs, John Alexander Breakwell, and Antony Willits Merz, we have an opportunity to present formerly unpublished photographs (Figs. 1–6).

The significance of the homicidal chauffeur game is also that it stimulated the appearance of other problems with the same dynamic equations as in the classic

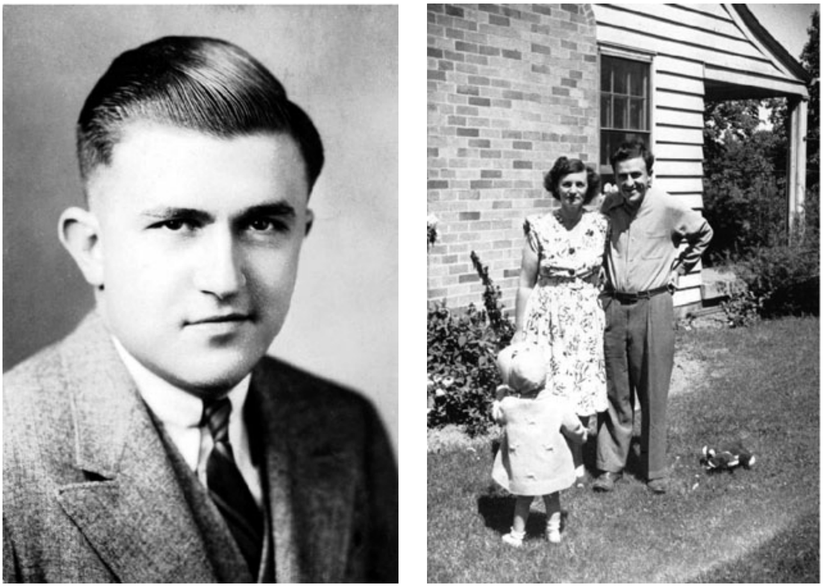


Fig. 1 *Left picture*: Rufus Isaacs (about 1932–1936). *Right picture*: Rose and Rufus Isaacs with their daughter Ellen in Hartford, Connecticut before Isaacs went to Notre Dame University in about 1945

Fig. 2 Rose and Rufus Isaacs embarking on a cruise in their 40s or 50s

Fig. 3 Rufus Isaacs at his retirement party, 1979

statement, but with different objectives of the players. The most famous among them is the surveillance-evasion problem considered in papers by John Breakwell, Joseph Lewin, and Geert Jan Olsder.

Fig. 4 John Breakwell at a Stanford graduation

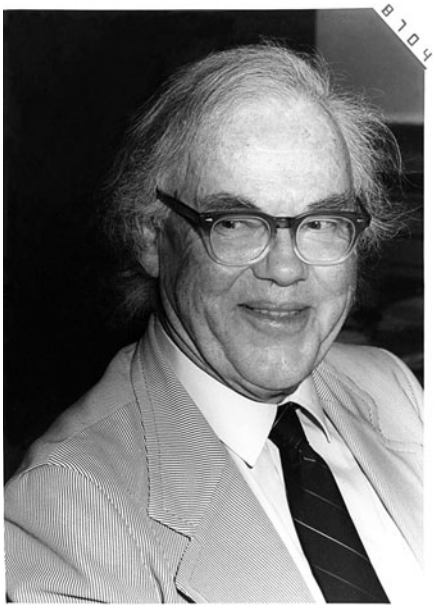


Fig. 5 John Breakwell (April 1987)

A very interesting variant of the homicidal chauffeur game is investigated in the papers by Pierre Cardaliaguet, Marc Quincampoix, and Patrick Saint-Pierre. The objectives of the players are the usual ones, whereas the constraint on the control of the evader depends on the distance between him and pursuer.

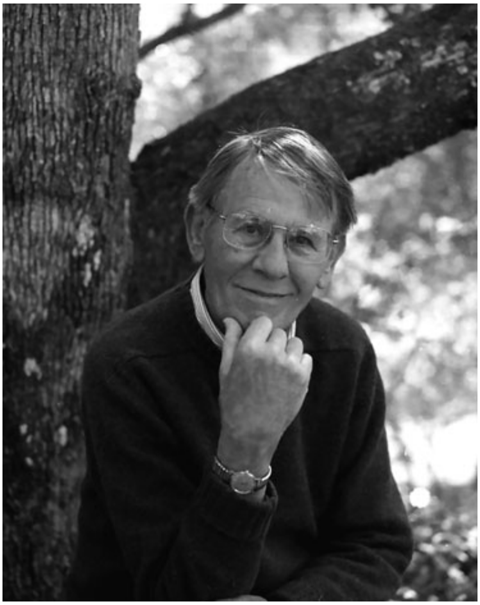


Fig. 6 Antony Merz (March 2008)

We also consider a statement where the pursuer is reinforced: he becomes more agile.

The description of the above-mentioned problems is accompanied by the presentation of numerical results for the computation of level sets of the value function using an algorithm developed by the authors. The algorithm is based on the approach for solving differential games worked out in the scientific school of Nikolai Nikolaevich Krasovskii (Ekaterinburg).

In the last section of the chapter, we mention some works using the homicidal chauffeur game as a test example for computational methods. Also, the two-target homicidal chauffeur game is noted as a very interesting numerical problem.

2 Classic Statement by R. Isaacs

Denote the players by the letters P and E. The dynamics read

$$\begin{aligned} P : \dot{x}_p &= w \sin\theta & \qquad E : \dot{x}_e &= v_1 \\ \dot{y}_p &= w \cos\theta & \dot{y}_e &= v_2 \\ \dot{\theta} &= wu/R, \ \ |u| \le 1 & v &= (v_1, v_2)', \ \ |v| \le \rho. \end{aligned}$$

Here, w is the magnitude of linear velocity, R is the minimum radius of turn. By normalizing the time and geometric coordinates, one can achieve that $w = 1$, $R = 1$. As a result, in the dimensionless coordinates, the dynamics have the form

$$\begin{aligned} P : \dot{x}_p &= \sin\theta & E : \dot{x}_e &= v_1 \\ \dot{y}_p &= \cos\theta & \dot{y}_e &= v_2 \\ \dot{\theta} &= u, \ \ |u| \le 1 & v &= (v_1, v_2)', \ |v| \le \nu. \end{aligned} \tag{1}$$

Choosing the origin of the reference system at the position of player P and directing the y-axis along P's velocity vector, one arrives [14] at the following system

$$\dot{x} = -yu + v_x, \quad \dot{y} = xu - 1 + v_y;$$
$$|u| \le 1, \ v = (v_x, v_y)', \ |v| \le \nu.$$

The objective of player P having control u at his disposal is, as soon as possible, to bring the state vector to the target set M being a circle of radius r with the center at the origin. The second player which steers using control v strives to prevent this. The controls are constructed based on a feedback law.

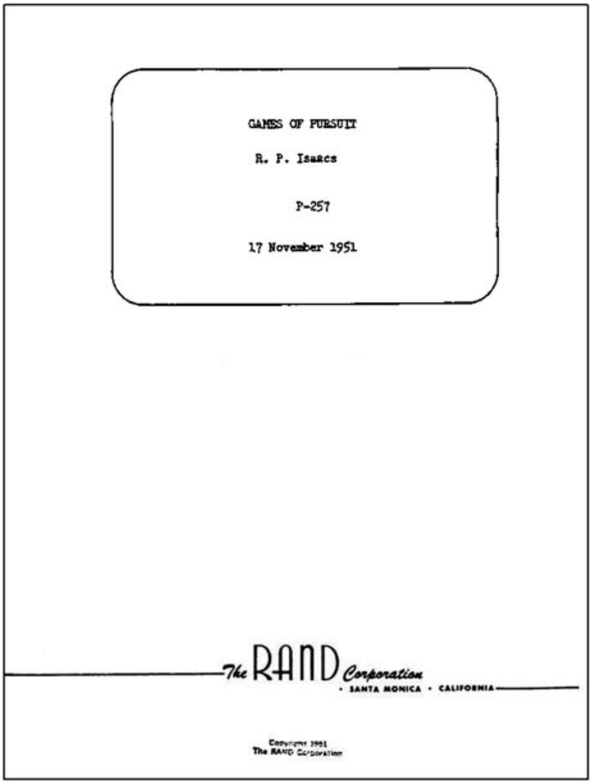
One can see that the description of the problem contains two independent parameters ν and r.

R. Isaacs investigated the problem for some parameter values using his method for solving differential games. The basis of the method is the backward computation of characteristics for an appropriate partial differential equation. First, some primary region is filled out with regular characteristics, then a secondary region is filled out, and so on. The final characteristics in the plane of state variables coincide with optimal trajectories.

As it was noted, the homicidal chauffeur game was first described by R. Isaacs in his report of 1951. The title page of this report is given in Fig. 7.

Figure 8a shows a drawing from the book [14] by R. Isaacs. The solution is symmetric with respect to the vertical axis. The upper part of the vertical axis is a singular line. Forward time optimal trajectories meet this line at some angle and then go along it toward the target set M. According to the terminology by R. Isaacs, the line is called universal. The part of the vertical axis adjoining the target set from below is also a universal singular line. Optimal trajectories go down along it. The rest of the vertical axis below this universal part is dispersal: two optimal paths emanate from every point of it. On the barrier line $\mathscr{B}$, the value function is discontinuous. The side of the barrier line where the value of the game is smaller will be called positive. The opposite side is negative.

The equivocal singular line emanates tangentially from the terminal point of the barrier (Fig. 8b). It separates two regular regions. Optimal trajectories that come to the equivocal curve split into two paths: the first one goes along the curve, and the second one leaves it and comes to the regular region on the right (optimal trajectories in this region are shown in Fig. 8a).



GAMES OF PURSUIT

R. P. Isaacs

P-257

17 November 1951

The RAND Corporation

• SANTA MONICA • CALIFORNIA

Copyright 1951
The RAND Corporation

Fig. 7 Title page of the first report [13] by R. Isaacs for the RAND Corporation

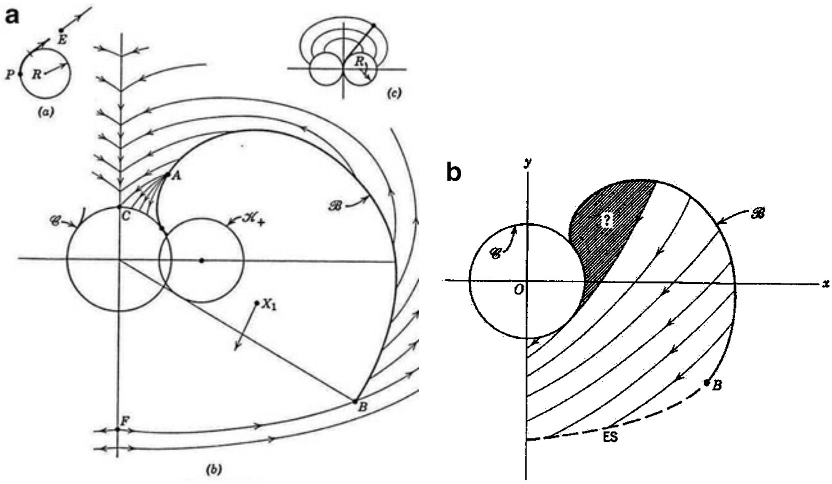


Fig. 8 Pictures by R. Isaacs from [14] explaining the solution to the homicidal chauffeur game. The solution is symmetric with respect to the vertical axis. **(a)** Solution structure in the primary region. **(b)** Solution in the secondary region

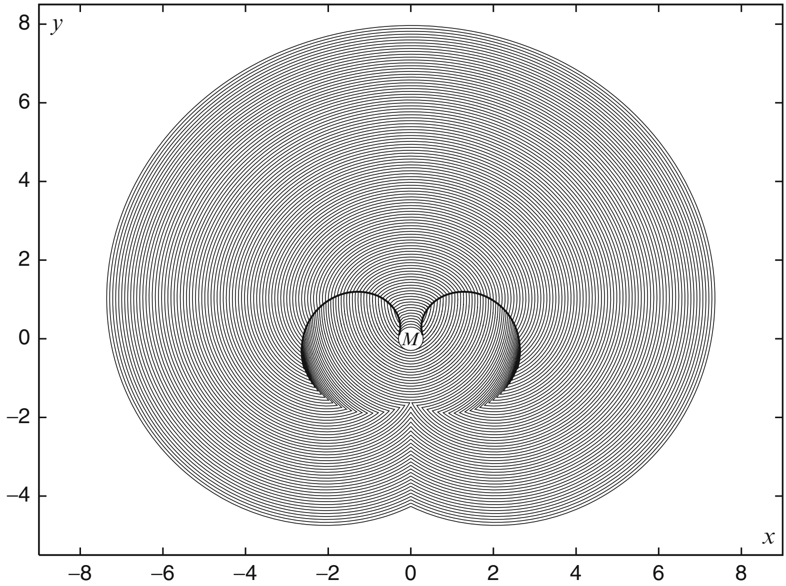


Fig. 9 Level sets of the value function for the classical problem; game parameters $\nu = 0.3$ and $r = 0.3$; backward computation is done till the time $\tau_f = 10.3$ with the time step $\Delta = 0.01$, output step for fronts $\delta = 0.1$

The equivocal curve is described through a differential equation which cannot be integrated explicitly. Therefore, any explicit description of the value function in the region between the equivocal and barrier lines is absent. The most difficult for the investigation is the "rear" part (Fig. 8b, shaded domain) denoted by R. Isaacs with a question mark. He could not obtain a solution for this domain.

Figure 9 shows level sets $W(\tau) = \{(x, y) : V(x, y) \leq \tau\}$ of the value function $V(x, y)$ for $\nu = 0.3, r = 0.3$. The numerical results presented in Fig. 9 and in subsequent figures are obtained using the algorithm proposed in [29]. The lines on the boundary of the sets $W(\tau)$, $\tau > 0$, consisting of points (x, y), where the equality $V(x, y) = \tau$ holds, will be called fronts (isochrones). Backward construction of the fronts, beginning from the boundary of the target set, constitutes the basis of the algorithm. A special computer program for the visualization of graphs of the value function in time-optimal differential games has been developed by Vladimir Lazarevich Averbukh and Oleg Aleksandrovich Pykhteev [2].

The computation for Fig. 9 is done with the time-step $\Delta = 0.01$ till the time $\tau_f = 10.3$. The output step for fronts is $\delta = 0.1$. Figure 10 presents the graph of the value function. The value function is discontinuous on the barrier line and on a part of the boundary of the target set. In the case considered, the value function is smooth in the above-mentioned rear region.

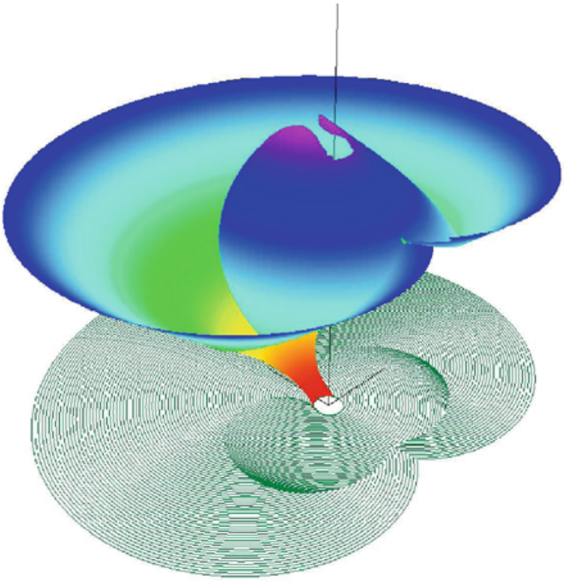


Fig. 10 Graph of the value function; $\nu = 0.3$, $r = 0.3$

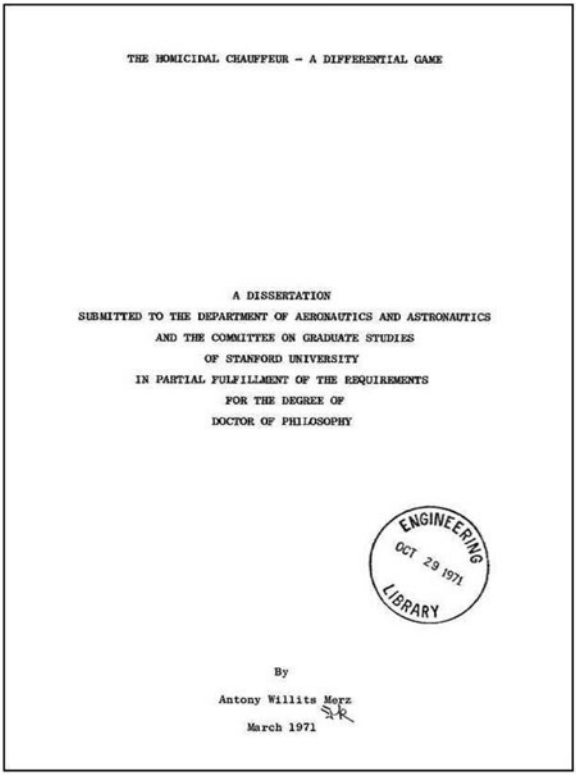
3 Investigations by J. Breakwell and A. Merz

J. Breakwell and A. Merz continued the investigation of the homicidal chauffeur game in the setting by R. Isaacs. Their results are partly and very briefly described in the papers [6, 23]. A complete solution is obtained by A. Merz in his PhD thesis [22]. The title page of the thesis is shown in Fig. 11.

A. Merz divided the two-dimensional parameter space into 20 subregions. He investigated the qualitative structure of optimal paths and the type of singular lines for every subregion. All types of singular curves (dispersal, universal, equivocal, and switch lines) described by R. Isaacs for differential games in the plane appear in the homicidal chauffeur game for certain parameter values. In the thesis, A. Merz suggested to distinguish some subtypes of singular lines and consider them separately. Namely, he introduced the notion of focal singular lines which are universal ones, but with tangential approach of optimal paths. The value function is nondifferentiable on the focal lines.

Figure 12 presents a picture and a table from the thesis by A. Merz that demonstrate the partition of two-dimensional parameter space into subregions with certain systems of singular lines (A. Merz used symbols γ, β to denote parameters, and he called singular lines exceptional lines).

The thesis contains many pictures explaining the type of singular lines and the structure of optimal paths. By studying them, one can easily detect tendencies in the behavior of the solution depending on the change of the parameters.

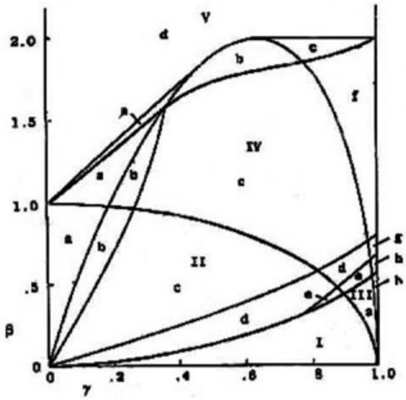


THE HOMICIDAL CHAUFFEUR – A DIFFERENTIAL GAME

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF AERONAUTICS AND ASTRONAUTICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

By
Antony Willits Merz
March 1971

Fig. 11 The title page of the PhD thesis by A. Merz



DISTRIBUTION OF EXCEPTIONAL LINES

Region	Sub-region	Exceptional Lines Present in Each Subregion											
		S	UL	PDL_y	PDL	EDL_a	EDL_c	$1D^+$	$1D^-$	EL	SL	SE	FL
I		x	x										
II	a	x	x	x					x		x		
	b	x	x	x					x	x	x		
	c	x	x	x					x	x			
	d	x	x	x					x	x		x	
	e	x	x	x					x	x		x	x
III	a	x	x			x		x					
	b	x	x				x	x					
IV	a		x	x	x			x	x		x		
	b		x	x	x			x	x	x	x		
	c		x	x	x	x		x	x	x			
	d		x	x	x	x		x	x	x		x	
	e		x	x	x	x		x	x	x		x	x
	f		x	x	x		x	x	x	x			
	g		x	x	x		x	x	x	x		x	
	h		x	x	x		x	x	x	x		x	x
V	a		x	x				x					
	b		x	x		x		x					
	c		x	x			x	x					
	d		x	x									

Fig. 12 Decomposition of two-dimensional parameter space into subregions

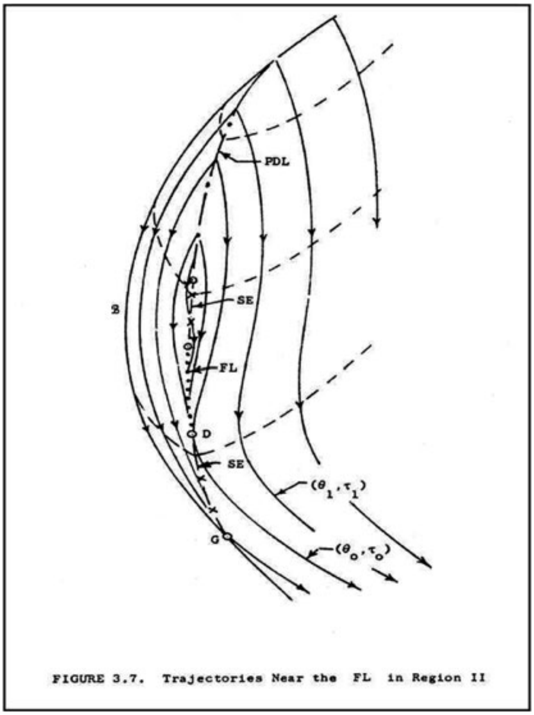


Fig. 13 Structure of optimal paths in the rear part for subregion IIe

In Fig. 13, the structure of optimal paths in that part of the plane that adjoins the negative side of the barrier is shown for the parameters corresponding to subregion IIe. This is the rear part denoted by R. Isaacs with a question mark. For subregion IIe, the situation is very complex.

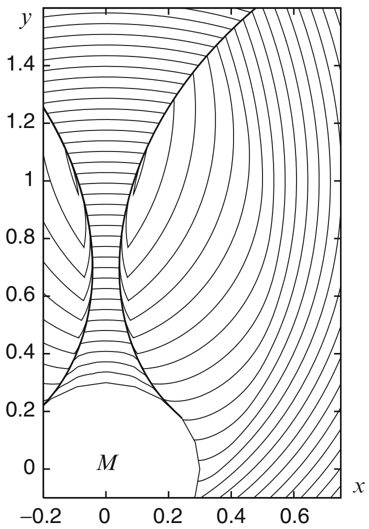
Symbol PDL denotes the dispersal line controlled by player P. Two optimal trajectories emanate from every point of this line. Player P controls the choice of the side to which trajectories come down. Singular curve SE (the switch envelope) is specified as follows. Optimal trajectories approach it tangentially. Then one trajectory goes along this curve, and the other (equivalent) one leaves it at some angle. Therefore, line SE is similar to an equivocal singular line. The thesis contains arguments according to which the switch envelope should be better considered as an individual type of singular line.

Symbol FL denotes the focal line. The dotted curves mark boundaries of level sets (in other words, isochrones or fronts) of the value function.

The value function is not differentiable on the line composed of the curves PDL, SE, FL, and SE.

The authors of this chapter undertook many efforts to compute value functions for parameters from subregion IIe. But this was not successful, because corner points that must be present on fronts to the negative side of the barrier were absent.

Fig. 14 Level sets of the value function for parameters from subregion IId; $\nu = 0.7$, $r = 0.3$; $\tau_f = 35.94$, $\Delta = 0.006$, $\delta = 0.12$



One of the possible explanations to this failure can be the following: the effect is so subtle that it cannot be detected even for very fine discretizations. The computation of level sets of the value function for the subregions where the solution structure changes very rapidly with varying parameters can be considered as a challenge for differential game numerical methods being presently developed by different scientific teams.

Figure 14 shows computational results for the case where fronts have corner points in the rear domain. However, the values of parameters correspond not to subregion IIe but to subregion IId. In that case, singular curve SE remains, but focal line FL disappears.

For some subregions of parameters, barrier lines on which the value function is discontinuous disappear. A. Merz described a very interesting transformation of the barrier line into two, close to each other, dispersal curves of players P and E. In this case, there exist optimal paths that both go up and down along the boundary of the target set. The investigation of such a phenomenon is of great theoretical interest.

Figure 15a presents a picture from the thesis by A. Merz that corresponds to subregion IVc (A. Merz as well as R. Isaacs used the symbol φ to denote the control of player P, denoted u in this chapter). Numerically constructed level sets of the value function are shown in Fig. 15b. When examining Fig. 15b, it might seem that some barrier line exists, but this is not true. We have exactly the case like the one shown in Fig. 15a.

Enlarged fragments of numerical constructions are given in Figs. 16, 17 (the scale of y-axis in Fig. 17 is increased with respect to that of x-axis). The curve consisting of fronts' corner points above the accumulation region of fronts is the dispersal line of player E. The curve composed of corner points below the accumulation

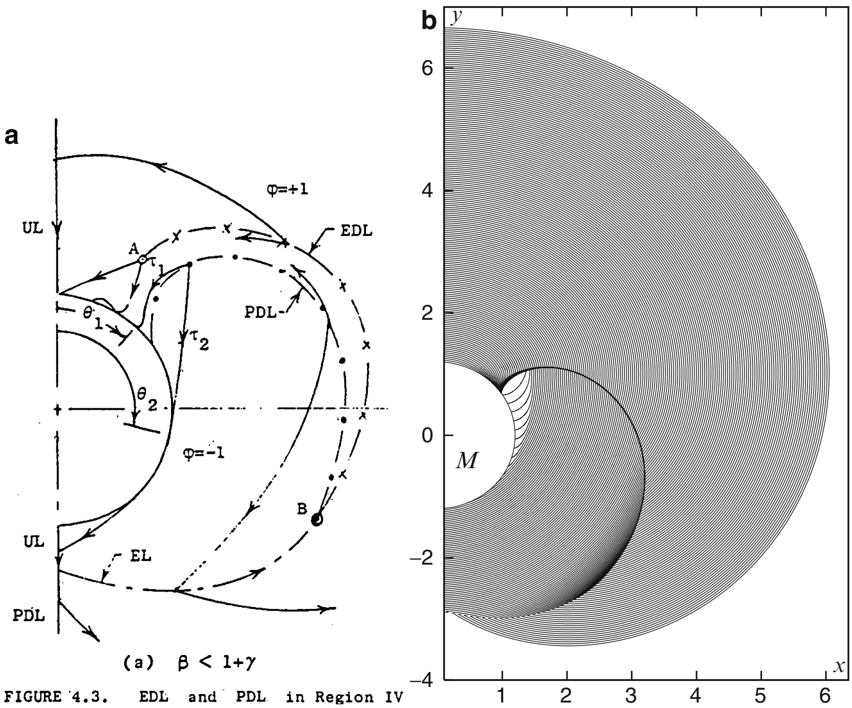


Fig. 15 (**a**) Structure of optimal trajectories in subregion IVc. (**b**) Level sets of the value function; $\nu = 0.7$, $r = 1.2$; $\tau_f = 24.22$, $\Delta = 0.005$, $\delta = 0.1$

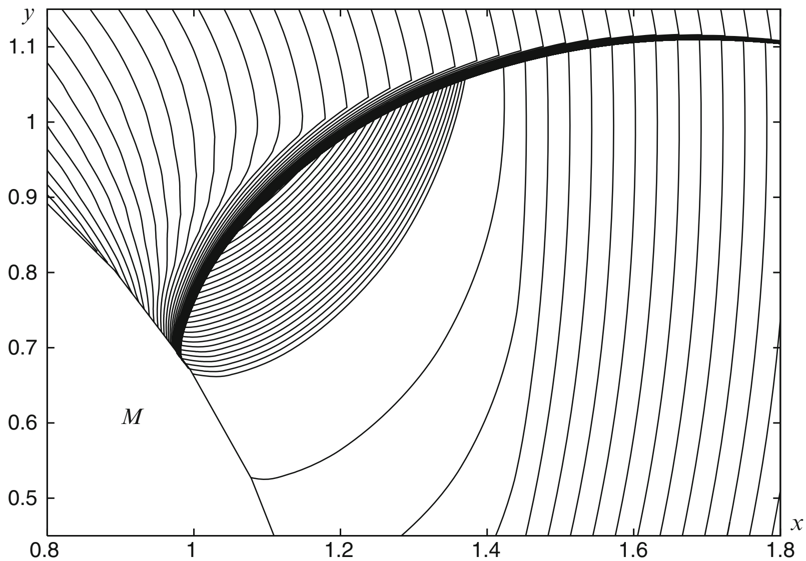


Fig. 16 Enlarged fragment of Fig. 15b; $\tau_f = 24.22$. Output step for fronts close to the time τ_f is decreased up to $\delta = 0.005$

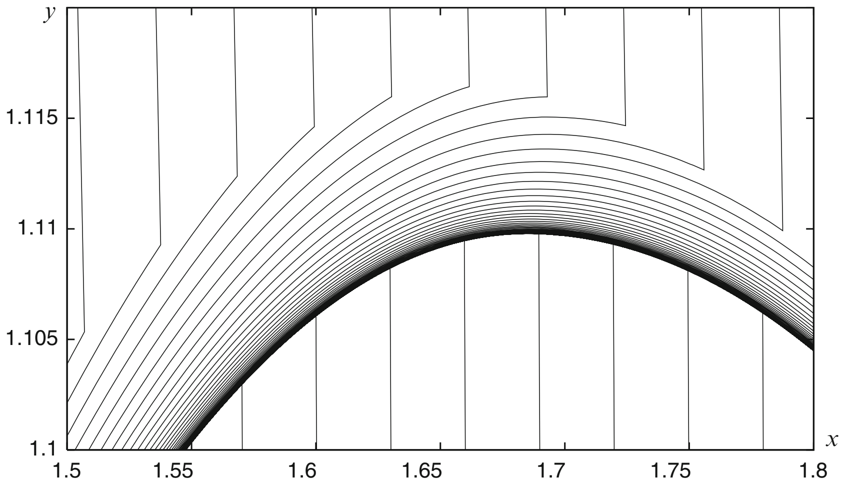


Fig. 17 Enlarged fragment of Fig. 15b

region is the dispersal line of player P. The value function is continuous in the accumulation region. To see where (in the considered part of the plane) the point with a maximum value of the game is located, additional fronts are shown. The point with the maximum value has coordinates $x = 1.1$, $y = 0.92$. The value function at this point is equal to 24.22.

4 Surveillance-Evasion Game

In the PhD thesis by J. Lewin [18] (performed as well under the supervision of J. Breakwell), in the joint paper by J. Breakwell and J. Lewin [19], and also in the paper by J. Lewin and G. J. Olsder [20], both dynamics and constraints on the controls of the players are the same as in Isaacs' setting, but the objectives of the players differ from those in the classic statement. Namely, player E tries to decrease the time to reach the target set M, whereas player P strives to increase that time. In [18] and [19], the target set is the complement (with respect to the plane) of an open circle centered at the origin. In [20], the target set is the complement of an open cone with the apex at the origin.

The meaning related to the original context concerning two moving vehicles is the following: player E tries, as soon as possible, to escape from some detection zone attached to the geometric position of player P, whereas player P strives to keep his opponent in the detection zone as long as possible. Such a problem was called the surveillance-evasion game. To solve it, J. Breakwell, J. Lewin, and G. J. Olsder used Isaacs' method.

One picture from the thesis by J. Lewin is shown in Fig. 18a, and one picture from the paper by J. Lewin and G. J. Olsder is given in Fig. 18b.

In the surveillance-evasion game with the conic target set, examples of transition from finite values of the game to infinite values are of interest and can be easily constructed.

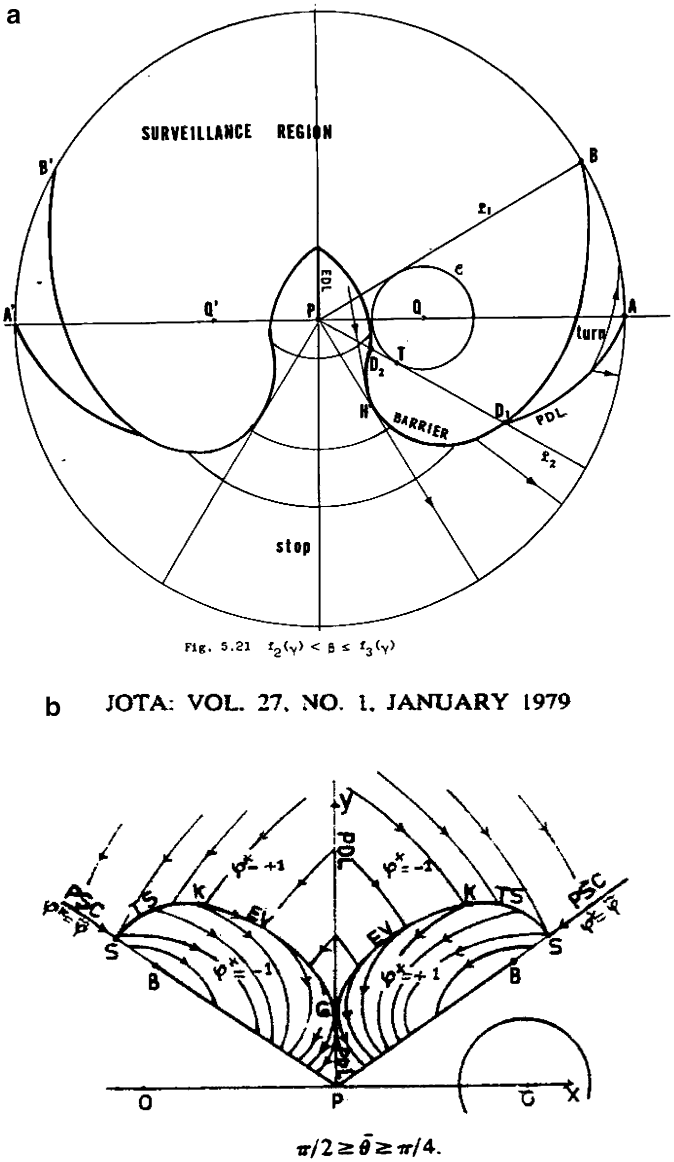


Fig. 18 (**a**) Picture from the PhD thesis by J. Lewin. Detection zone is a circle. (**b**) Picture from the paper by J. Lewin and G. J. Olsder. Detection zone is a convex cone

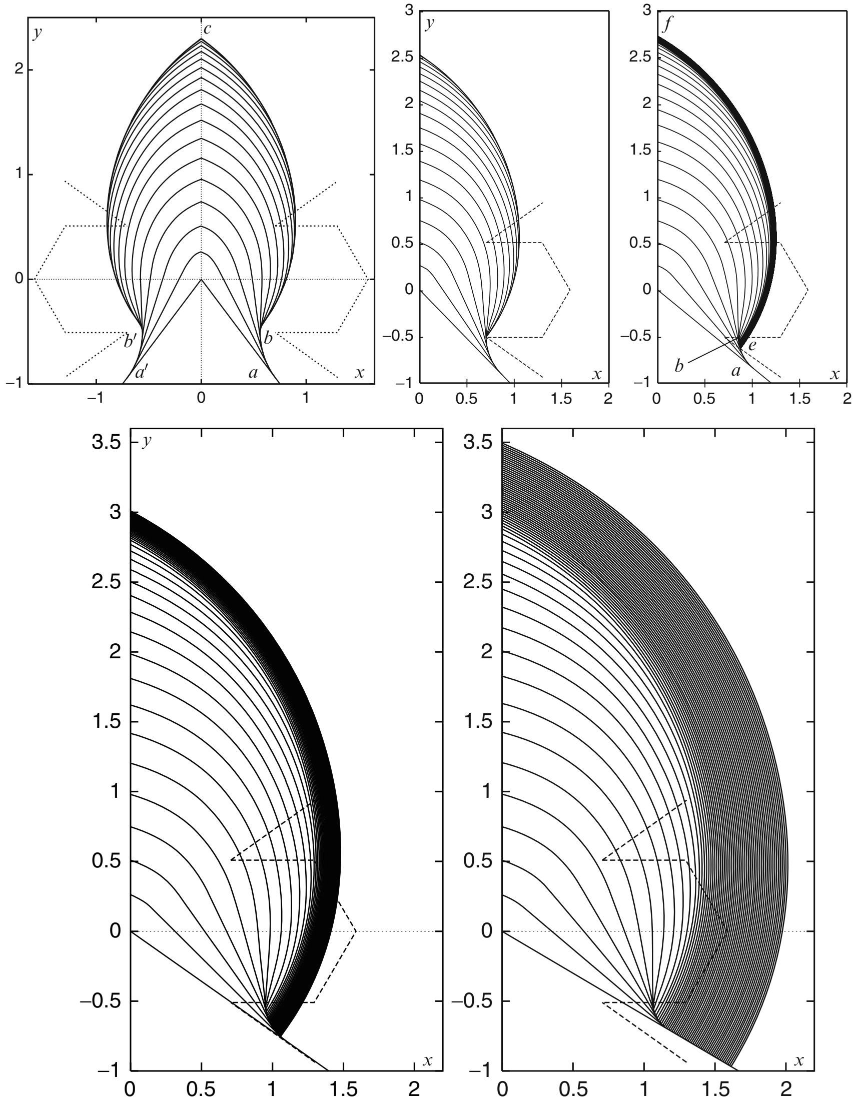


Fig. 19 Surveillance-evasion game. Change of the front structure depending on the semi-angle α of the nonconvex detection cone; $\nu = 0.588$, $\Delta = 0.017$, $\delta = 0.17$

Figure 19 shows level sets of the value function for five values of parameter α which specifies the semi-angle of the nonconvex conic detection zone. Since the solution to the problem is symmetric with respect to y-axis, only the right half-plane is shown for four of five figures. The pictures are ordered from greater to smaller α.

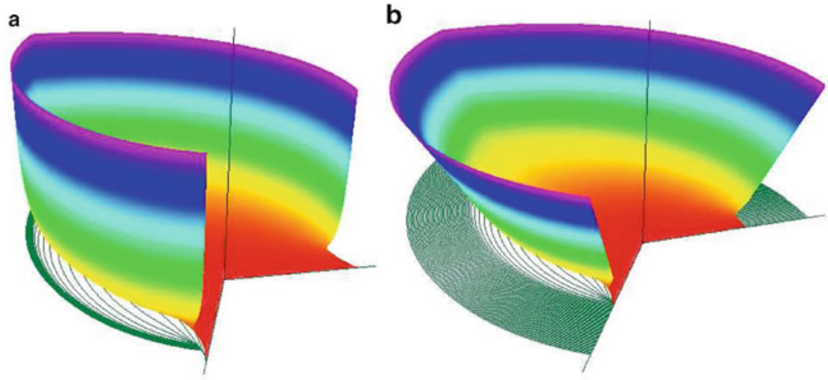


Fig. 20 Value function in the surveillance-evasion game. (**a**) $\nu = 0.588$, $\alpha = 130°$. (**b**) $\nu = 0.588$, $\alpha = 121°$

In the first picture, the value function is finite in the set that adjoins the target cone and is bounded by the curve $a'b'cba$. This set is filled out with the fronts (isochrones). The value function is zero within the target set. Outside the union of the target set and the set filled out with the fronts, the value function is infinite. In the third picture, a situation of the accumulation of fronts is presented. Here, the value function is infinite on the line fe and finite on the arc ea. The value function has a finite discontinuity on the arc be. The graph of the value function corresponding to the third picture is shown in Fig. 20a.

The second picture demonstrates a transition case from the first to the third picture.

In the fifth picture, the fronts propagate slowly to the right and fill out (outside the target set) the right half-plane as the backward time τ goes to infinity. Figure 20b gives a graph of the value function for this case. The fourth picture shows a transition case between the third and fifth pictures.

5 Acoustic Game

Let us return to problems where player P minimizes and player E maximizes the time to reach the target set M. In papers [8, 9], P. Cardaliaguet, M. Quincampoix, and P. Saint-Pierre have considered an "acoustic" variant of the homicidal chauffeur problem suggested by Pierre Bernhard [4]. It is supposed that the constraint ν on the control of player E depends on the state (x, y). Namely,

$$\nu(x, y) = \nu^* \min\left\{1, \sqrt{x^2 + y^2}/s\right\}, \ s > 0.$$

Here, ν^* and s are parameters of the problem.

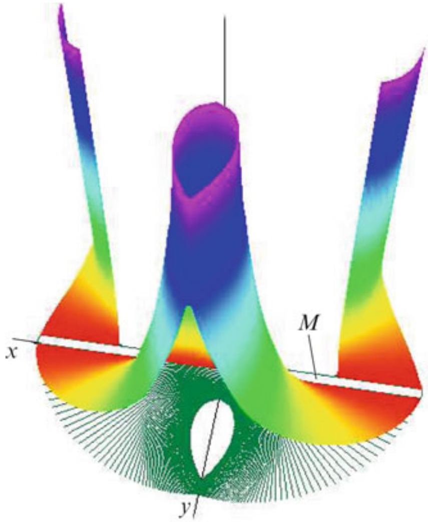


Fig. 21 Graph of the value function in the acoustic problem; $\nu^* = 1.5$, $s = 0.9375$

The applied aspect of the acoustic game is the following: object E should not be very loud if the distance between him and object P becomes less than a given value s.

P. Cardaliaguet, M. Quincampoix, and P. Saint-Pierre investigated the acoustic problem using their own numerical method for solving differential games which is based on viability theory [1]. It was revealed that one can choose parameter values in such a way that the set of states where the value function is finite will contain a hole, where the value function is infinite. Such a case can be obtained especially easily when the target set is a rectangle stretched along the horizontal axis.

Figure 21 shows an example of the acoustic problem with the hole. The graph of the value function is shown. The value of the game is infinite outside the set filled out with the fronts. An exact theoretical description of the arising hole and the computation (both analytical and numerical) of the value function near the boundary of the hole seems to be a very complex problem.

6 Game with a More Agile Player *P*

The dynamics of player P in Isaacs' setting is the simplest one among those used in mathematical publications for the description of car motion (or aircraft motion in the horizontal plane). In this model, the trajectories are curves of bounded curvature. In the paper [21] by Andrey Andreevich Markov published in 1889, four problems related to the optimization over curves with bounded curvature have been considered. The first problem (Fig. 22) can be interpreted as a time-optimal control problem where a car has the dynamics of player P. Similar interpretation can be given to the main theorem (Fig. 23) of the paper [11] by Lester E. Dubins

Нѣсколько примѣровъ рѣшенія особаго рода задачъ о наибольшихъ и наименьшихъ величинахъ.

А. А. Маркова.

ЗАДАЧА 1.

Между данными точками A и B (см. фиг. 1-ю) провести кратчайшую кривую линію при слѣдующихъ двухъ условіяхъ: 1) радіусъ кривизны нашей кривой повсюду долженъ быть не меньше данной величины ρ, 2) въ точкѣ A касательная къ нашей кривой должна имѣть данное направленіе AC.

РѢШЕНІЕ.

Пусть M одна изъ точекъ нашей кривой, а прямая NMT соотвѣтствующая касательная.

Fig. 22 Fragment of the first page of the paper [21] by A. Markov. “Problem 1: Find a minimum length curve between given points A and B provided that the following conditions are satisfied: 1) the curvature radius of the curve should not be less than a given quantity ρ everywhere, 2) the tangent to the curve at point A should have a given direction AC. Solution: Let M be a point of our curve, and the straight line *NMT* be the corresponding tangent...”

AMERICAN

JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

Volume LXXIX, Number 3

JULY, 1957

ON CURVES OF MINIMAL LENGTH WITH A CONSTRAINT ON AVERAGE CURVATURE, AND WITH PRESCRIBED INITIAL AND TERMINAL POSITIONS AND TANGENTS.*[1]

By L. E. Dubins.

We have now established our main result:

Theorem I. *Every planar R-geodesic is necessarily a continuously differentiable curve which is either* (1) *an arc of a circle of radius R, followed by a line segment, followed by an arc of a circle of radius R; or* (2) *a sequence of three arcs of circles of radius R; or* (3) *a subpath of a path of type* (1) *or* (2).

Fig. 23 Two fragments of the paper by L. Dubins published in 1957. The name "car" is used neither by A. A. Markov nor by L. Dubins. A. A. Markov mentions problems of railway construction. In modern works on theoretical robotics [17], an object with the classical dynamics of player P is called "Dubins' car".

The next in complexity is the car model from the paper by James A. Reeds and Lawrence A. Shepp [32]:

$$\dot{x}_p = w \sin\theta, \quad \dot{y}_p = w\cos\theta, \quad \dot{\theta} = u;$$
$$|u| \le 1, \ |w| \le 1.$$

The control u determines the angular velocity of motion. The control w is responsible for the instantaneous change of the linear velocity magnitude. In particular, the car can instantaneously reverse the direction of motion. A noninertia change of the linear velocity magnitude is a mathematical idealization. But, citing [32, p. 373], "for slowly moving vehicles, such as carts, this seems like a reasonable compromise to achieve tractability".

It is natural to consider problems where the range for changing the control w is $[a, 1]$. Here, $a \in [-1, 1]$ is a parameter of the problem. If $a = 1$, Dubins' car is obtained. For $a = -1$, one arrives at Reeds-Shepp's car.

Let us replace in (1) the classic car by a more agile car. Using the transformation to the reference coordinates, we obtain

$$\dot{x} = -yu + v_x, \quad \dot{y} = xu - w + v_y;$$
$$|u| \le 1, \ w \in [a, 1], \ v = (v_x, v_y)', \ |v| \le \nu. \tag{2}$$

Player P is responsible for the controls u and w, player E steers with the control v.

Note that J. Breakwell and J. Lewin investigated the surveillance-evasion game [18, 19] with the circular detection zone in the assumption that, at every time instant, player P either moves with the unit linear velocity or remains immovable. Therefore, they actually considered dynamics like (2) with $a = 0$.

The homicidal chauffeur game where player P controls the car which is able to change his/her linear velocity magnitude instantaneously was considered by the authors of this chapter in [30]. The dependence of the solution on the parameter a specifying the left end of the constraint to the linear velocity magnitude was investigated numerically.

Figure 24 shows the level sets of the value function for $a = -0.1$, $\nu = 0.3$, $r = 0.3$. The computation is done backward in time till $\tau_f = 4.89$. Precisely, this value of the game corresponds to the last outer front and to the last inner front adjoining to the lower part of the boundary of the target circle M. The front structure is well visible in Fig. 25 showing an enlarged fragment of Fig. 24. One can see a nontrivial character of changing fronts near the lower border of the accumulation region. The value function is discontinuous on the arc dhc. It is also discontinuous outside M on two short barrier lines emanating tangentially from the boundary of M. The right barrier is denoted by ce.

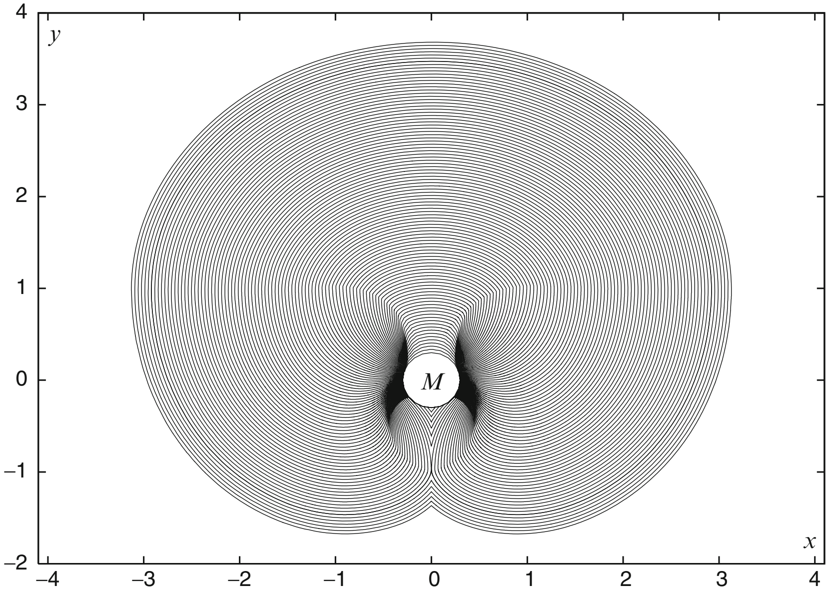
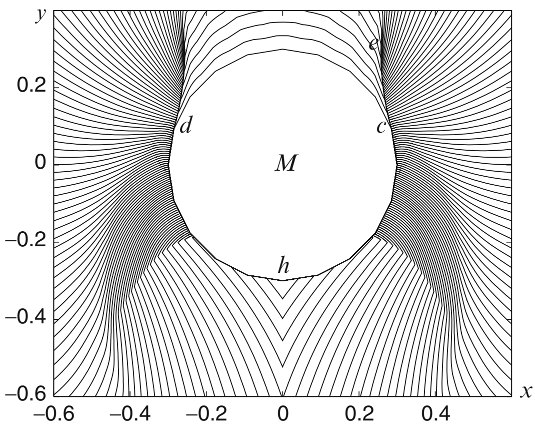


Fig. 24 Level sets of the value function in the homicidal chauffeur game with more agile pursuer; $a = -0.1$, $\nu = 0.3$, $r = 0.3$; $\tau_f = 4.89$, $\Delta = 0.002$, $\delta = 0.05$

Fig. 25 Enlarged fragment of Fig. 24



When solving time-optimal differential games of the homicidal chauffeur type (with discontinuous value function), the most difficult task is the construction of optimal (or ε-optimal) strategies of the players. Let us demonstrate such a construction using the last example.

We construct ε-optimal strategies using the extremal aiming procedure [15, 16]. The computed control remains unchanged during the next step of the discrete control scheme. The step of the control procedure is a modeling parameter. The strategy of player $P(E)$ is defined using the extremal shift to the nearest point (extremal repulsion from the nearest point) of the corresponding front. If the trajectory comes to a prescribed layer attached to the positive (negative) side of the discontinuity line of the value function, then a control which pushes away from the discontinuity line is utilized.

Let us choose two initial points $a = (0.3, -0.4)$ and $b = (0.29, 0.1)$. The first point is located in the right half-plane below the front accumulation region, the second one is close to the barrier line on its negative side. The values of the game in the points a and b are $V(a) = 4.225$ and $V(b) = 1.918$, respectively.

In Fig. 26, the trajectories for ε-optimal strategies of the players are shown. The time-step of the control procedure is 0.01. We obtain that the time of reaching the target set M is equal to 4.230 for the point a and 1.860 for the point b. Figure 26c shows an enlarged fragment of the trajectory emanating from the initial point b. One can see a sliding mode along the negative side of the barrier.

Figure 27 presents trajectories for nonoptimal behavior of player E and optimal action of player P. The control of player E is computed using a random number generator (random choice of vertices of the polygon approximating the circle constraint of player E). The reaching time is 2.590 for the point a and 0.300 for the point b. One can see how the second trajectory penetrates the barrier line. In this case, the value of the game calculated along the trajectory drops jump-wise.

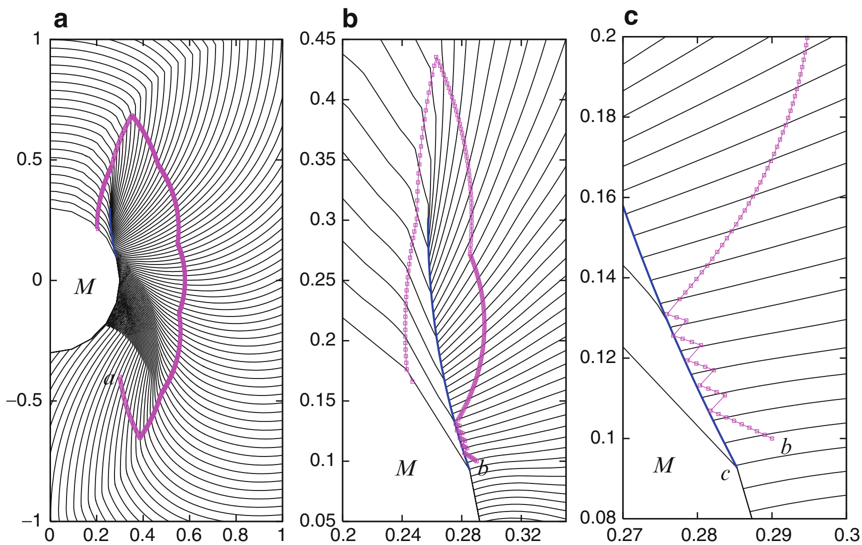


Fig. 26 Homicidal chauffeur game with more agile pursuer. Simulation results for optimal motions. (**a**) Initial point $a = (0.3, -0.4)$. (**b**) Initial point $b = (0.29, 0.1)$. (**c**) Enlarged fragment of the trajectory from the point b

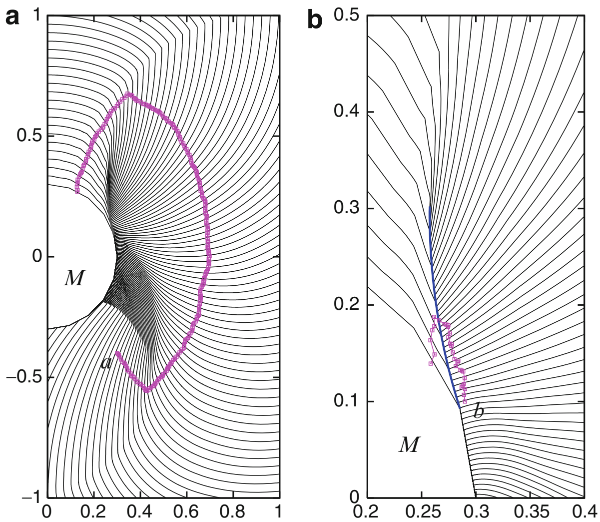


Fig. 27 Homicidal chauffeur game with more agile pursuer. Optimal behavior of player P and random action of player E. (**a**) Initial point $a = (0.3, -0.4)$. (**b**) Initial point $b = (0.29, 0.1)$

7 Homicidal Chauffeur Game as a Test Example

Presently, numerical methods and algorithms for solving antagonistic differential games are intensively developed. Often, the homicidal chauffeur game is used as a test or demonstration example. Some of these papers are [3, 25–27, 29, 31]. In the reference coordinates, the game is of the second order in the phase variable. Therefore, one can apply both general algorithms and algorithms taking into account the specifics of the plane. The nontriviality of the dynamics is in that the control u enters the right-hand side of the two-dimensional control system as a factor by the state variables, and that the constraint on the control v can depend on the phase state. Moreover, the control of player P can be two-dimensional, as it is in the modification discussed in Sect. 6, and the target set can be nonconvex like in the problem from Sect. 4.

Along with the antagonistic statements of the homicidal chauffeur problem, some close but nonantagonistic settings are known as being of great interest for numerical investigation. In this context, we note the two-target homicidal chauffeur game [12] with players P and E, each attempting to drive the state into his target set without being first driven to the target set of his opponent. For the first time, two-target differential games were introduced in [5]. The applied interpretation of such games can be a dogfight between two aircrafts or ships [10, 24, 28].

Acknowledgements The first author is supported by RFBR (grant 09–01–00436) and Program of the Presidium of RAS "Mathematical control theory".

References

1. Aubin, J.-P.: Viability Theory. Birkhäuser, Basel (1991)
2. Averbukh, V.L., Ismagilov, T.R., Patsko, V.S., Pykhteev, O.A., Turova, V.L.: Visualization of value function in time-optimal differential games. In: Handlovičová, A., Komorníková, M., Mikula, K., Ševčovič, D. (eds.) Algoritmy 2000, 15th Conference on Scientific Computing, pp. 207–216. Vysoke Tatry – Podbanske, Slovakia (2000)
3. Bardi, M., Falcone, M., Soravia, P: Numerical methods for pursuit-evasion games via viscosity solutions. In: Bardi, M., Raghavan, T. E. S., Parthasarathy, T. (eds.) Stochastic and Differential Games: Theory and Numerical Methods, Ann. of the Int. Soc. of Dyn. Games, Vol. 4, pp. 105–175. Birkhäuser, Boston (1999)
4. Bernhard, P., Larrouturou, B.: Étude de la barrière pour un problème de fuite optimale dans le plan. Rapport de recherche. INRIA, Sophia-Antipolis (1989)
5. Blaquière, A., Gérard, F., Leitmann, G.: Quantitative and Qualitative Differential Games. Academic Press, New York (1969)
6. Breakwell, J.V., Merz, A.W.: Toward a complete solution of the homicidal chauffeur game. In: Proc. of the 1st Int. Conf. on the Theory and Application of Differential Games, pp. III-1–III-5. Amherst, Massachusetts (1969)
7. Breitner, M.: The genesis of differential games in light of Isaacs' contributions. J. Optim. Theory Appl. **124**(3), 523–559 (2005)
8. Cardaliaguet, P., Quincampoix, M., Saint-Pierre, P.: Numerical methods for optimal control and differential games. Ceremade CNRS URA 749, University of Paris - Dauphine (1995)
9. Cardaliaguet, P., Quincampoix, M., Saint-Pierre, P.: Set-valued numerical analysis for optimal control and differential games. In: Bardi, M., Raghavan, T. E. S., Parthasarathy, T. (eds.) Stochastic and Differential Games: Theory and Numerical Methods, Ann. of the Int. Soc. of Dyn. Games, Vol. 4, pp. 177–247. Birkhäuser, Boston (1999)
10. Davidovitz, A., Shinar, J.: Two-target game model of an air combat with fire-and-forget all-aspect missiles. J. Optim. Theory Appl. **63**(2), 133–165 (1989)
11. Dubins, L.E.: On curves of minimal length with a constraint on average curvature and with prescribed initial and terminal positions and tangents. Am. J. Math. **79**, 497–516 (1957)
12. Getz, W.M., Pachter, M.: Two-target pursuit-evasion differential games in the plane. J. Optim. Theory Appl. **34**(3), 383–403 (1981)
13. Isaacs, R.: Games of pursuit. Scientific report of the RAND Corporation, Santa Monica (1951)
14. Isaacs, R.: Differential Games. Wiley, New York (1965)
15. Krasovskii, N.N.: Control of a Dynamic System. The Minimum Problem of a Guaranteed Result. Nauka, Moscow (1985) (in Russian)
16. Krasovskii, N.N., Subbotin, A.I.: Game-Theoretical Control Problems. Springer, New York (1988)
17. Laumond, J.-P. (ed.): Robot Motion Planning and Control. Lect. Notes in Control and Inf. Sci. **229**. Springer, New York (1998)
18. Lewin, J.: Decoy in pursuit-evasion games. PhD thesis, Stanford University (1973)
19. Lewin, J., Breakwell, J.V.: The surveillance-evasion game of degree. J. Optim. Theory Appl. **16**(3–4), 339–353 (1975)
20. Lewin, J., Olsder, G.J.: Conic surveillance evasion. J. Optim. Theory Appl. **27**(1), 107–125 (1979)
21. Markov, A.A.: Some examples of the solution of a special kind of problem on greatest and least quantities. Soobscenija Charkovskogo Matematiceskogo Obscestva **2, 1**(5, 6), 250–276 (1889) (in Russian)
22. Merz, A.W.: The homicidal chauffeur – a differential game. PhD thesis, Stanford University (1971)
23. Merz, A.W.: The homicidal chauffeur. AIAA Journal **12**(3), 259–260 (1974)
24. Merz, A.W.: To pursue or to evade – that is the question. J. Guid. Control Dyn. **8**(2), 161–166 (1985)

25. Meyer, A., Breitner, M.H., Kriesell, M.: A pictured memorandum on synthesis phenomena occurring in the homicidal chauffeur game. In: Martin-Herran, G., Zaccour, G. (eds.) Proceedings of the Fifth International ISDG Workshop, pp. 17–32. Int. Soc. of Dyn. Games, Segovia (2005)
26. Mikhalev, D.K., Ushakov, V.N.: Two algorithms for approximate construction of the set of positional absorption in the game problem of pursuit. Autom. Remote Control **68**(11), 2056–2070 (2007)
27. Mitchell, I.: Application of level set methods to control and reachability problems in continuous and hybrid systems. PhD Thesis, Stanford University (2002)
28. Olsder, G.J., Breakwell, J.V.: Role determination in aerial dogfight. Int. J. Game Theory **3**, 47–66 (1974)
29. Patsko, V.S., Turova, V.L.: Level sets of the value function in differential games with the homicidal chauffeur dynamics. Int. Game Theory Review **3**(1), 67–112 (2001)
30. Patsko, V.S., Turova, V.L.: Numerical investigation of the value function for the homicidal chauffeur problem with a more agile pursuer. In: Bernhard, P., Gaitsgory, V., Pourtallier, O. (eds.) Advances in Dynamic Games and Their Applications: Analytical and Numerical Developments, Ann. of the Int. Soc. of Dyn. Games, Vol. 10, pp. 231–258. Birkhäuser, Boston (2009)
31. Raivio, T., Ehtamo, H.: On numerical solution of a class of pursuit-evasion games. In: Filar, J.A., Mizukami, K., Gaitsgory, V. (eds.) Advances in Dynamic Games and Applications, Ann. of the Int. Soc. of Dyn. Games, Vol. 5, pp. 177–192. Birkhäuser, Boston (2000)
32. Reeds, J.A., Shepp, L.A.: Optimal paths for a car that goes both forwards and backwards. Pac. J. Math. **145**(2), 367–393 (1990)

Collision Avoidance Strategies for a Three-Player Game

Sriram Shankaran, Dušan M. Stipanović, and Claire J. Tomlin

Abstract Collision avoidance strategies for a game with three players, two pursuers and one evader, are constructed by determining the semipermeable curves that form the barrier. The vehicles are assumed to have the same capabilities, speed, and turn-rates. The game is assumed to be played on a two-dimensional plane. We consider avoidance strategies for a particular form of the game defined in the following way: the pursuers are assumed to act noncooperatively, the evader upon realizing that one (or both) of the pursuers can cause capture, takes an evasive action. We find states from which the pursuer can cause capture following this evasive action by the evader. The envelope of states that can lead to capture is denoted by the barrier set. Capture is assumed to have occurred when one (or both) pursuers have reached within a circle of radius, l, from the evader. The usable part and its boundary are first determined along with the strategy along the boundary. Semipermeable curves are evolved from the boundary. If two curves intersect (they have a common point), the curves are not extended beyond the intersection point. As in the game of two cars, universal curves and the characteristics that terminate and emanate from the universal curve are used to fill voids on the barrier surface. For the particular game (and associated strategies) considered in this paper, numerical simulations suggest that the enlarged set of initial states that lead to capture is closed. As the game considered here is a subset of the more complete game, when two pursuers try to cause capture of a single evader, the avoidance strategies are most likely to belong to the set of strategies for the complete game.

S. Shankaran (✉)
General Electric Research Center, Niskayuna, NY, USA
e-mail: sshankaran@gmail.com

D.M. Stipanović
Department of Industrial and Enterprise Systems Engineering, and Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Illinois, USA
e-mail: dusan@illinois.edu

C.J. Tomlin
Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, USA
e-mail: tomlin@eecs.berkeley.edu

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_13,

1 Introduction

It can be argued that the exposition of barriers and optimal trajectories in differential games was laid out in Isaacs's book [1]. Since this seminal work, there has been a countable number of attempts at characterizing the barriers that arise in differential games. The first complete solution for the game of the homicidal chauffeur was provided by Merz in his doctoral dissertation [3] and the breadth and depth of this analysis remains astounding. The solution to the Game of Two Cars by Merz [2] provided a solution technique to determine the barrier when the two cars had equal capabilities, namely speed and turn-rates. The solution procedure, though, was built on a careful study of the possible optimal trajectories that can emanate from the usable part of the target boundary and hence as its strength provided the solution to the Game of Degree along with the barrier. The extension of this approach to other games or games involving more players requires the arduous endeavor of conjecturing the trajectories that lead to the barrier.

The later work of Pachter and Getz [4] determines the barrier for the Game of Two Cars by constructing the curves (or trajectories) on the barrier. A combination of geometric and feasibility arguments were used to determine the particular combinations of control inputs that enabled curves on the barrier to be constructed, terminated and "voids" to be filled by curves emanating from a universal curve. While this approach can provide optimal avoidance strategies, it appears that this approach could experience difficulties when extended to higher dimensions or other games of the same dimension.

There seems to be a period of dormancy until around 2000 when the use of level-sets, dynamic implicit surfaces and numerical solutions to the Hamilton–Jacobi equation was used in an ingenious way by Mitchell, Bayen, and Tomlin [8] to recover the barrier for the game of two cars and other problems of interest in aircraft collision-avoidance. While this approach does not require explicit computation of the trajectories and the optimal inputs, the curse of dimensionality handicaps the approach to be applied in higher dimensions without support from numerical algorithms and computational architecture (although one can argue that significant strides are being made on both of these fronts). Additionally, it is not clear whether the optimal inputs to avoid collision can be easily extracted from the static zero-level set that results from this computation.

The more recent series of articles by Patsko and Turova [7] explore an alternate approach that resides in the idea of finding roots of the Hamiltonian for families of semipermeable curves that can arise in differential games. They focus on a particular set of dynamics, namely the one associated with the homicidal chauffeur, and suggest an algorithm that can evolve normals to the barrier based on the existence of particular sets of roots for the Hamiltonian for the corresponding points in state-space. The immediate difficulty of this approach is the maze produced by the combinatorial problem of many families of semipermeable curves that can exist for a differential game but as its advantage provides a computationally/numerically tractable algorithm. The other difficulty with this approach is the possibility of multiple solutions when the system of equations that provide the normal to the barrier at each point along the semipermeable curve is underdetermined. The authors explicitly mention that their approach while not universal provides an elegant approach to the particular problem of the homicidal chauffeur.

1.1 Isaacs's Approach [1]

As all existing methods have some history/origin in Isaacs' approach, we briefly study its tenets. Given a dynamical system, $\dot{x} = f(x, \sigma)$, where $x, f \in R^n$ and $\sigma \in R^m$ are the control inputs of the players and n is the dimension of the state-space and m is the dimension of the control space, Isaacs uses RPEs to evolve trajectories on the barrier that are of the form, $\mathring{x} = -f(x, \sigma)$, where $^\circ$ is used to represent temporal derivatives in backward time. Hence, $\mathring{x} = -\dot{x}$. The barrier is characterized in terms of its normal, $\nu = \nu_i, i = 1, n$, and can be written in a Hamiltonian form as

$$\min_p \max_e H(\nu, f) = \min_p \max_e \langle \nu, f \rangle = 0,$$

where the min and max are performed by the minimizing and maximizing players in the game. It must be noted that the class of dynamics that we are interested in, results in Hamiltonians that have a separable form and hence the order of the minimizing and maximizing operation does not alter the solution. Upon taking derivatives of the Hamiltonian, Isaacs writes "adjoint" equations for the evolution of the normals to the barrier. In our game, the optimal inputs are assumed to lie in a constrained interval in one-dimensional space, $[-1, 1]$, and hence we can perform the optimization *a priori* to obtain the inputs as functions of x and ν. The next step is to determine the boundary of the usable part from the target and having determined this, one can conceptually evolve trajectories that will be curves on the barrier. A sufficiently dense collection of these curves can then characterize the barrier. These curves can intersect (in which case they need to be terminated), converge (in which case rules to extend all, some or none of the curves have to be devised), or diverge (in which case the "void" left in the barrier surface has to be filled). Isaacs left the details of conclusively proving in the affirmative or negative to the readers. It turns out that much of the difficulty in determining the barrier for the game of two cars stems from this fact. However, the mathematical infrastructure is, nevertheless, the bedrock of the analysis in this study.

1.2 Patsko and Turova [7]

In a recent series of articles, Patsko and Turova expounded an alternate algorithm that determines families of semipermeable curves associated with the roots of the Hamiltonian. For the homicidal chauffeur game, they identified two families of semipermeable curves and due to the convexity of the Hamiltonians, each family had two possible roots. If a point in state-space had one root, one semipermeable curve passed through it and if there were two roots, two curves passed through it and so on.

An algorithm was then suggested to determine these semipermeable curves which can be outlined as follows:

- Consider a dynamical process, $\dot{x} = \Pi \quad l(x, \sigma)$. Here, x is the evolution in state-space of the semipermeable curve, Π is a rotation matrix that rotates l, the normal to the barrier by $\pm\pi/2$ degrees. The choice of $\pm$ depends on the nature of the normal as defined in Patsko and Turova [7].
- Determine whether any or all Hamiltonians, $H(l, x)$ have a root for a given x. If a root exists, compute it and evolve the dynamical system in time.
- The roots of the Hamiltonian are determined numerically either from an approximation to a circular vectogram or from possible values on the line interval that the optimal inputs can assume.

The manner in which Patsko and Turova suggested identification of the various families of semipermeable curves immediately posses the problem of exploding combinations in higher dimensions in addition to defining the rotation matrix, Π, to determine the unique normal on the perpendicular plane for three-dimensional games. The more important difficulty is that it is possible that semipermeable curves can emanate from portions bounded by the boundary of the usable part and hence these need to be determined to define the capture zone correctly. Patsko and Turova do this using arguments from geometry that identify intersecting subsets of the centers for radii of circular motion for the phase dynamics in the homicidal chauffeur. Whether it is feasible to do this for other games in an automated manner is not clear and seems to be an open problem yet to be addressed.

1.3 Bardi, Falcone and Dolcetta's Approach [5, 6]

We briefly point to the work by Bardi, Falcone, and Dolcetta that solves the Hamilton–Jacobi–Isaacs Partial Differential Equation (PDE) using a discrete-time formulation. As this approach rests on solving the HJ equation (in a discrete sense) it shares the disadvantages of the Mitchell, Bayen, and Tomlin [8] approach.

1.4 Outline of the Paper

We now focus on our approach which follows from Isaacs and Merz but extends it to a game of three cars to determine interesting features of the semipermeable surface. Isaacs construction of solutions to various differential game problems was rooted in obtaining Retrograde–Path–Equations (RPEs) derived from the HJI PDE equation. In the game of degree, where one is interested in obtaining optimal trajectories, the RPEs are ordinary differential equations that determine the evolution of the state of the system and the derivatives of the value function. In the game of kind, where one is interested in obtaining avoidance trajectories, the RPEs determine the evolution of the state and the normal to the barrier surface. The game of kind that determines possible states that can lead to capture will be the focus of this study.

In the next section, the dynamics and the RPEs are derived. Section 3 discusses the terminal conditions and the optimal inputs for the different players. Section 4 discusses different characteristics that form the barrier. Section 5 shows the results for a game of three cars. Conclusions and future directions of research are discussed in Sect. 6.

2 Dynamics and RPEs

The dynamics of the vehicles are assumed to be

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = W\cos(\phi_i); \quad \frac{\mathrm{d}y_i}{\mathrm{d}t} = W\sin(\phi_i); \quad \frac{\mathrm{d}\phi_i}{\mathrm{d}t} = \sigma_i,$$

where $-1 \leq \sigma \leq 1$ is the control input that the vehicles choose and W is assumed to be 1. Capture is defined by the time taken by one of the pursuers to reach a radius, l, of the evader. The two pursuers are referred to with subscripts 1 and 2 and the evader is labeled with a subscript e. Figure 1 shows the relative coordinate

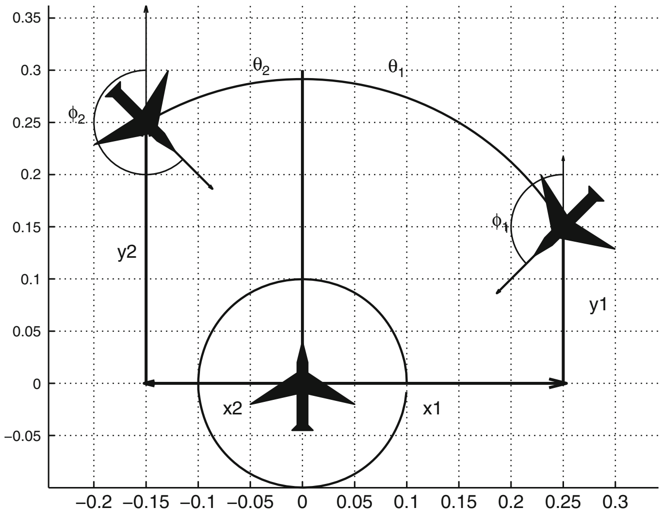


Fig. 1 Relative Coordinate System; all angles measured positive clockwise

system used to solve the game. To reduce the dimensionality of the problem, the dynamics is written in relative coordinates with the evader fixed at the origin as

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = -\sigma_e y_1 + \sin(\phi_1); \quad \frac{\mathrm{d}y_1}{\mathrm{d}t} = -1 + \sigma_e x_1 + \cos(\phi_1); \quad \frac{\mathrm{d}\phi_1}{\mathrm{d}t} = -\sigma_e + \sigma_{p1}$$

$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = -\sigma_e y_2 + \sin(\phi_2); \quad \frac{\mathrm{d}y_2}{\mathrm{d}t} = -1 + \sigma_e x_2 + \cos(\phi_2); \quad \frac{\mathrm{d}\phi_2}{\mathrm{d}t} = -\sigma_e + \sigma_{p2}.$$

The pay-off for this game is of the following type,

$$J = \min\left\{\left(\int_0^{t_1} 1\mathrm{d}t, \int_0^{t_2} 1\mathrm{d}t\right)\right\},$$

where t_1 and t_2 are the time-to-capture for pursuer 1 and 2 respectively. The evader chooses σ_e so as to maximize J and the pursuers choose σ_{p_i} so as to minimize J. As the pay-off is of the nondifferentiable type, it is not easy to adapt Isaacs's approach for the game of degree but we are currently only concerned with collision avoidance strategies, which are derived from the following definition of the barrier surface:

$$\min_{\sigma_{p1}} \min_{\sigma_{p2}} \max_{\sigma_e} < \nu \cdot f > = 0.$$

The RPEs for the game of kind can be written as

$$\begin{aligned} \overset{\circ}{x} &= -f(x, \sigma_e, \sigma_1, \sigma_2); \left(\overset{\circ}{\nu_{x_i}}\right) = \nu_{y_i}\sigma_e; \left(\overset{\circ}{\nu_{y_i}}\right) \\ &= -\nu_{x_i}\sigma_e; \left(\overset{\circ}{\nu_{\phi_i}}\right) = \nu_{x_i}\cos(\phi_i) - \nu_{y_i}\sin(\phi_i); \quad i = 1, 2, \end{aligned}$$

where $\nu = [\nu_{x_i}, \nu_{y_i}, \nu_{\phi_i}]$ is the normal to the barrier (semipermeable surface), the time derivatives (represented with a $^\circ$symbol) correspond to integration in backward time, and the indices, i, correspond to the relative coordinate system of the two pursuers. To integrate the above system of coupled ODEs, the state of the game at terminal time is to be determined.

The pay-off for the game is the time it takes for either one of the pursuers to reach the capture circle surrounding the evader. The strategy set falls under the category formalized by: the evader chooses its control input first, the pursuers having knowledge of this input, pick their inputs, and then all three players execute their inputs simultaneously. We note that the evader's strategy is optimal with respect to the more dangerous pursuer and not optimal with respect to the other pursuer. The pursuers do not cooperate in deciding their strategy but play a noncooperative game, where each pursuer tries to reach the capture circle independently.

3 Terminal Conditions and Optimal Inputs

The terminal states of the two pursuers can be written in terms of their distances, β and γ, the angular positions, θ_{01} and θ_{02} and the relative headings, ϕ_{01} and ϕ_{02}, as

$$x_1(T) = (\beta\sin(\theta_{01}), \beta\cos(\theta_{01}), \phi_{01})$$
$$x_2(T) = (\gamma\sin(\theta_{02}), \gamma\cos(\theta_{02}), \phi_{02})$$

When the game terminates, either $\beta = l$ or $\gamma = l$ or both β and $\gamma = l$. At capture, the relative radial velocity for both pursuers can be a) zero for one and nonzero for the other or b) zero for both. The radial velocity for each pursuer can be (after manipulation of the kinematic equations) written as

$$\frac{\mathrm{d}r_i}{\mathrm{d}t} \leq \cos(\theta_{0i} - \phi_{0i}) - \cos(\theta_{0i}); \quad i = 1, 2. \tag{1}$$

For the game of kind, the boundary of the usable part is determined by the equality relation in (1) for the pursuer on the capture circle, where

$$\phi_{0i} = 2\theta_{0i} \quad \text{or} \quad \phi_{0i} = 0$$

and the inequality relation holds for the other pursuer that can be rewritten as

$$\cos(\theta_{0i} - \phi_{0i}) - \cos(\theta_{0i}) < 0.$$

Hence, the boundary of the usable part in (x_2, y_2, ϕ_2) coordinates spans the circumferential extent from $(0, \phi/2)$. A HJI-like equation can be written for the barrier surface

$$\min_{\sigma_{p1}} \min_{\sigma_{p2}} \max_{\sigma_e} < \nu \cdot f >= 0.$$

If the components of the normal to the barrier, ν, are written as

$$\nabla\nu = [\nu_1, \nu_2, \nu_3, \nu_4, \nu_5, \nu_6] = [\nu_{x_1}, \nu_{y_1}, \nu_{\phi_1}, \nu_{x_2}, \nu_{y_2}, \nu_{\phi_2}],$$

the control inputs (assuming that the players pick their inputs independently) are

$$\sigma_{p1} = -\mathrm{sign}(\nu_{\phi_1}); \sigma_{p2} = -\mathrm{sign}(\nu_{\phi_2});$$
$$\sigma_e = \mathrm{sign}(-\nu_{x_1} y_1 + \nu_{y_1} x_1 - \nu_{\phi_1} - \nu_{x_2} y_2 + \nu_{y_2} x_2 - \nu_{\phi_2}).$$

In the game of two cars, at termination, the elements of the normal can be written as $[\sin(\theta_{01}), \cos(\theta_{01}), 0]$. In the game of three cars, at termination, only one of the pursuers has zero radial velocity with respect to the evader. The components of the normal in the coordinate system corresponding to this pursuer will be similar to the game of two cars. The radial velocity of the other pursuer is decreasing and hence the values of the normals are typically indeterminate. (Termination with both pursuers with zero radial velocity is a special situation that can be handled separately.) Hence, in general, if pursuer 1 has zero radial velocity at termination, the components of the normal can be written as

$$\nabla \nu = [\nu_1, \nu_2, \nu_3, \nu_4, \nu_5, \nu_6] = [\sin(\theta_{01}), \cos(\theta_{01}), 0, \nu_{x_2}, \nu_{y_2}, \nu_{\phi_2}]$$

The optimal inputs for the game of kind can be written in terms of the retrograde path equations as

$$\begin{aligned} \sigma_e &= \operatorname{sign}(S), \quad \sigma_{p1} = -\operatorname{sign}(\nu_3), \quad \sigma_{p2} = -\operatorname{sign}(\nu_6), \\ S &= -\nu_1 y_1 + \nu_2 x_1 - \nu_3 - \nu_4 y_2 + \nu_5 x_2 - \nu_6. \end{aligned}$$

3.1 Evader's Input

Let us define S as,

$$S = S_1 + S_2,$$

where S_1, S_2 contain terms related to pursuer 1 and 2 as

$$\begin{aligned} S_1 &= -\nu_1 y_1 + \nu_2 x_1 - \nu_3 \\ S_2 &= -\nu_4 y_2 + \nu_5 x_2 - \nu_6 \\ \sigma_e &= \operatorname{sign}(S_1 + S_2) \end{aligned}$$

$S_1 = 0$ at termination and hence it would seem that evader's control input depends on S_2. However, the evader needs to maintain zero radial velocity with respect to the first pursuer and hence we will need to look at the retrograde derivatives of S_1 to determine the evader's input,

$$\begin{aligned} \overset{\circ}{S}_1 &= \sin(\theta_{01}) \\ \overset{\circ}{S}_2 &= \nu_4 \\ \sigma_e &= \operatorname{sign}(\sin(\theta_{01}) + \nu_4). \end{aligned}$$

Although the evader's input depends on the position of both the closer pursuer (meeting it with zero radial velocity) and the second pursuer that is farther away (trying to reduce its radial velocity), the evader will react to the closer pursuer, failing which capture time will be smaller. Hence, the evader's input will be $+1$ for positive θ_{01} and -1 for negative θ_{01}.

When $\theta_{01} = 0$, the evader's input can be computed from second retrograde derivatives of S_1 and can be shown to take one value from the set $\{0, -1, 1\}$. Unlike the game of two cars, the evader's input can be 0 at termination (when the two pursuers are symmetrically placed on the capture circle).

3.2 Pursuer's Input

The pursuers' control inputs at termination are

$$\sigma_{p1} = -\text{sign}(\nu_3) \quad \sigma_{p_2} = -\text{sign}(\nu_6)$$

As we have denoted that pursuer 1 is on the capture circle at termination, its control input is identically zero and we turn to the retrograde equations to determine the inputs. If $S_1 = \nu_3$,

$$\begin{aligned}\overset{\circ}{S}_1 &= \sin(\theta_{01} - \phi_1)\\ \sigma_{p1} &= -\text{sign}(\sin(\theta_{01}))\end{aligned}$$

This pursuer has input $+1$ for negative θ_{01} and -1 for positive θ_{01}. When $\theta_{01} = 0$, the second retrograde derivatives yield

$$\sigma_{p1} = \text{sign}(\sigma_{p1})$$

leading to inputs from $\{0, -1, +1\}$.

The control inputs for the other pursuer are more involved to compute as we do not know the values of the corresponding components of the normal at termination. Hence, it will be assumed that this purser can have all possible inputs from the set $\{0, -1, +1\}$.

Hence, in character, this game is similar to the game of two cars and hence it will be useful to obtain solutions to that game. This game has a known solution that results in a closed barrier. Intuitively, the barrier for the game of three cars will be "larger" in the sense that additional portions of the state space can form initial conditions from which capture can be ensured. The central theme of the paper is to construct these additional portions of the state-space, which will result in an "enlarged" barrier for the game of three players.

4 Retrograde Solutions

For fixed control inputs along optimal trajectories in the game of kind and game of degree, retrograde solutions to the equations of motion can be found in terms of the terminal state [2] for the two relative coordinate systems as follows,

4.1 $\sigma_e = -\sigma_{pi} = \pm 1$

$$\begin{aligned}x_i &= x_{0i}\cos(\tau) + \sigma_e(1 - \cos(\tau) + y_{0i}\sin(\tau) + \cos(\phi_i) - \cos(\phi_{0i} + \sigma_e\tau))\\ y_i &= y_{0i}\cos(\tau) + \sin(\tau) - \sigma_e(x_{0i}\sin(\tau) + \sin(\phi_i) - \sin(\phi_{0i} + \sigma_e\tau))\\ \phi_i &= \phi_{0i} + 2\sigma_e\tau \end{aligned} \tag{2}$$

4.2 $\sigma_e = \sigma_{pi} = \pm 1$

$$
\begin{aligned}
x_i &= x_{0i}\cos(\tau) + \sigma_e(1 - \cos(\tau) + y_{0i}\sin(\tau) + \cos(\phi_{0i} + \sigma_e\tau) - \cos(\phi_i)) \\
y_i &= y_{0i}\cos(\tau) + \sin(\tau) - \sigma_e(x_{0i}\sin(\tau) + \sin(\phi_{0i} + \sigma_e\tau) - \sin(\phi_i)) \\
\phi_i &= \phi_{0i}
\end{aligned} \tag{3}
$$

4.3 $\sigma_{pi} = 0, \sigma_e = \pm 1$

$$
\begin{aligned}
x_i &= x_{0i}\cos(\tau) - \tau\sin(\phi_i) + \sigma_e(1 - \cos(\tau) + y_{0i}\sin(\tau)) \\
y_i &= y_{0i}\cos(\tau) - \tau\cos(\phi_i) + \sin(\tau) - \sigma_e x_{0i}\sin(\tau) \\
\phi_i &= \phi_{0i} + \sigma_e\tau
\end{aligned} \tag{4}
$$

5 Results

The computational approach we adopt is the following:

1. Determine the avoidance strategies for the game of two cars, one pursuer and one evader.
2. Enumerate the optimal inputs of the evader in the game of two cars.

For each input of the evader,

1. Assume that the evader plays optimally with respect to the first pursuer but nonoptimally with respect to the second pursuer.
2. Position the second pursuer such that it will potentially enlarge the barrier surface. It is intuitively clear that certain initial positions will enlarge the barrier surface and some will not (for example, when the second pursuer is on the same side as the first pursuer but further away, one would not expect any changes to the barrier surface). Most of the changes to the barrier surface happen when the pursuers "bracket" the evader, so the second pursuer is to the left of the evader if the first pursuer is to the right (and vice-versa).
3. Determine the change to the barrier surface for the second pursuer due to the nonoptimal play of the evader.

Intuitively, we would expect that the barrier surface will differ from that of the game of two cars, when the pursuers bracket the evader in the true game space. This restricts the possible states from which the second pursuer can cause capture even when the evader is playing optimally only with respect to the first pursuer. If both pursuers are on the same side of the evader, the game will typically reduce to a one-on-one game.

5.1 Game of Kind for Two Cars

We solve this problem it in a coordinate system fixed with respect to the pursuer and make the appropriate changes to interpret the solution in a coordinate system fixed with respect to the evader. We determine the coordinate (x, y) that the RPEs pass through for a given ϕ plane and build the three-dimensional barrier surface from a series of two-dimensional boundary curves.

It can be easily argued that on the left boundary of the usable part, the pursuer is turning left and the evader is turning right and on the right boundary the directions of rotation are reversed. As the characteristics starting from the boundary on the left (trajectories corresponding to (2), $\sigma_p = -1$) show decreasing relative heading, a portion of the left barrier can easily be constructed by evolving these trajectories and determining their positions on a particular ϕ plane. For characteristics from the right boundary of the usable part, the relative heading is increasing and hence to determine the barrier along an interval of the relative heading coordinate, those characteristics from the boundary corresponding to smaller relative heading have to be considered (trajectories corresponding to (2), $\sigma_p = 1$).

Turning to the point in state-space corresponding to $(\theta, \phi) = (0, 0)$, a universal curve starts at this point and follows the trajectory of (4). However, the evolution of this trajectory and the trajectory corresponding to (2) for $\phi < 0$ leave a void on the barrier surface. We can fill this void by characteristics that meet at the universal curve. By evolving the universal curve and the tributaries that meet it on either side, the left barrier can be completely recovered and an indication of the portion of the boundary that gives rise to characteristics that lead to the right barrier can be identified. For portions of the left barrier, we evolve along the universal curve for a certain time interval, τ_1, and then change the trajectory to one given by (2). For the right barrier, we evolve along the universal curve for a period of time, τ_1, and then switch to the trajectory given by (3) with $\sigma_p = 1$.

The game of two cars has an envelope barrier and hence the trajectories that lead to the capture circle on the barrier are reflected in the game of degree. We refer the reader to Merz [2] and the SUDAAR report of Ian Mitchell for details that enable construction of the barrier.

Table 1 shows the time-to-reach the capture circle for initial conditions that lie on the barrier surface. (The markers in the table correspond to the plot in Fig. 2. $0^+ = 0 + \delta$ represents time-to-capture that is slightly larger than 0). The range of values is useful when extending the analysis to the game with three players as it enables us to determine the character of the game for different initial conditions.

Table 1 Envelope of time-to-reach for different portions of the barrier

Trajectory Type	Min Time	Max Time
Triangle (left)	0^+	$\pi/2$
Triangle (right)	0^+	$\pi/2$
Circle	0^+	$3\pi/4$
Diamond	0^+	$11\pi/12$
Star	0^+	$11\pi/12$

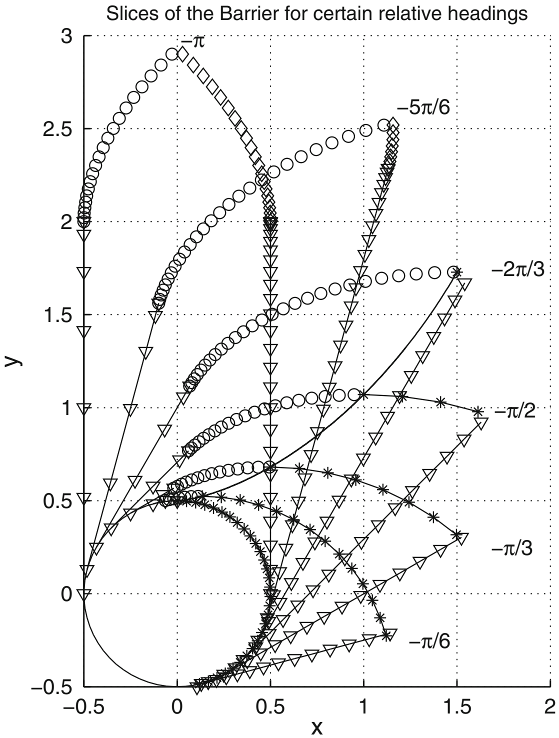


Fig. 2 Barrier projected onto the (x, y) plane for capture circle radius of 0.5. The points marked with a triangle reach the capture circle with the pursuer and evader turning in opposite directions. The evader is turning clockwise for points on the left and counterclockwise for points on the right. The points marked with a circle evolve to the capture circle through the universal curve (pursuer not turning and evader turning clockwise). The points marked with a diamond also reach the capture circle through an universal curve (pursuer not turning and evader turning counterclockwise). The points marked with a star reach the capture circle through the universal curve but with the players initially turning in the same direction

5.2 Game of Kind for Three Cars

We begin construction of the barrier for the game of three cars with the barrier for the game of two cars. Before we proceed we restate the payoff for the players in this game. The payoff for the evader is the time for capture by either one of the pursuers. For the pursuers, the payoff is the time it takes for either one of the pursuers to reach the capture circle.

The objective of this section is to determine how the barrier for the game of three cars is different from that of the game of two cars. As stated before, one would expect additional portions of the state-space to be included by the barrier in the game of three cars. These additions can be obtained in the following manner. The barrier surface for the game of two cars is the union of a left and right portion (also referred to as the left and right barrier). Hence, the barrier for the game of three cars will contain additions to the left and right portions. The enlarged barrier is formed by the union of the additions to the right, additions to the left and the barrier for the game of two cars. In the following sections, the additions to the left and right barrier surfaces are constructed. They will sometimes be referred to as enlargements (or increase in left or right barrier) as they play the role of enlarging the barrier when compared to the game of two cars.

The approach to determining the increase to the left and right barriers follows the sequence outlined in the beginning of Sect. 5. When the evader is playing optimally to escape from the first pursuer, the game of two cars suggests that the evader is turning away from the first pursuer. The evader is also on the left portion of the barrier for the first pursuer. In this scenario, if the second pursuer finds the evader to its right (and hence the pursuers have bracketed the evader) it provides the second pursuer to capture the evader from positions that were outside its original right barrier (the one from the game of two cars). While this might raise the question of nonoptimal play for the evader with respect to the second pursuer, when bracketed, the evader has very little choice but to prolong the time for capture. The same line of reasoning can be used to determine the increase to the left barrier.

5.2.1 Initial Positions of the Second Pursuer that Eliminate it From the Game

From the estimate of the barrier for the game of two cars, we first seek to determine positions of the second pursuer that result in a game between one of the pursuers and the evader. To achieve this, we allow the game to start from a relative orientation that places one of the pursuer on the barrier surface and enables the second pursuer to replicate the control inputs of the evader. As the evader's control inputs are ± 1 for all initial positions on the barrier surface, we evolve characteristics from all points on the capture circle where the second pursuer can cause a reduction in the radial velocity at termination. These points are given by the following relation:

$$\cos(\phi_{02} - \theta_{02}) - \cos(\theta_{02}) \leq 0.$$

Note that the game terminates with the first pursuer on the boundary of the usable part such that its radial velocity with respect to the evader is 0. Evolving all characteristics from the capture circle over a given period of time, τ, provides an envelope curve consisting of positions for which the second pursuer while turning with the evader can reach a circle of radius γ (need not be equal to β) while the first pursuer captures the evader. As the second pursuer is turning with the evader, the slices of the reach surface are easy to compute (from 3) and are shown in Figs. 3 ($\sigma_e = 1$) and 4 ($\sigma_e = -1$).

Intuitively, one would expect that initial states in Figs. 3 and 4 would be part of the enlarged barrier. However, it was observed during numerical simulations that the “triangle” points in Figs. 5 and 6 are either to the “left” or to the “right” of the “diamond” points in Figs. 3 and 4 hence suggesting that the pursuer has an alternate strategy to reach the capture circle in a shorter time. In other words, there exists

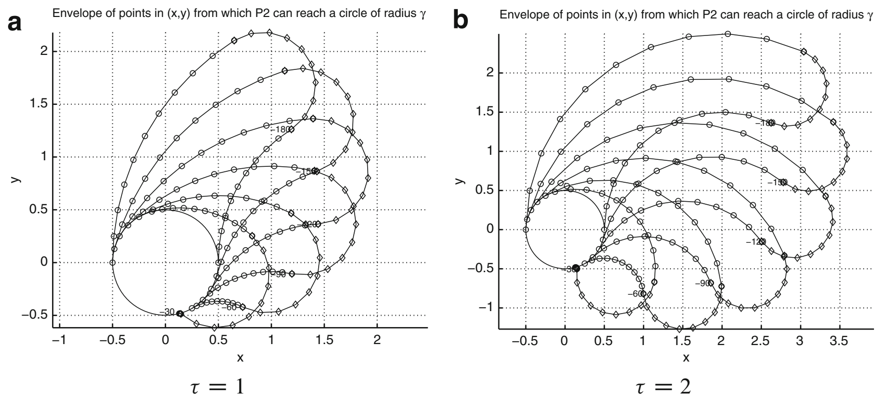


Fig. 3 Positions from which the second pursuer minimizes its distance from the evader while the first pursuer catches the evader, $\sigma_e = 1$. (**a**) $\tau = 1$ and (**b**) $\tau = 2$

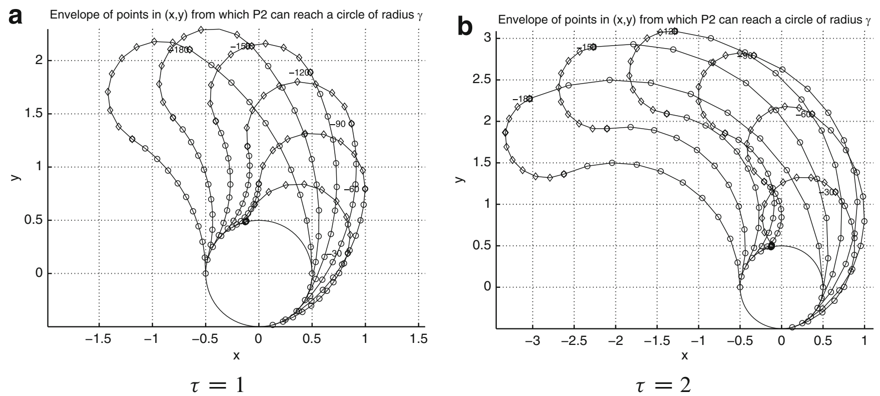


Fig. 4 Positions from which the second pursuer minimizes its distance from the evader while the first pursuer catches the evader, $\sigma_e = -1$. (**a**) $\tau = 1$ and (**b**) $\tau = 2$

a set of initial points from which if the evader turns with the closer pursuer, the pursuer will eventually cause capture. But there might exist initial conditions from which this pursuer can cause capture earlier. Based on the numerical simulations, it was observed that a two-part trajectory for the closer pursuer where it is not initially turning and then turns with the evader identifies points that are closer to the capture circle than the "diamond" points in Figs. 3 and 4. These are the "triangle" points in Figs. 5 and 6. A portion of the "circle" points (the point from the right of the capture circle is associated with the right barrier and vice versa) in Figs. 3 and 4 are part of the enlarged barrier and they initially extend from the capture circle (curves formed with "star" points in Figs. 5 and 6). Trajectories along which the first pursuer is not turning emanate from the "circle" points and form portions of the closed barrier ("triangle" points in Figs. 5 and 6).

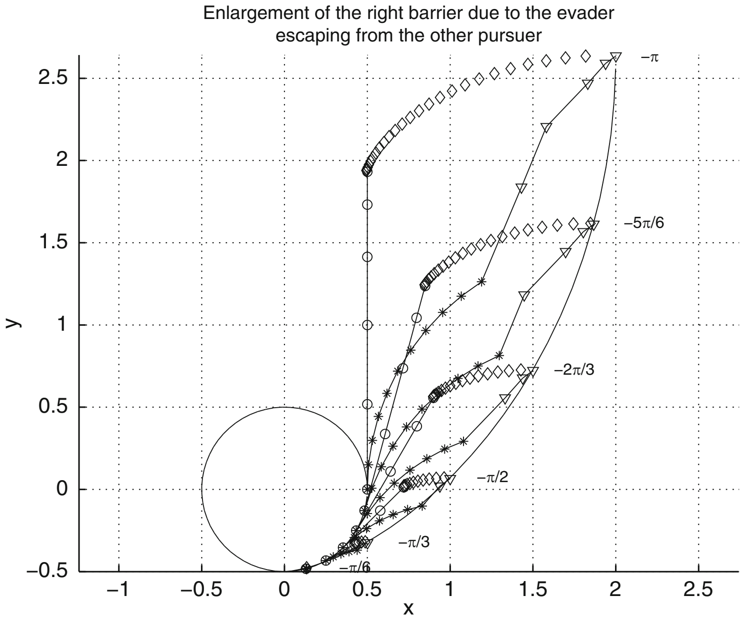


Fig. 5 Enlargement of right barrier projected onto the (x, y) plane. "Circle" points correspond to the pursuer and the evader turning in opposite directions, the "diamond" points correspond to the pursuer and the evader initially turning in opposite directions and then the pursuer not turning (universal curve), the "star" points correspond to the pursuer and the evader turning in the same direction, the "triangle" points correspond to the pursuer initially not turning and then switching to turning with the evader

5.2.2 Increase in the Right Barrier

To determine the change to the right barrier, we assume that the initial configuration is such that the evader is closer to the "first" pursuer and is trying to avoid earlier termination by escaping from this pursuer. This scenario leads to the increase in the size of the barrier for the "second" pursuer. If the evader was playing optimally with respect to the second pursuer, its control input would have been -1.

The first bounding surface of the enlarged right barrier is obtained from characteristics corresponding to (3), with $\sigma_p = 1$, from points on the capture circle corresponding to

$$\theta = \frac{\phi}{2} + \pi$$

These points are shown as curves with "stars" in Fig. 5. The universal curve evolving from the bottom of the capture circle is also shown (as a solid line with no markers) but is not part of this surface.

The next surface corresponds to "circle' points in Fig. 5 and is obtained by evolving trajectories when the evader is turning clockwise and the pursuer is turning in the opposite direction. Proceeding as in the game of two cars, we can determine a

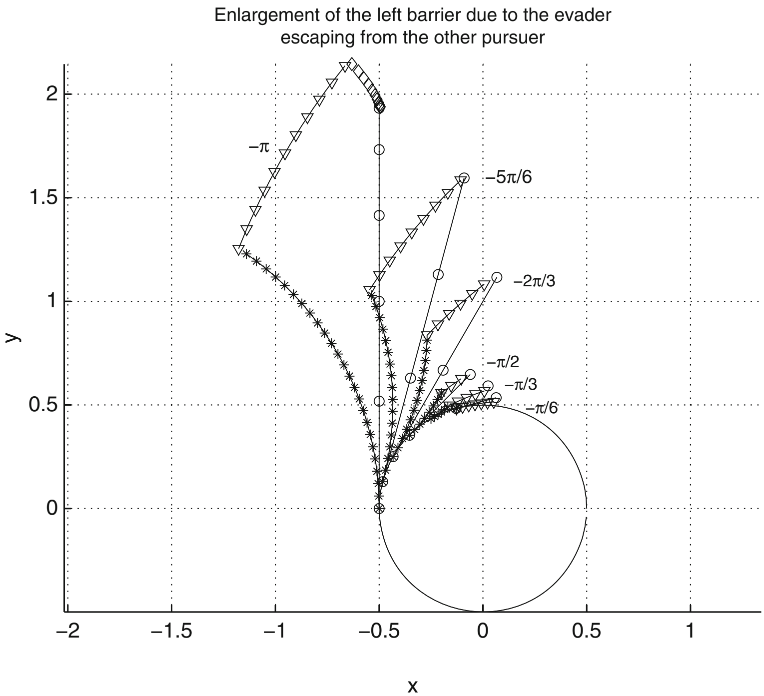


Fig. 6 Enlargement of left barrier projected onto the (x, y) plane. "Circle" points correspond to the pursuer and the evader turning in opposite directions, the "diamond" points correspond to the pursuer and the evader initially turning in opposite directions and then the pursuer not turning (universal curve), the "star" points correspond to the pursuer and the evader turning in the same direction, the "triangle" points correspond to the pursuer initially not turning and then switching to turning with the evader

portion of this additional barrier abutting the original right barrier by computing the (x, y) coordinates for each ϕ plane for a trajectory that evolves with (2). We see that this portion of the additional barrier joins the original right barrier but the straight line portions are shorter. Please note that the straight line portions are computed using a different strategy than in the game of two cars. The straight line portions are shorter than in the game of two cars as the inputs used to compute it are reversed. In the game of two cars, these points were also computed using (2) but the inputs were reversed.

Now, we evolve a universal curve from the point $(\theta, \phi) = (\pi, 0)$. As in the game of two cars there exist points from which the "second" pursuer and the evader are turning in opposite directions ("diamond" points in Fig. 5) and reach the capture circle through the universal curve.

The last surface that closes the extended right barrier consists of points from which the pursuer is initially not turning but switches to turning with the evader to cause capture. These points are shown in "triangle" in Fig. 5. As noted before, the presence of these "triangle" points is the reason why the "diamond" points in Fig. 3 are not included in Fig. 5.

Table 2 Envelope of time-to-reach for different portions of the enlarged right barrier

Trajectory Type	Min Time	Max Time
Circle	0^+	$\pi/2$
Diamond	0^+	π
Star	0^+	0.95
Triangle	0^+	$0.1{+}5\pi/6$

Intuitively, this set of trajectories provides the necessary extension of the right barrier. As the evader's initial position moves further right, it is to be expected that the motion of the second pursuer will change from (a) turning in opposite directions, to (b) turning in opposite directions and then not turning, to (c) initially not turning and then turning with the evader, and finally (d) turning with the evader.

Table 2 shows the range of values for the time-to-reach for the second pursuer.

In relation to the barrier for the game of two cars, we observe the following for the enlarged right barrier shown in Fig. 5. The "circle" points that form the straight line portion are shorter when compared to the barrier of the game of two cars. (Please observe that the straight line portions of the barrier in the game of two cars have different lengths on the left and right of the capture circle). For all relative headings between the second pursuer and the evader, the enlarged right barrier for the second pursuer while encompassing more states to the right does not extend as far as in the game of two cars. The physical interpretation of this situation is this: for states just to the right of the right barrier for the game of two cars, in the game of three cars, the second pursuer is not able to cause capture even when the evader is escaping from the first pursuer (unless these states happen to lie within the sets shown in Fig. 5). For states outside of this, the evasive inputs of the evader enables escape from both pursuers. It is conceivable that for certain initial states (for example, the "diamond" points in Fig. 3) the second pursuer can cause capture while the evader is escaping the first pursuer and hence they should be part of the enlarged right barrier in Fig. 5. For these states, the second pursuer can wait until the evader reaches a particular portion of the enlarged right and execute a strategy appropriate that portion. As noted before, these strategies are in general harder to compute comprehensively and hence we do not include them here. Please note that only a small set of "triangle" points were computed to Fig. 5 to enable easier visualization of the barrier surface.

5.2.3 Increase in the Left Barrier

The increase in the left barrier is computed using similar arguments as for the right barrier. The evader is assumed to be reacting to the "first" pursuer to maximize the capture time. This allows the "second" pursuer to enlarge the possible states from which it can cause capture. Hence, the pursuer we refer to the in this section is the "second" pursuer.

When the evader uses a -1 control input, the left barrier of the "second" pursuer can increase in shape. The first bounding surface of this enlarged barrier is obtained

Table 3 Envelope of time-to-reach for different portions of the enlarged left barrier

Trajectory Type	Min Time	Max Time
Circle	0^+	$\pi/2$
Diamond	$0.3+\pi/2$	$0.75+\pi/2$
Star	0^+	0.9
Triangle	0^+	$0.1+\pi$

by evolving characteristics corresponding to (2) with $\sigma_{p_2} = +1$ from points on the capture circle corresponding to

$$\theta = \frac{\phi}{2}.$$

These points are shown as curves with "star" points in Fig. 6.

The next bounding surface, shown with "circle" points in Fig. 6, corresponds to the trajectories in which the pursuer and the evader are turning in opposite directions (the evader is turning counterclockwise). These trajectories evolve without a switch to reach the capture circle. The "diamond" points correspond to characteristics in which the pursuer and the evader are turning in opposite directions initially but the pursuer switches its input to 0 to reach the capture circle through the universal curve from the top of the capture circle. The "triangle" points form the final surface that closes this increased left barrier, and they correspond to points from which the pursuer is initially not turning but then switches its input to turn with the evader to reach the capture circle.

Table 3 shows the time-to-reach the capture circle for the second pursuer. Note that the "diamond" points are not present on each slice of the barrier and hence the initial states from here have a min-time-to-capture. All other points "meet" the capture circle and hence have a min-time of 0.

In relation to the barrier for the game of two cars, we observe the following for the enlarged left barrier shown in Fig. 6. The "circle" points in the enlarged barrier fall on the "triangular" points of the left barrier in Fig. 2. For larger relative headings ($[-\pi, -2\pi/3]$) between the second pursuer and the evader, there seems to be a noticeable increase in the states from which capture can occur. For small relative headings, the corresponding increase is small. For the barrier in the game of two cars, we observed a portion of the right barrier marked with "diamond" points, where the pursuer and the evader were initially turning in opposite directions and the pursuer stopped turning. We observed a similar strategy on the left barrier extending only for a small portion of the slices in the enlarged left barrier for the game of two cars (shown in "diamond" points).

6 Conclusions

The method of characteristics is used for a game with three players to determine the increase in the set of states from which "capture" can occur. The different singular curves and surfaces on this extended barrier are currently being characterized.

It appears that the enlarged set of states is closed for the strategies considered in the study. It is quite possible that there exist other optimal strategies not considered in this study that will alter the capture set. The approach presented in this study can, in theory, be extended to games with more pursuers but the construction is tedious. An alternate analysis based on symplectic techniques seems to be capable of simplifying the process while adding rigor to the analysis.

The Matlab codes that perform the computations of the barrier are available from the authors on request.

Acknowledgements The authors would like to acknowledge Profs. V.S. Patsko and V.L. Turova for their initial proof-reading of the version submitted to the Symposium of Dynamic Games at Wroclaw, Poland in 2008. Their comments helped to identify misprints and unsupported logical conclusions that we have since tried to address. In addition, the comments and suggestions of two anonymous reviewers have greatly helped in improving the quality of the paper, to make precise the particular game studied here and improve the presentation and format of the results. The first author also wishes to acknowledge early support from ONR under the MURI grant when this work was initiated.

References

1. Isaacs, R.: Differential Games. Wiley, New York (1965)
2. Merz, A.W.: The Game of Two Identical Cars. Journal of Optimization Theory and Applications **9**(5), 324–343 (1972)
3. Merz, A.W.: The Homicidal Chauffeur. Ph.D. Thesis, Department of Aeronautics and Astronautics, Stanford University (1971)
4. Pachter, M., Getz, W.M.: The Geometry of the Barrier in the "Game of Two Cars". Optimal Control Applications and Methods **1**, 103–118 (1980)
5. Bardi, M., Dolcetta, C.I.: Optimal Control and Viscosity Solutions for Hamilton-Jacobi-Bellman Equations. Birkhäuser, Boston (1997)
6. Falcone, M., Bardi, M.: An approximation scheme for the minimum time function. SIAM Journal of Control and Optimization **28**(4), 950–965 (1990)
7. Patsko, V.S., Turova, V.L.: Families of semipermeable curves in differential games with the homicidal chauffeur dynamics. Automatica **40**(12), 2059–2068 (2004)
8. Mitchell, I., Bayen, A., Tomlin C.: Computing reachable sets for continuous dynamic games using level set methods. IEEE Transactions on Automatic Control **50**(7), 947–957 (2005)

Part III
Evolutionary Games

1. *Elvio Accinelli, Juan Gabriel Brida and Edgar J. Sanchez Carrera,* Imitative Behavior in a Two-Population Model
2. *Pierre Bernhard*, ESS, Population Games, Replicator Dynamics: Dynamics and Games if not Dynamic Games
3. *Yezekael Hayel, Hamidou Tembine, Eitan Altman and Rachid El-Azouzi,* A Markov Decision Evolutionary Game for Individual Energy Management
4. *David M. Ramsey*, Mutual Mate Choice with Multiple Criteria

Imitative Behavior in a Two-Population Model

Elvio Accinelli, Juan Gabriel Brida, and Edgar J. Sanchez Carrera

Abstract We study an evolutionary game with two asymmetric populations where agents from each population are randomly paired with members of the other population. We present two imitation models. In the first model, dissatisfaction drives imitation. In the second model, agents imitate the successful. In the first model, we use a simple reviewing rule, while in the second model we use a proportional imitation rule where switching depends on agents comparing their payoffs to others' payoffs. We show that such imitative behavior can be approximated by a replicator dynamic system. We characterize the evolutionarily stable strategies for a two asymmetric populations normal form game and we show that a mixed strategy is evolutionary stable if and only if it is a strict Nash equilibrium. We offer one clear conclusion: whom an agent imitates is more important than how an agent imitates.

1 Introduction

The replicator dynamics defines an explicit model of a selection process, specifying how population shares associated with different pure strategies in a game evolve over time. Taylor [11] considers a two-population model of the replicator dynamics and conceives a large population of agents who are randomly matched over time to play a finite, symmetric, two-player game, just as in the setting for evolutionary stability. We break the symmetry by introducing differences in payoffs and behaviors. We aim to find a dynamic system representing the imitative behavior inside a

E. Accinelli
Facultad de Economia, UASLP, Mexico
e-mail: elvio.accinelli@eco.uaslp.mx

J.G. Brida
School of Economics, Free University of Bozen, Bolzano, Italy
e-mail: juangabriel.brida@unibz.it

E.J.S. Carrera (✉)
Department of Economics, University of Siena, Italy
e-mail: sanchezcarre@unisi.it

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_14,

population facing another population, in a game where each agent is trying to do his/her best. We assume that all agents in the populations are infinitely lived and interact forever. Each agent adheres to some pure strategy for some time interval while periodically reviewing his/her strategy, and occasionally he/she will adopt a new strategy as a consequence of review [14].[1] We analyze an evolutionary social process, as defined by [8], with rational agents who do not follow pre-programmed strategies. The imitation rule consists of two parts. First, players may acquire their behavioral response by copying the behaviors of those who they perceive as successful. Second, they may acquire their responses by acting to maximize their gains given their beliefs about how others will act. Other influences are also at work, including conformism or dissatisfaction. Hence, we propose two imitation models.

1. Agents follow a pure imitation rule; they adopt the strategy of *"the first person they meet on the street"*. Agents randomly select another agent from the population and adopt the strategy of that agent. This simplistic hypothesis captures behaviors encountered in several populations.
2. Agents take into account their limited cognitive capacities and use imperfectly observed local information to update their behaviors. We assume that the probability that an agent adopts a certain strategy by reviewing agents to imitate correlates positively with the payoff currently expected by changing to that strategy. Although each agent imperfectly observes local information, successful strategies will tend to spread while unsuccessful strategies will tend to die out.
3. We show that these behavioral rules result in an adjustment process that can be approximated by the replicator dynamics.

To adopt an asymmetric approach is crucial in economics. Consider the following examples: economic models with incumbents and entrants in oligopolistic markets, social theory where we need to consider the relationships between migrants and residents with nonhomogeneous behaviors, or tourism economics where residents and tourists constitute two nonhomogeneous populations with different attitudes toward – and perceptions about – tourism development efforts or environmental quality.

Our model follows the framework of [12] and [9, 10]. Nevertheless, our main object is to analyze the dynamic characteristics of the Nash equilibria in an asymmetric context. These previously cited works point out how an agent who faces repeated choice problems will imitate others who obtained high payoffs.[2] Despite

[1] Björnerstedt and Weibull [3] study a number of models where agents imitate others in their own population, and they show that a number of payoff-positive selection dynamics, including different versions of the replicator dynamics, may be so derived.

[2] Schlag [9] analyzes what imitation rules an agent should choose, when he/she occasionally has the opportunity to imitate another agent in the same player position, but is otherwise constrained by severe restrictions on information and memory. He/she finds that if the agent wants a learning rule that leads to nondecreasing expected payoffs over time in all stationary environments, then the agent should (a) always imitate (not experiment) when changing strategy, (b) never imitate an agent whose payoff realization was worse than his/her own, and (c) imitate agents whose

their basic similarity, these two models differ along at least two dimensions: (1) the informational structure ("whom agents imitate") and, (2) the behavioral rule ("how agents imitate"). While agents in Vega-Redondo's model observe their immediate competitors, in Schlag's model agents observe others who are just like them, but play in different groups against different opponents. Apesteguia et al. [2] shows that the difference in results between the two models occurs because of the different informational assumptions rather than the different adjustment rules. So, whom an agent imitates is more important than how an agent imitates; We confirm this affirmation.

The paper is organized as follows. Section 2 describes the model. Section 3 introduces the pure imitation model where each agent implements his/her own strategy by imitating because of dissatisfaction, and we discuss the properties of the dynamic and Nash equilibria. Section 4 presents a model of imitation of successful agents. Section 5 characterizes the evolutionary stable strategy in the framework of an asymmetric normal form game. Section 6 concludes the paper.

2 The Model

Let us assume that agents, at a given period of time and in a given territory, comprise two populations: residents, R, and migrants, M. Each population splits into two clubs. The split depends on the strategy agents play or the behavior that agents follow. Suppose that these strategies are: *to admit or not to admit marriage with a member of the other population*$\{m, nm\}$. Let $x^\tau \in R^2_+$ be the vector $x^\tau = (x^\tau_m, x^\tau_{nm})\ \forall\ \tau \in \{R, M\}$, normalized to $x^\tau_m + x^\tau_{nm} = 1$, where each entry is the share of agents in the respective club over the total population. We introduce the symbol Δ to represent the 2-simplex in R^2, i.e:

$$\Delta = \{(x_1, x_2) \in R^2 : x_1 + x_2 = 1,\ x_1 \geq 0, x_2 \geq 0\}.$$

So, x^τ represents a distribution on the behavior of the agents in the population defined by $\tau \in \{R, M\}$.

Each period an i-strategy agent, $i \in \{m, nm\}$, from population $\tau \in \{R, M\}$, reviews his/her strategy with probability $r^\tau_i(x)$ asking himself/herself whether he/she may or may not change his/her current strategy, where $x = (x^M, x^R)$. The time-rate at which agents from each population review their strategy choice depends on the current performance of the agent's strategy relative to the current population state.

Let $p^\tau_{ij}(x)$ be the probability that a reviewing i-strategy agent adopts some pure strategy $j \neq i$, $\forall j \in \{m, nm\}$. Clearly, $r^\tau_i(x) p^\tau_{ij}(x)$ is the probability that an agent changes from the ith to jth club.

In the following, $e_m = (1, 0)$ and $e_{nm} = (0, 1)$ indicate vectors of pure strategies, m or nm, independent of population τ. Thus, the *outflow* from club i in

---

payoff realizations are better than his/her own with a probability that is proportional to this payoff difference.

population τ is $q_i^\tau r_i^\tau(x)p_{ij}^\tau(x)$ and the *inflow* is $q_j^\tau r_j^\tau(x)p_{ji}^\tau(x)$, where $q_i^\tau = q^\tau x_i^\tau$ is the number of i-strategists from population τ and q^τ represents all the agents from population τ, hence $q^\tau = q_i^\tau + q_j^\tau$. For any given population τ (assuming that the size of the population τ is constant) by the law of large numbers, we may model these processes as deterministic flows. Rearranging terms, we obtain the following:

$$\dot{x}_i^\tau = r_j^\tau(x)p_{ji}^\tau(x)x_j^\tau - r_i^\tau(x)p_{ij}^\tau(x)x_i^\tau, \ \forall i, j \in \{m, nm\}, \ j \neq i, \ \tau \in \{R, M\}. \tag{1}$$

System (1) represents the interaction between agents with different behaviors, $i \in \{m, nm\}$, in a given population $\tau \in \{R, M\}$, under a certain environmental state definite by the distributions $x = (x^M, x^R) = (x_m^M, x_{nm}^M; x_m^R, x_{nm}^R) \in \Delta \times \Delta$. In fact, (1) shows the dynamics of polymorphic populations with agents changing their behavior under imitation pressure and it allows us to analyze and to characterize the possible steady states for these populations.

The aim of this work is to characterize evolutionary stable distributions or population's states, in which each member of each population adopts a behavior that is the best possible given the behavior of those in their own population and given the characteristics of those in the other population. This is precisely a stable Nash equilibrium for our two-population game.

If for each agent in a given population there are n different possible behavior (pure strategies), then by Δ we represent the set of possible distributions on the set of possible behavior. So, by $x \in \Delta_s$ we represent a distribution on the set of the pure strategies, i.e, is a mixed strategy.

Let us define the following real numbers:

$$u^M(x, y) = \sum_{j=1}^{n}\sum_{i=1}^{n} u^M(e_i, y)x_i y_j$$

and

$$u^R(y, x) = \sum_{i=1}^{n}\sum_{j=1}^{n} u^R(e_j, x)x_i y_j,$$

where $u^M(e_i, y)$ is the agent's expected payoff from population M following the ith behavior, given that population R has the distribution y over the set of his/her own possible behaviors and analogously for $u^R(e_j, x)$. Recall that the asymmetry corresponds to migrants (M) and residents (R).

Definition 1. In a two-population normal form game with n different behaviors in each population, a pair of distributions (or a mixed strategy profile) (x, y) is a Nash equilibrium if the following inequalities are verified $u^M(x, y) \geq u^M(z, y)\ \forall\ z \in \Delta$ and $u^R(y, x) \geq u^R(w, x)$ for all $w \in \Delta$.

Intuitively, x is a best response of population-player M when population-player R displays y. Similarly, y is a best response of population-player R, given that player M displays x. $\Delta = \left\{x \in R^n_+ : \sum_{i=1}^n x_i = 1\right\}$ is the set of possible distributions over the set of the different behaviors, (the share of each behavior) or pure strategies, in a given population. In the next section, we will work with $n = 2$ different possible behaviors symbolized by m and nm, inside each population.

Consider the following example. Imagine a country where two different populations inhabit the same area. Call them 'migrants' and 'residents'. Both populations conform by disjoint "clubs" of agents with different behaviors, and the behavior of such agents may change by imitation. The main question here is *what kind of behavior will survive in the population if agents change because of imitative pressure?*

3 Imitation and Dissatisfaction

We introduce here the first evolutionary model – a simple imitation model. Each agent observes the performance of one other agent (see for instance [1, 3, 9, 10]). An agent's decision to stick to a strategy/club or to change strategy is a function of the type of agent he/she encounters in his/her own population. For a model of pure imitation, all reviewing agents adopt the strategy of *the first person that they meet in the street*, picking this person at random from the population.

We consider that:

(a) An agent's decision depends upon the utility associated with his/her behavior, given the composition of the other population, represented by the notation $u^\tau(e_i, x^{-\tau})$ (where τ represents the population to which the agent following the ith behavior belongs and $-\tau \in \{R, M\}$, $-\tau \neq \tau$) and on the characteristics of populations represented by $x = (x^R, x^M)$. So:

$$r_i^\tau(x) = f_i^\tau(u^\tau(e_i, x^{-\tau}), x).$$

We interpret the function $f_i^\tau(u^\tau(e_i, x^{-\tau}), x)$ as the propensity of a member of the ith club considering switching membership as a function of the expected utility gains from switching. Agents with less successful strategies review their strategy at a higher rate than agents with more successful strategies.

(b) Having opted for a change, an agent will adopt the strategy followed by the first population-player to be encountered (his/her neighbor), i. e., for any $\tau \in \{R, M\}$:

$$p(i \to j/\text{i considers a change to j}) = p_{ij}^\tau = x_j^\tau,\ i, j \in \{m, nm\}.$$

Considering *(a)* and *(b)*, (1) can be written as:

$$\dot{x}_i^\tau = x_j^\tau f_j^\tau(u^\tau(e_j, x^{-\tau}))x_i^\tau - x_i^\tau f_i^\tau(u^\tau(e_i, x^{-\tau}))x_j^\tau$$

or

$$\dot{x}_i^\tau = (1 - x_i^\tau)x_i^\tau \left[f_j^\tau(u^\tau(e_j, x^{-\tau})) - f_i^\tau(u^\tau(e_i, x^{-\tau}))\right].$$

This is the general form of the dynamic system representing the evolution of a two-population and four-club structure (two "clubs" in each population). It provides a system of four simultaneous equations with four state variables, where each state variable is the population share of the club members. However, given the normalization rule, $x^{\tau}_{m} + x^{\tau}_{nm} = 1$, for each $\tau = \{R, M\}$, (3) can be reduced to two equations with two independent state variables. Taking advantage of this property, henceforth we select variables x^{R}_{m} and x^{M}_{nm} with their respective equations.

To grapple with the problem, let us assume f^{τ}_{i} is population specific, but the same across all its components independent of club membership. Assume, furthermore, that it is linear in utility levels. Thus, the propensity to switch will be decreasing in the level of the utility:

$$f^{\tau}_{i}(u^{\tau}(e_i, x^{-\tau})) = \alpha^{\tau} - \boldsymbol{\beta}^{\tau} u^{\tau}(e_i, x^{-\tau}) \in [0, 1]$$

with $\alpha^{\tau}, \beta^{\tau} \geq 0$ and $\frac{\alpha^{\tau}}{\beta^{\tau}} \geq u^{\tau}(e_i, x^{-\tau})$ $\forall i$ and x^{τ}. To get a full linear form, we assume

$$u^{\tau}(e_i, x^{-\tau}) = e_i A^{\tau} x^{-\tau},\ i \in \{m, nm\};$$

in other words, utility is a linear function of both variables, through a population-specific matrix of weights or constant coefficients, $A^{\tau} \in \mathscr{M}_{2\times 2}$, $(\tau \in \{R, M\})$. The latter assumption implies that utility levels reflect population-specific properties, i.e., different preference structures over outcomes. The previous model therefore reduces to a much simpler model:

$$\dot{x}^{\tau}_{m} = \boldsymbol{\beta}^{\tau}\left[u^{\tau}(e_m, x^{-\tau}) - u^{\tau}(x^{\tau}, x^{-\tau})\right] x^{\tau}_{m}$$

equivalently

$$\dot{x}^{\tau}_{m} = \boldsymbol{\beta}^{\tau} x^{\tau}_{m}(1 - x^{\tau}_{m})\left[(1, -1)A^{\tau} x^{-\tau}\right],\ \tau \in \{R, M\}.$$

or in full:

$$\begin{cases} \dot{x}^{R}_{m} = \beta^{R} x^{R}_{m}(1 - x^{R}_{m})(a^{R} x^{M}_{m} + b^{R}) \\ \\ \dot{x}^{M}_{m} = \beta^{M} x^{M}_{m}(1 - x^{M}_{m})(a^{M} x^{R}_{m} + b^{M}) \end{cases}, \tag{2}$$

whose coefficients a^{M} and b^{R} depend upon the entries of the two population-specific matrices, A^{M} and A^{R}, respectively.

3.1 Dynamic Stability and Nash Properties

System (2) admits five stationary states or dynamic equilibria, i.e.

$$(0, 0),\ (0, 1),\ (1, 0),\ (1, 1) \text{ and a positive interior equilibrium } (\bar{x}^{R}_{m}, \bar{x}^{M}_{m}),$$

where

$$\bar{x}_m^R = -\frac{b^R}{a^R}, \quad \bar{x}_m^M = -\frac{b^M}{a^M}.$$

The most interesting case occurs when $\bar{P} = (\bar{x}_m^R, \bar{x}_m^M)$, characterizing an equilibrium lying in the interior of the square $\mathscr{C} = [0,1] \times [0,1]$, which happens when:

$$0 < -\frac{b^R}{a^R} < 1 \text{ and } 0 < -\frac{b^M}{a^M} < 1,$$

where $\left(a^R, a^M\right)$ and $\left(b^R, b^M\right)$ have opposite signs. These equilibria can be interpreted as follows:

- The trivial equilibrium obtains when none of the residents is inclined to admit migrants, nor is any migrant inclined to mix with residents.
- Another equilibrium obtains at the opposite corner, where the sharing clubs involve all of their respective population. Here, reciprocal integration of the two populations is complete. The two remaining border equilibria show a different club dominating the two populations and in a sense constitute a mismatch between strategies.
- Finally, we have the interior equilibrium. This equilibrium implies that a certain percentage of each population is well disposed to the other population while the rest only accept marriage between members of their own populations.

We now proceed to assess the stability of the five equilibria.

Proposition 1. *There exists a consistent coefficients range,*

$$\left\{-\frac{1}{2} < \left(b^R, b^M\right) < 0 < \left(a^R, a^M\right)\right\},$$

which ensures that the steady states $(1,1)$ *and* $(0,0)$ *are asymptotically stable equilibria, while* $(1,0)$ *and* $(0,1)$ *are nonstable nodes and* $\left(\bar{x}_m^R, \bar{x}_m^M\right)$ *is a saddle point.*

Proof. We can judge whether the five equilibria are stable or not via analyzing the Jacobian Matrix associated with the system (2), i.e.,

$$J = \begin{bmatrix} \beta^R(1-2x_m^R)(a^R x_m^M + b^R) & \beta^R a^R x_m^R(1-x_m^R) \\ \beta^M a^M x_m^M(1-x_m^M) & \beta^M(1-2x_m^M)(a^M x_m^R + b^M). \end{bmatrix}$$

The values of the matrix depend on the population-specific matrices. Recall that equilibria fitting $\det(J) > 0$ and $tr(J) < 0$ are asymptotically stable. Then, we have the following cases:

1. $x_m^R = x_m^M = 1$, the evaluated Jacobian in this case is

$$J = \begin{bmatrix} -\beta^R(a^R + b^R) & 0 \\ 0 & -\beta^M(a^M + b^M) \end{bmatrix}.$$

When $\left(b^R, b^M\right)$ are positive and $\left(a^R, a^M\right)$ negative numbers, then,

$$\det(J) = \left(-\beta^R(a^R + b^R)\right) \cdot \left(-\beta^M(a^M + b^M)\right) > 0$$
$$tr(J) = -\beta^R(a^R + b^R) - \beta^M(a^M + b^M) < 0$$

and then the equilibrium point $(1, 1)$ is a stable node. Otherwise, if $\left(b^R, b^M\right)$ are negative and $\left(a^R, a^M\right)$ positive numbers, then, $\det(J) > 0$ and $tr(J) < 0$, and then the equilibrium point $(1, 1)$ is asymptotically stable. Therefore, $(1, 1)$ is always a stable node.

2. $x_m^R = x_m^M = 0$, the evaluated Jacobian is

$$J = \begin{bmatrix} \beta^R b^R & 0 \\ 0 & \beta^M b^M \end{bmatrix}.$$

Note that, if $\left(b^R, b^M\right)$ are positive and $\left(a^R, a^M\right)$ negative numbers, it implies that

$$\det(J) = \left(\beta^R b^R\right) \cdot \left(\beta^M b^M\right) > 0$$
$$tr(J) = \left(\beta^R b^R\right) + \left(\beta^M b^M\right) > 0,$$

and then the equilibrium point $(0, 0)$ is a nonstable node. Anyway, with $\left(b^R, b^M\right)$ negative and $\left(a^R, a^M\right)$ positive numbers, $\det(J) > 0$ and $tr(J) < 0$, then the equilibrium point $(0, 0)$ is an asymptotically stable.

3. $x_m^R = 1$, $x_m^M = 0$, the evaluated Jacobian is

$$J = \begin{bmatrix} -\beta^R b^R & 0 \\ 0 & \beta^M(a^M + b^M) \end{bmatrix}.$$

If $\left(b^R, b^M\right)$ are positive and $\left(a^R, a^M\right)$ negative numbers, it implies that

$$\det(J) = \left(-\beta^R b^R\right) \cdot \left(\beta^M(a^M + b^M)\right) > 0$$
$$tr(J) = \left(\beta^R b^R\right) + \left(\beta^M b^M\right) < 0,$$

and then the equilibrium point $(1, 0)$ is an asymptotically stable node. Otherwise, if $\left(b^R, b^M\right)$ are negative and $\left(a^R, a^M\right)$ positive numbers, then, $\det(J) > 0$ and $tr(J) > 0$, and therefore the equilibrium point $(1, 0)$ is an unstable node, which is consistent with the fact that $(0, 0)$ and (1,1) are asymptotically stable nodes.

4. $x_m^R = 0$, $x_m^M = 1$, the evaluated Jacobian is

$$J = \begin{bmatrix} \beta^R(a^R + b^R) & 0 \\ 0 & -\beta^M b^M \end{bmatrix}.$$

If $\left(b^R, b^M\right)$ are positive and $\left(a^R, a^M\right)$ negative numbers, it implies that

$$\det(J) = \left(\beta^R(a^R + b^R)\right) \cdot \left(-\beta^M b^M\right) > 0$$
$$tr(J) = \left(\beta^R(a^R + b^R)\right) + \left(-\beta^M b^M\right) < 0,$$

and then the equilibrium point (0, 1) is an asymptotically stable node. Otherwise, if $\left(b^R, b^M\right)$ are negative and $\left(a^R, a^M\right)$ positive numbers, then, $\det(J) > 0$ and $tr(J) > 0$, and then the equilibrium point (0, 1) is an unstable node, which is consistent with the fact that (0,0) and (1,1) are asymptotically stable nodes.

5. The interior equilibrium $\bar{x}^R_m = -\frac{b^R}{a^R}$ and $\bar{x}^M_m = -\frac{b^M}{a^M}$. Evaluating this point in the Jacobian yields

$$J = \begin{bmatrix} \beta^R\left(1 + 2\dfrac{b^R}{a^R}\right)\left(-a^R\dfrac{b^M}{a^M} + b^R\right) & -\beta^R b^R\left(1 + \dfrac{b^R}{a^R}\right) \\ -\beta^M b^M\left(1 + \dfrac{b^M}{a^M}\right) & \beta^M\left(1 + 2\dfrac{b^M}{a^M}\right)\left(-a^M\dfrac{b^R}{a^R} + b^M\right) \end{bmatrix}.$$

If $-\frac{1}{2} < b^R < 0$ and $-\frac{1}{2} < b^M < 0$ and $\left(a^R, a^M\right)$ positive numbers, then, $\det(J) < 0$ and then the equilibrium point $\left(\bar{x}^R_m, \bar{x}^M_m\right)$ is a saddle point. □

Next the Fig. 1 draws the vector field of the replicator dynamic system in (2).

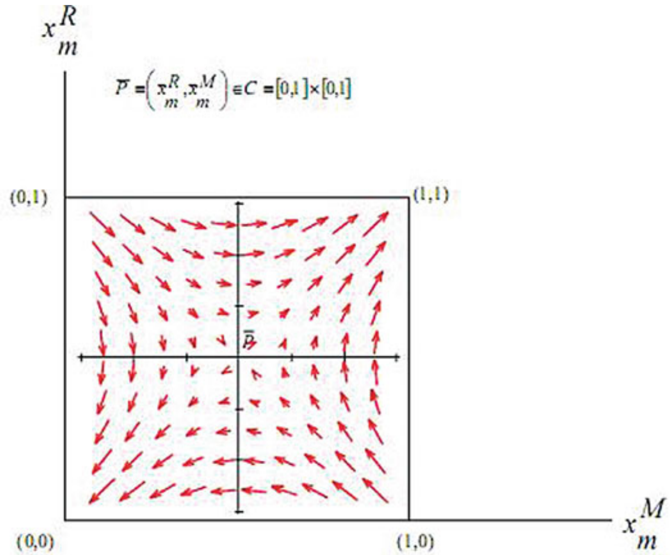


Fig. 1 Solution orbits for the system $\left(\dot{x}^R_m, \dot{x}^M_m\right)$

4 Picking the Most Successful Strategy

The rule now states that an agent must choose an action with the best payoff with a probability proportional to the expected payoff. In other words, a migrant (resident) will change to a strategy played by another member of his/her population if and only if the alternative strategy brings greater expected utility.

We assume that every reviewing strategist samples at random from his/her own population and that he/she can observe, with some noise, an average payoff of the sampled agent's strategy.

Let us assume that a reviewing strategist, from population $\tau \in \{R, M\}$, with behavior (club) $j \in \{m, nm\}$, meets another agent with an alternative behavior (club) $i \neq j$, and that, he/she will change from j to i if and only if:

$$u^\tau(e_i, x^{-\tau}) > u^\tau(e_j, x^{-\tau}).$$

The utility of each agent depends on his/her own strategy and on the characteristics of the agents of the other population. Further, we assume that there is some uncertainty in the strategies' estimated utility, so that each reviewing strategist must estimate the value of $u^\tau(e_i, x^{-\tau})$.

Let D be the estimator for the difference $u^\tau(e_i, x^{-\tau}) - u^\tau(e_j, x^{-\tau})$. Let $P^\tau(\bar{D} \geq 0)$ be the probability that the estimated $\bar{D} = u^\tau(e_i, x^{-\tau}) - \bar{u}^\tau(e_j, x^{-\tau})$ is positive.

In such a way, let p^τ_{ji} be the probability that a j-strategist becomes an i-strategist. This probability is given by

$$p^\tau_{ji} = x^\tau_{ji} P^\tau(\bar{D} \geq 0),$$

where x^τ_{ji} is the probability that a j-strategist meets an i-strategist, $\forall\ \tau \in \{R, M\}$. Let us assume that $P^\tau(\bar{D} \geq 0)$ depends upon the true value of the difference $u^\tau(e_i, x^{-\tau}) - u^\tau(e_j, x^{-\tau})$, which is unknown to the ith agent. That is to say

$$P^\tau(\bar{D} \geq 0) = \phi^\tau\left(u^\tau(e_i, x^{-\tau}) - u^\tau(e_j, x^{-\tau})\right), \tag{3}$$

where $\phi : R \to [0, 1]$ is a differentiable probability distribution. Therefore, the probability that a j-strategist, from population τ, estimates a positive value $\bar{D}$ increases with the true value of the difference $u^\tau(e_j, x^{-\tau}) - u^\tau(e_i, x^{-\tau})$. For the sake of simplicity, let $u^\tau(e_i, x^\tau)$ be linear, i.e.,

$$u^\tau(e_i, x^{-\tau}) = e_i A^\tau x^{-\tau}. \tag{4}$$

Thus, from (3) and (4) the probability that a j-strategist changes to the ith strategy is:

$$p^\tau_{ji}(u^\tau(e_i - e_j, x^{-\tau})) = \phi^\tau(u^\tau(e_i - e_j, x^{-\tau}))x^\tau_i.$$

Hence, the change in the share of the i-strategist will be given by the probability that a j-strategist becomes an i-type weighted by the relative number of j-strategists minus the probability that an i becomes a j-strategist, likewise weighted:

$$\dot{x}_i^\tau = \left[x_j^\tau p_{ji}^\tau - p_{ij}^\tau x_i^\tau\right] x_i^\tau .$$

In this case, the equation is:

$$\dot{x}_i^\tau = x_j^\tau x_i^\tau \left[\phi^\tau(u^\tau(e_j - e_i, x^{-\tau})) - \phi^\tau(u^\tau(e_i - e_j, x^{-\tau}))\right],$$

and a first-order approximation is given by:[3]

$$\begin{aligned}\dot{x}_i^\tau &= x_i^\tau x_j^\tau [\tau\phi^{\tau\prime}(0,)\left[u^\tau(e_j - e_i, x^{-\tau}) - u^\tau(e_j - e_i, x^{-\tau}))\right] = \\ &= 2\phi^{\tau\prime}(0) u^\tau(e_i - x^\tau, x^{-\tau}) x_i^\tau .\end{aligned}$$

Then, in a neighborhood of an interior stationary point, the dynamics is approximately represented by a replicator dynamics multiplied by a constant. Stability analysis of the local type can therefore be carried out using the linear part of the nonlinear system.[4]

In the special case where ϕ^τ is linear:

$$\phi^\tau = \lambda_\tau + \mu^\tau u^\tau(e_j - e_i, x^{-\tau}),$$

where λ^τ and μ^τ verify:

$$0 \leq \lambda^\tau + \mu^\tau u(x^\tau, x^{-\tau}) \leq 1, \ \forall\ x \in \left\{z \in R_+^2 : \max_{i=1,2} z_i \leq 1\right\}.$$

Thus, for each behavior i and each population τ the following equation holds:

$$\dot{x}_i^\tau = 2\mu^\tau u^\tau(e_i - x^\tau, x^{-\tau}) x_i^\tau, \tag{5}$$

which is merely a replicator dynamics. Hence, stability analysis is similar to the analysis of the model of pure imitation driven by dissatisfaction. Hence, it is more important who an agent imitates than how an agent imitates.

---

[3] For this dynamics the set of interior stationary state coincides with the set of interior Nash equilibria.

[4] If the equilibrium is nonhyperbolic.

5 Evolutionarily Stable Strategies in Asymmetric Games

A key concept in evolutionary game theory is that of an *evolutionarily stable strategy (ESS)* (see [5,13,14]).[5] We introduce a definition of ESS in the framework of our model where there are two populations playing an asymmetric normal form game (see [4] and [6]).

To introduce this concept, let us consider the following characterization: Let $N_{\tau j}$ be the total of $j-$strategists, $j \in \{m, nm\}$, in population, $\tau \in \{M, R\}$, and let $H = N_{Mm} + N_{Mnm} + N_{Rm} + N_{Rnm}$ be the total inhabitants in the country. Let us denote by $x_{1R} = \frac{N_{Rm}}{H}$, $x_{2R} = \frac{N_{Rnm}}{H}$, $x_{1M} = \frac{N_{Mm}}{H}$, $x_{2M} = \frac{N_{Mnm}}{H}$. Then, we denote by

$$\Delta = \left\{x \in R^4_+ : x_{1R} + x_{2R} + x_{1M} + x_{2M} = 1\right\}$$

the $(k-1)-$simplex of R_k in our case $k = 4$. Also, we introduce the symbolism: $x = (x^R, x^M)$ where $x^\tau = (x_{1\tau}, x_{2\tau})$; $\forall \tau \in \{M, R\}$.

Let us now introduce the following notation:

$$\bar{x}^R_m = \frac{H}{|R|}x_{1R},\ \bar{x}^R_{nm} = \frac{H}{|R|}x_{2R},\ \bar{x}^M_m = \frac{H}{|M|}x_{1M},\ \bar{x}^M_{nm} = \frac{H}{|M|}x_{2M},$$

where $|R|$ is the total number of agents in population R, $|M|$ is the total of agents in population M, and $\bar{x}^\tau = (\bar{x}^\tau_m, \bar{x}^\tau_{nm}),\ \forall\ \tau \in \{R, M\}$.

Definition 2. We say that $x \in \Delta$ is an *evolutionarily stable strategy (ESS)*, if and only if for every $y \in \Delta$, $y \neq x$ there exists some $\bar{\varepsilon}_y \in (0,1)$ such that for all $\epsilon \in (0, \bar{\varepsilon}_y)$ and with $w = \varepsilon y + (1-\varepsilon)x$ where $w = (w^R, w^M) = \varepsilon(y^R, y^M) + (1-\varepsilon)(x^R, x^M)$, then:

$$u^\tau(\bar{x}^\tau, \bar{w}^{-\tau}) > u^\tau(\bar{y}^\tau, \bar{w}^{-\tau}),\quad \forall\ \tau \in \{R, M\}.$$

Intuitively, we say that a distribution x is an ESS if the incumbent strategy x^τ does better in the post-entry population than the alternative strategy $y^\tau\ \ \forall\tau$. In our definition, agents of a given population do not play against agents in their own population, so the second-order conditions in the definition of an ESS become superfluous.

The next theorem shows the equivalence between our definition of ESS in the framework of asymmetric normal form games and the definition of strict Nash equilibrium.[6]

Theorem 1. *If the distribution* $x = (x_{1R}, x_{2R}, x_{1M}, x_{2M})$ *is an ESS in the sense of Definition 2, then the profile of mixed strategies* $\bar{x} = (\bar{x}^R, \bar{x}^M) = (\bar{x}^R_m, \bar{x}^R_{nm}, \bar{x}^M_m, \bar{x}^M_{nm})$ *is a strict Nash equilibrium for the two population normal*

[5] This concept was originally applied to Biology (see [7]).

[6] Recall that a Nash equilibrium $s = (x, y)$ is called strict if and only the profile s is the unique best reply against itself.

form game and conversely, for each strict Nash equilibrium $\bar{x}$, there exists a distribution x that is an ESS.

Proof. Let $S^\tau, \tau \in, \{M, R\}$ be the set of mixed strategies of the populations M and R, respectively. The theorem follows immediately from our definition of ESS (*Definition* 2), the definition of strict Nash equilibrium and the continuity of the functions $u_\tau : S^\tau \times S^{-\tau} \to R$. Note that from the definition of ESS it follows that $u_\tau(\bar{x}^\tau, \bar{x}^{-\tau}) > u_\tau(\bar{y}^\tau, \bar{x}^{-\tau})\ \forall\ y^\tau \neq x^\tau \in S^\tau$ and $\forall \tau$ so $\bar{x}^\tau$ is the only best response against $\bar{x}^{-\tau}\ \tau \in \{M, R\}$. □

Note that a distribution $x \in \Delta$ defines an interior stationary point for each one of the dynamic systems (5) if and only if

$$u^\tau(e_i, \bar{x}^{-\tau}) = u(\bar{x}^\tau, \bar{x}^{-\tau}).$$

This means that x^τ is a best response for population $\tau = \{M, R\}$ against $x^{-\tau}$ so, the profile $(x^\tau, x^{-\tau})$ is a Nash equilibrium. And conversely, if the profile $(x^\tau, x^{-\tau})$ is a Nash equilibrium, then it is a stationary point for these dynamic systems.

Theorem 2. *If the distribution $x = (x_{1R}, x_{2R}, x_{1M}, x_{2M})$ is an ESS in the sense of Definition 2, then the profile of mixed strategies $\bar{x} = (\bar{x}^R, \bar{x}^M)$ is a globally asymptotically stable stationary point of (5) or (2).*

Proof. Consider the relative entropy function $H_x : Q_x \to R$, where Q_x is the set of the distributions $y \in \Delta$ that assign positive probability to every state that has positive probability in x. In fact, Q_x is the union of the interior of the simplex Δ an the minimal boundary face containing x.

$$H_x(y) = -\sum_{i \in C(x)} x_i \ln\left(\frac{y_i}{x_i}\right) = -\sum_{i \in C(x^M)} x_i^M \ln\left(\frac{y_i^M}{x_i^M}\right) - \sum_{i \in C(x^R)} x_i^R \ln\left(\frac{y_i^R}{x_i^R}\right).$$

By $C(x)$ we denote the support of x, i.e the set of coordinates not equal to 0.

As it is easy to see, if x is ESS, then $H_x(y)$ verify:

1. $H_x(y) = 0$ if and only if $y = x$.
2. $\dot{H}_x(y) = -\sum_{i=1}^4 \dot{y}_i x_i =$

$$= -k\left[\left(u^M(x^M, y^R) - u^M(y^M, y^R)\right) + \left(u^R(x^R, y^M) - u^R(y^R, y^M)\right)\right],$$

where:

$$k = \begin{cases} \beta^\tau \text{ if the system is (2)}, \\ \\ 2\mu^\tau \text{ if the system is (5)}. \end{cases}$$

3. Thus, if x is an ESS, then $\dot{H}_x(y) < 0\ \forall\ y \in U \cap Q_x$ where U is a neighborhood of x.

□

In this context this theorem shows that, when a population has a distribution that is an ESS, if this composition is perturbed (in a relative neighborhood of x), then the performance of all strategies in this population decreases with respect to the original distribution x. In other words, if x is an ESS, then, not only is it the best distribution within each group, given the behavior of the population in the given country, it is also a stable strategy in the sense that every change in the distribution of the behavior within a group implies worse payoff for the deviating group. Moreover, if x is an ESS and the corresponding $\bar{x}_m = (\bar{x}_m^R, \bar{x}_m^M)$ is an interior stationary point for the dynamic system given by (2) or by (5), it follows that for every solution $\xi(t, z_0)$ of these systems of equations, with initial conditions z_0 in the interior of the square $[0, 1] \times [0, 1]$ implies that $\xi(t, z_0)_{t \to \infty} \to \bar{z}$. This means that $\bar{z}$ is a globally asymptotically stable stationary point, in the sense of attracting all interior initial state. Hence, if the distribution of the population $x = (x_{1R}, x_{2R}, x_{1M}, x_{2M})$ in $t = t_0$ is perturbed to a new distribution y, then the evolution of this perturbed distribution, whose state in each time t is denoted by $\phi(t, y)$, where $\phi(t_0, y) = y$, evolves according to one of the dynamic systems (2) or (5) , and y will always approach the initial distribution x.

6 Concluding Remarks

We showed that imitative behavior can be represented by the replicator dynamics with two specific cases: (a) pure imitation driven by dissatisfaction where all reviewing agents imitate the first person they encounter, and (b) the successful reviewing strategist chooses the actions with the best payoffs, with a probability proportional to the expected payoff.

As long as the linear approximation is mathematically valid, the second differs from the first, yet we obtain similar evolutionary dynamics for the two models. Similar conclusions can be applied to the relationship between ESS, Nash equilibrium, and the Replicator Dynamics.

Hence, we conclude that for imitation rules that use random sampling, *"whom an agent imitates is more important than how an agent imitates"*. We justify the extension of the simpler model based on a description of an observed behavior: agents choose an expected maximization of benefits, since *they do what others do when they can imitate successful strategies*.

Finally, we introduced the concept of ESS for a model of two asymmetric populations and we show that for every ESS there is a strict Nash equilibrium and conversely.

Acknowledgements Elvio Accinelli thanks the partial support obtained from the project C07-FAI-11-46.82 of the UASLP Secretara de Posgrado. Many thanks to the International Society of Dynamic Games. Thanks to Simon Halliday, Krzysztof Szajowski, and two anonymous referees for their helpful comments. The usual disclaimer applies. Juan Gabriel Brida thanks the partial support obtained from the project "The Contribution of Tourism to Economic Growth" financed by the Free University of Bolzano.

References

1. Alos-Ferrer, Weidenholzer, S.: Imitation, Local Interactions and Efficiency. Economics Letters **93**(2), 163–168 (2006)
2. Apesteguia, J., Huck, S., Oechssler, J.: Imitation-theory and experimental evidence. Journal of Economic Theory **136**, 217–235 (2007)
3. Björnerstedt, J., Weibull, J.: Nash equilibrium and evolution by imitation. In: Arrow, K. et al. (eds.) The Rational Foundations of Economic Behaviour. Macmillan, Hampshire (1996)
4. Hines, W.G.S.: Three Characterizations of Population Strategy Stability. Journal of Applied Probability **17**, 333–340 (1980)
5. Hofbauer, J., Sigmund, K.: Evolutionary Games and Population Dynamics. Cambridge University Press, Cambridge (2002)
6. Rice, S.H.: Evolutionary Theory: Mathematical and Conceptual Foundations. Sinauer Associates, Sunderland (2004)
7. Maynard Smith, J.: The theory of games and the evolution of animal conflict. Journal of Theoretical Biology **47**, 209–222 (1974)
8. Samuelson, L., Zhang, J.: Evolutionary Stability in Asymmetric Games. Journal of Economic Theory **57**, 363–391 (1992)
9. Schlag, K.H.: Why Imitate, and if so, How? A Boundedly Rational Approach to Multi-Armed Bandits. Journal of Economic Theory **78**(1), 130–156 (1998)
10. Schlag, K.H.: Which One Should I Imitate. Journal of Mathematical Economics **31**(4), 493–522 (1999)
11. Taylor, P.: Evolutionarily stable strategies with two types of player. Journal of Applied Probability **16**, 76–83 (1979)
12. Vega-Redondo, F.: The evolution of Walrasian behavior. Econometrica **65**, 375–384 (1997)
13. Vega-Redondo, F.: Evolution, Games, and Economic Behavior. Oxford University Press, Oxford, New York (1996)
14. Weibull, W.J.: Evolutionary Game Theory. MIT, Cambridge (1995)

ESS, Population Games, Replicator Dynamics: Dynamics and Games if not Dynamic Games

Pierre Bernhard

Abstract We review some classical definitions and results concerning Evolutionarily Stable Strategies (E.S.S.) with special emphasis on their link to Wardrop equilibrium, and on the nonlinear case where the fitness accrued by an individual depends nonlinearly on the state of the population. On our way, we provide a simple criterion to check that a linear finite dimensional Wardrop equilibrium – or Nash point in the classical E.S.S. literature – satisfies the second-order E.S.S. condition. We also investigate a bifurcation phenomenon in the replicator equation associated with a population game. Finally, we give two nontrivial examples of Wardrop equilibria in problems where the strategies are controls in a dynamic system.

1 Introduction

Wardrop equilibrium, E.S.S., and related concepts form the game theoretic foundation of the investigation of population dynamics under evolution or learning behaviors as depicted by the replicator equation or more generally adaptive dynamics.

Pioneered by John Glenn Wardrop in the context of road traffic as far back as 1952, [21] these concepts have been given a new impetus after their re-discovery and extension by John Maynard-Smith and co-workers in the mid seventies, [10, 12] in the context of theoretical biology and evolution theory. The introduction by Taylor and Jonker [19] of the replicator equation gave its solid mathematical grounds to the intuition of stability present from the inception. Since then, a large body of literature has appeared under the generic name of evolutionary game theory. See e.g. [7, 9, 20, 22]. Nowadays, routing problems have become a hot topic again with the advent of the INTERNET and, more recently, ad hoc networks, together with learning in populations. Old topics such as optimal transportation [4] have been

P. Bernhard (✉)
University of Nice-Sophia Antipolis and INRIA Sophia Antipolis Méditerranée, France
e-mail: Pierre.Bernhard@sophia.inria.fr

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_15,

renewed by these new problems. All share the characteristic that they investigate the collective effect of rational, usually selfish, behavior of individual agents on a large population of similar ones. We actually witness the emergence of a new population dynamics paradigm, where games are one of the underlying fundamental tools.

The aim of this paper is to better disseminate those ideas in the traditional Dynamic Games community, stressing among other things the links with transportation theory, and maybe offering some new results on our way.

2 Wardrop Equilibrium and E.S.S.

Let us recall the foundations of Evolutionary Stable Strategy theory, and in that process define our notation.

2.1 Generating Function and Fitness Function

The setup introduced here is rather general, covering both finite and infinite strategy sets, as well as linear and nonlinear generating functions. Yet, many results are restricted to finite strategy sets, or further to the *matrix case*: finite strategy set and linear generating function.

2.1.1 Population and Strategy Distribution

A large population of identical agents interact with each other. Each chooses a strategy $x \in X$ in a given *strategy set* X. The nature of the set X is an important feature of the theory. In the first part of this article, we consider the cases $X = \{x_1, x_2, \dots, x_n\}$ finite or $X \subset \mathbb{R}^n$ a compact subset of $\mathbb{R}^n$. The notation $\mathrm{d}x$ will mean the Lebesgue measure if X is continuous and the discrete measure if it is finite, so that in the latter case, $\int_X f(x)\mathrm{d}x = \sum_{i=1}^n f(x_i)$. In the last part, we shall consider an infinite dimensional strategy set.

We are interested in the share or *proportion* of agents using each strategy, say $p(x)$. Technically, p is a positive measure of mass one over X, and can therefore be viewed as a probability measure. We let $\Delta(X)$ denote the set of such probability measures over X. It is finite dimensional if and only if X is finite. The measure $p(A) = \int_A p(\mathrm{d}x)$ of a subset $A \subset X$ is the probability that an agent picked "at random" with uniform probability over the population use a strategy of the subset A of strategies. Then the mean value of a function f is $\mathbb{E}^p f = \langle p, f \rangle = \int_X f(x)p(\mathrm{d}x)$.

If X is finite, we shall let $p(x_i) =: p_i$, we identify the measure p with the vector of $\mathbb{R}^n$ with components p_i, and any scalar function f over X as the vector with components $f_i = f(x_i)$. Then p belongs to the simplex $\Sigma_n \subset \mathbb{R}^n$ identified with $\Delta(X)$, and $\mathbb{E}^p f = \langle p, f \rangle = \sum_{i=1}^n p_i f(x_i)$.

In applications to road traffic, or to routing of messages in a communication network, x is a choice of route linking the desired origin and destination points in a route network. In evolutionary biology, x is viewed as a phenotypic *trait* resulting from an underlying genotypic structure. Many other examples can be given.

The biological literature contains interesting discussions of the difference between the setup we just described, considered a genetic polymorphism, and a population, say with a finite set X of possible behaviors, where all agents have the same genotype, dictating a probabilistic – or *mixed* – choice of behavior, so that they all have the same probabilities p_i of behaving according to "strategy" x_i, for $i = 1, 2, \dots, n$. As a result, due to the law of large numbers, in such a genetically monomorphic population, the *proportion* of animals adopting the behavior x_i at each instant of time will be p_i, although the individuals using each x_i may vary over time.

While these discussions are highly significant in terms of biologic understanding – e.g. some mixtures might be impossible to produce with a genetically monomorphic population, or the transmission of the behavior over generations may be different in both cases, the more so if sexual reproduction is involved – we shall not be concerned with them. The populations are statistically the same in both cases. As a consequence, their dynamics will be considered the same.[1]

2.1.2 Fitness and the Generating Function

The various possible strategies induce various benefits to their users. It is assumed all along that

(a) There exists a scalar, real, measure of reward for each participant. In evolutionary biology, this measure, called "fitness", may be reproductive efficiency – the excess of birth rate over the death rate per animal. In the example of road traffic, it will be the opposite of the time spent to reach its destination, etc.
(b) This reward is a function of the strategy used by the particular agent and the *state* of the population, defined as the *proportions*, the measure p.

Hence, an agent using strategy x in a population in a state p has a reward – a fitness – $G(x, p)$. We call G the *generating function*. It is assumed measurable in x, uniformly bounded and weakly continuous in p. We shall also use the notation $G(p)$ to mean the function $x \mapsto G(x, p)$, hence in the finite case, the vector with components $G_i(p) = G(x_i, p)$, $i = 1, \dots, n$.

A special case of interest is the *linear case* where $p \mapsto G(x, p)$ is linear. This is the case if, say, there is a reward $H(x, y)$ to an agent of type x meeting an agent of type y, and "meeting" happens at random with uniform probability, so that the average fitness of an agent of type x is

$$G(x, p) = \int_X H(x, y) p(\mathrm{d}y). \tag{1}$$

[1] A difference in the stability analysis mentioned by [11] is due to a questionable choice of dynamics in the discrete time case, which we will not follow.

An individual with a mixed strategy q, or equivalently a subpopulation in a state q, in a population of overall state p will get an average fitness

$$F(q,p) = \langle q, G(p)\rangle = \int_X G(x,p)q(\mathrm{d}x).$$

It follows from this definition that F is always linear with respect to its first argument, but not necessarily with respect to the second one. We shall refer to the *linear case* to mean linearity of G w.r.t. p, hence of F w.r.t. its second argument.

2.2 *Stable Population States*

2.2.1 Wardrop Equilibrium

Consider a population in a state p. Assume that in that population, some individuals mutate, creating a small subpopulation of total relative mass ε and state q. The overall population now is in a state

$$q_\varepsilon = \varepsilon q + (1-\varepsilon)p.$$

We say that the subpopulation *invades* the original population if[2] $F(q,q_\varepsilon) \geq F(p,q_\varepsilon)$. The original population will be considered evolutionarily stable if it is protected against invasion by any (single)[3] mutation. We therefore state:

Definition 1. A distribution p is an *Evolutionarily Stable Strategy (E.S.S.)* if

$$\forall q \in \Delta(X), \exists \varepsilon_0 : \forall \varepsilon \leq \varepsilon_0, \quad F(q,q_\varepsilon) < F(p,q_\varepsilon). \tag{2}$$

It follows from the dominated convergence theorem that the fitness function F inherits the continuity of G with respect to its second argument. Therefore, letting ε go to zero, one immediately sees that a necessary condition for p to be an E.S.S. is

$$\forall q \in \Delta(X), \quad F(q,p) \leq F(p,p). \tag{3}$$

J.-G. Wardrop [21] was considering a population of drivers in a road network. Their strategy is the choice of a route in the network. The time of travel is a function of the occupation of the route chosen. And if the total population is fixed, this is a

[2] As a mathematician, we take this as our *axiomatic definition* of invading. The relationship to the biological concept as well as the choice of a large inequality here is a discussion left to the biologists. Our convention here is that of the biological literature [11].

[3] The case of simultaneous mutations is more complex.

function of the proportion q_i of that population that uses the route considered. We quote from [21]:

> Consider the case of a given flow of traffic Q which has the choice of D alternative routes from a given origin to a given destination, numbered $1, 2, \ldots D$. [...] Consider two alternative criteria based on these journey times which can be used to determine the distribution on the routes, as follows
>
> (a) The journey times on all routes actually used are equal, and less than those which would be experienced by a single vehicle on any unused route.
>
> (b) The average journey time is minimum.
>
> The first criterion is quite a likely one in practice, since it might be assumed that traffic will tend to settle down into an equilibrium situation in which no driver can reduce his journey time by choosing a new route. [...]

Notice that the argument developed is one of stability. It should be further noticed that his "generating function" $t(q)$, the travel time on a route, is nonlinear, of the form $t_i = b_i/(1 - q_i/p_i)$ for constants b_i and p_i that depend on the network and the rest of its usage.

To extend Wardrop's definition to an infinite set X, we are led to introduce

$$\forall p \in \Delta(X), \quad Z(p) = \{x \mid G(x, p) < \max_{y \in X} G(y, p)\},$$

and we end up with the *Wardrop condition*, which is equivalent to condition (3):

$$p(Z(p)) = 0. \tag{4}$$

In recognition of Wardrop's anteriority, and following the standard terminology in routing theory, we let

Definition 2. A distribution p satisfying (3), or equivalently (4), is called a *Wardrop equilibrium*.

Link with Nash equilibrium One may notice that (3) is equivalent to stating that (p, p) is a (symmetric) Nash equilibrium of the two-player game, where the rewards J^1 and J^2 of the players are $J^1(q^1, q^2) = F(q^1, q^2)$, $J^2(q^1, q^2) = F(q^2, q^1)$. As a consequence, most of the literature on evolutionary games uses the phrase "Nash equilibrium" where we follow the usage of the transportation literature with "Wardrop equilibrium". Recognizing a Nash equilibrium makes the equivalence of (3) and (4) a trivial extension of Von Neumann's *equalization theorem* [13, 14], although in the biology literature, it is usually attributed to [6].

2.2.2 Second Order E.S.S. Condition

Not all Wardrop equilibria are E.S.S., because (3), or equivalently (4), is only a necessary condition for (2) to hold. Let the *best response map* $B(\cdot)$ be defined as

$$\forall p \in \Delta(X), \quad B(p) = \{r \in \Delta(X) \mid F(r, p) = \max_{q \in \Delta(X)} F(q, p)\}.$$

Notice that $B(p)$ is precisely the set of distributions r that satisfy $r(Z(p)) = 0$, so that Wardrop's condition (4) also reads $p \in B(p)$.

Proposition 1. *In the linear case, a Wardrop equilibrium p is an E.S.S. if and only if*

$$\forall q \in B(p)\backslash\{p\}, \quad F(p,q) > F(q,q). \tag{5}$$

or equivalently

$$\forall q \in B(p)\backslash\{p\}, \quad F(q-p,q-p) < 0. \tag{6}$$

Proof. The proof is elementary and can be found in [18].

Notice that $F(r,r)$ is a quadratic form. Therefore, this last condition clearly shows that the second E.S.S. condition (5) is in fact a second-order condition.

We stay for the time being with the linear case. Let a Wardrop equilibrium p be fixed, and $X_1(p) = X - Z(p)$, a set of measure one for distributions in $B(p)$. Let H_1 be the restriction of H to $X_1(p) \times X_1(p)$, and F_1 be the corresponding bi-linear function for measures (not necessarily positive) over $X_1(p)$. Let also $\mathbf{1}$ be the constant function equal to one over X (or over $X_1(p)$); and notice that $p - q$ in (6) is orthogonal to $\mathbf{1}$ in the space $\mathscr{M}(X_1)$ of measures over X_1. Let also X_2 be the support of p. From Wardrop's condition, it follows that $X_2 \subset X_1$. Define similarly the restriction H_2 of H to $X_2 \times X_2$ and $\mathscr{M}(X_2)$. We therefore have[4]:

Theorem 1. *In the linear case, a Wardrop equilibrium p is an E.S.S.*

(a) if the restriction of the quadratic form $F_1(r,r)$ to measures $r \in \mathbf{1}^\perp \subset \mathscr{M}(X_1)$ is negative definite,
(b) only if the restriction of the quadratic form $F_2(r,r)$ to measures $r \in \mathbf{1}^\perp \subset \mathscr{M}(X_2)$ is negative definite.

Proof. The sufficient condition derives trivially from the proposition. The restriction to X_2 in the necessary condition insures that p be in the relative interior of $\Delta(X_2)$, hence making it possible to violate the inequality in (2) if the quadratic form is not negative definite.

Concerning the nonlinear case, we need to introduce the following notation:[5] define $H(x,p) = \mathrm{D}_2 G(x,p)$ the derivative of $p \mapsto G(x,p)$, and its restriction H_2 to $X_2 \times \Delta(X_2)$. It follows from the proof of Theorem 3 below the following:

Theorem 2. *A necessary condition for a Wardrop equilibrium p to be an E.S.S. is that the restriction of the quadratic form $\langle r, H_2(p)r\rangle$ to $r \in \mathbf{1}^\perp \subset \mathscr{M}(X_2)$ be nonpositive definite.*

---

[4] It would suffice that negativity be required of the quadratic form $\langle q_1 - p_1, H_1(q_1 - p_1)\rangle$ for all $q_1 \in \Delta(X_1)$, which is less demanding. However, we offer no simple check of that property, it is why we stick with the definition given here.

[5] There is a slight overload of notation for H w.r.t. the linear case. It is resolved if we accept that $D_2 G(x,q)r = H(x,q)r = \int_X H(x,y,q)r(\mathrm{d}y)$.

Proof. Introduce the (negative) *score function* [6]

$$\mathscr{E}(\varepsilon, q) = F(q, q_\varepsilon) - F(p, q_\varepsilon),$$

and notice that $B(p) = \{q \mid \mathscr{E}(0, q) = 0\}$. Because F is always linear w.r.t. its first argument, it follows that $D_1\mathscr{E}(0, q) = \langle q - p, H(p)(q - p)\rangle$. Now, p is interior to $\Delta(X_2)$. Choosing q also in $\Delta(X_2)$, we get $D_1\mathscr{E}(0, q) = \langle q - p, H_2(p)(q - p)\rangle$, with $p - q$ in any direction. The result follows.

We may give a somewhat more explicit condition, but it is not clear that it be more useful. Let μ be the Lebesgue measure and $\mu_1 = \mu(X_1)$. Given any measure r over X_1, let $r(X_1) = r_1$, the measure $\hat{r} = r - (r_1/\mu_1)\mu$ is orthogonal to $\mathbb{1}$. So that the first condition of our first theorem, e.g., is that, for any r not colinear to μ in $\Delta(X_1)$,

$$F_1(r, r) - \frac{r_1}{\mu_1}(F_1(r, \mu) + F_1(\mu, r)) + \frac{r_1^2}{\mu_1^2}F_1(\mu, \mu) < 0.$$

This is again a quadratic form over $\Delta(X_1)$. It could be translated in terms of H_1. This does not seem very useful, except in the (finite linear) matrix case.

2.2.3 Matrix Case

Assume now that $X = \{x_1, x_2, \dots, x_n\}$ and that $G(p)$ is linear in p. We identify the function H of (1) with the matrix with elements $H_{ij} = H(x_i, x_j)$. We now have $G(p) = Hp$, and $F(q, p) = \langle q, Hp\rangle = q^t Hp$.

A distribution $p \in \Sigma_n$ is a Wardrop equilibrium if and only if, up to a reordering of the coordinates (hence of the x_i), it can be written as a composite vector with $p_1 \in \Sigma_{n_1} \subset \mathbb{R}^{n_1}$,

$$p = \begin{pmatrix} p_1 \\ 0 \end{pmatrix} \qquad \text{with furthermore} \qquad H\begin{pmatrix} p_1 \\ 0 \end{pmatrix} = \begin{pmatrix} f\,\mathbb{1} \\ z \end{pmatrix}$$

for the same decomposition, with $f \in \mathbb{R}$ and $z_j < f$ for every coordinate z_j of z. Here, $Z(p)$ is the set $\{x_{n_1+1}, \dots, x_n\}$, $B(p)$ is the set of all $q \in \Sigma_n$, which share the same decomposition $(q_1, 0)$ as p, and $F(p, p) = f$.

To investigate the second-order condition, we partition H according to the same decomposition displaying the restriction H_1 of H to $X_1 \times X_1$:

$$H = \begin{pmatrix} H_1 & H_{10} \\ H_{01} & H_{00} \end{pmatrix}.$$

[6] We borrow this name from [22].

The sufficient condition of Theorem 1 is now that the restriction of the quadratic form $\langle r_1, H_1 r_1\rangle$ to vectors r_1 orthogonal to the vector $\mathbb{1}$ of $\mathbb{R}^{n_1}$ be negative definite. Furthermore, call p_2 the subvector of the strictly positive coordinates of p_1. The necessary condition of Theorem 1 is that the corresponding submatrix H_2 generate a nonpositive definite restriction of the quadratic form $\langle r_2, H_2 r_2\rangle$ to vectors r_2 orthogonal to $\mathbb{1}$ (with appropriate dimension).

We now give a practical means of checking these conditions. Given a $m \times m$ matrix A, define $\sigma(A)$ as the $m-1 \times m-1$ matrix obtained by replacing each block of four adjacent entries

$$\begin{pmatrix} a_{i,j} & a_{i,j+1} \\ a_{i+1,j} & a_{i+1,j+1} \end{pmatrix}$$

in A by their *symmetric difference* $a_{i,j} - a_{i,j+1} - a_{i+1,j} + a_{i+1,j+1}$.

Proposition 2. *The restriction of the quadratic form $\langle r, Ar\rangle$ to the vectors r orthogonal to $\mathbb{1}$ is negative definite (resp. nonpositive definite) if and only if $\sigma(A) + \sigma(A)^t$ is negative definite (resp. non-positive definite).*

Proof. $\sigma(A) = P^t A P$ where P is the $m \times m-1$ injective matrix

$$P = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ -1 & 1 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{pmatrix}$$

whose range space is the vector subspace orthogonal to the vector $\mathbb{1}$.

2.3 Further Stability Concepts

Many variations of the concept of E.S.S. have been proposed. The only one we mention here is sometimes called *Evolutionary Robust Strategy* or E.R.S.[7]

Definition 3. A strategy distribution p is called an E.R.S. if there exists a weak neighborhood $\mathcal{N}$ of p such that

$$\forall q \in \mathcal{N} \setminus \{p\}, \quad F(p,q) > F(q,q).$$

Hence, in effect, for an E.R.S., the dominance inequality of (5) is requested of all q in a neighborhood and not only of the best response q's.

[7] It is closely related to J. Apaloo's concept of *Neighborhood Invader Strategies* [3].

Concerning the relationship of this concept with E.S.S., we recall the notation $H(x, p) = \mathrm{D}_2 G(x, p)$ the derivative of G with respect to p, and its restrictions H_1 to $X_1 \times \mathscr{M}(X_1)$. We state:

Definition 4. A Wardrop equilibrium p is said *regular* if there exists a $\Delta(X_1)$-neighborhood of p such that over that neighborhood $q \mapsto H_1(x, q)$ is Lipschitz continuous for all $x \in X$ and the restriction of the quadratic form $\langle r, H_1(q)r\rangle$ to measures r orthogonal to $\mathbf{1}$ in $\mathscr{M}(X_1)$ is negative definite, both uniformly[8] in q.

We have the following

Theorem 3.

(a) All E.R.S. are E.S.S.
(b) In the finite case, all regular Wardrop equilibria, and all E.S.S. in the linear case, are E.R.S.

The proof of *(a)* is elementary: applying the E.R.S. condition to $q_\varepsilon = p + \varepsilon(q - p)$ and using the linearity of F w.r.t. its first argument, we get condition (2). The proof of *(b)* is classical for the matrix case (see [9]). We extend it to the nonlinear regular case in the appendix.

The importance of the concept of E.R.S. for us lies in the fact that it is a (local) Lyapunov asymptotically stable point of the replicator dynamics.

2.4 Clutch Size Determination and Braess'Paradox

The transportation literature is familiar with Braess'paradox, seldom quoted in the E.S.S., biologically inspired literature. This generically refers to situations where improving the quality of the resource decreases the fitness of every individuals in the population. The following is such an example.

A species of parasitoids lays its eggs in its *hosts*, say the eggs of another species. This is a *gregarious* parasitoid, meaning that several offspring can be born from a single host. Yet, the probability that the parasitizing succeed, actually giving parasitoid newborns, decreases with the number of eggs in the host. Either all eggs laid in any host succeed, or all fail because the host died too early. The population of parasitoids is such that, as a first approximation, it may be assumed that exactly two parasitoids will lay some eggs in each host. We are interested in the number of eggs, or *clutch size*, x that the female parasitoids lay.

We consider a very simple case where only three parasitoids can be born from a single host. We assume that if only two parasitoid eggs are laid in a single host, they will survive (with probability one). If three are laid, the survival probability is π, and we assume $\pi > 1/3$. The only two sensible (pure) strategies are laying one or

[8] We shall use this definition in the finite case only, it is why we need not assume any regularity w.r.t. x.

two eggs, since at least one more will be present in the host due to superparasitism. The game matrix of this problem is as follows:

$$H = \begin{pmatrix} 1 & \pi \\ 2\pi & 0 \end{pmatrix}.$$

The Wardrop equilibrium is $p(\{x = 1\}) = \pi/(3\pi - 1)$. It is an E.S.S., since $\sigma(H) = 1 - 3\pi < 0$. The fitness of every individuals at equilibrium is

$$G_1(p) = G_2(p) = F(p, p) = \frac{2\pi^2}{3\pi - 1}.$$

For $\pi = 1/2$, this leads to $F(p, p) = 1$, while improving the survival probability of a group of three eggs to $\pi = 2/3$ leads to $p = 2/3$ and $F(p, p) = 8/9 < 1$. This is because raising the survival probability of three eggs attracts more animals to the strategy 2, making the outcome 0, the "waste of eggs", more frequent.

3 Replicator Dynamics

We shall consider only the finite case. The infinite case is far less well known. Some results concerning it, both old an new, can be found in [18].

3.1 Evolutionary Dynamics

Assume that "fitness" $G_i(p)$ measures the excess (may be negative) of the birth rate minus the death rate per individuals using the strategy x_i (with phenotype x_i) in a population of strategy distribution p. Let $n_i(t)$ be the number of individuals using x_i at time t. We also have for the strategy distribution

$$q_i(t) = \frac{n_i(t)}{\sum_k n_k(t)}.$$

If generations are discrete, with a time-step h, – say a population reproducing once a year – this yields in the continuous limit for the n_i's together with the definition of $q_i(t)$ [9]

$$q_i(t + h) = q_i(t) \frac{1 + hG_i(q(t))}{1 + hF(q(t), q(t))}.$$

[9] The classical theory summarized here departs from the discrete time dynamics of [11].

Looking at these dynamics over a very large time horizon, i.e. taking a larger and larger time unit, is equivalent to letting h go to zero. And the above equation converges to the *replicator dynamics*

$$\dot{q}_i = q_i[G_i(q) - F(q,q)]. \tag{7}$$

We summarize here the classical invariance and asymptotic properties of these equations:

Theorem 4.

(a) The replicator dynamics (7) leave each face of Σ_n invariant.
(b) All limit points of their trajectories are Wardrop equilibria.
(c) In finite dimension, all E.R.S. are locally asymptotically stable.

Proof. The proof is classical, and can be found, e.g., in [22] or [9].

This equation has been extensively used as a simple model of population dynamics under evolution [9, 11, 22] and also as a Nash equilibrium selection device [16].

3.2 Population Games

The ideas embedded into the derivation of the replicator dynamics may be extended to situations where several different populations are interfering, in effect, to Nash equilibria of classical games as opposed to Wardrop equilibria. These have been called "Population games" by W. Sandholm [17].

3.2.1 Population Game Dynamics

Let N populations interact. We denote the population number with upper indices $K, L = 1, 2, \dots, N$. Individuals of population K have n^K possible strategies (phenotypes). As previously, we let n_i^K be the number of individuals of population K using the strategy i, q_i^K be the proportion of such individuals in population K, and q^K be the n^K-dimensional vector of the q_i^K, $i = 1, 2, \dots, n^K$. We shall also make (parsimonious) use of the classical notation q^{-K} to mean the set of all q^L for $L \neq K$, and let q be the set of all q^K's. It is assumed that the fitness of an individual of population K using strategy i is a function $G_i^K(q^{-K})$, so that the collective fitness of population K is $F^K(q) = F^K(q^K, q^{-K}) = \langle q^K, G^K(q^{-K})\rangle$.

W. Sandholm has investigated several mechanisms by which the individuals may update their choices of strategies. Several of the most natural schemes lead to the dynamics

$$\dot{n}_i^K = n_i^K G_i^K(q^{-K}) \quad \text{and hence} \quad \dot{q}_i^K = q_i^K[G_i^K(q^{-K}) - F^K(q)].$$

3.2.2 Wolves and Lynxes

We restrict now our attention to linear two-population games with two strategies available in each population. This is the simplest possible case, and it already exhibits interesting features.

Let therefore H^1 and H^2 be two 2×2 matrices,

$$H^K = \begin{pmatrix} a^K & b^K \\ c^K & d^K \end{pmatrix}$$

so that the payoffs are $G^K(q^L) = H^K q^L$. The underlying two-player game is defined by the bi-matrix

$$\begin{array}{c|c|c|}
x^1 \backslash x^2 & 1 & 2 \\ \hline
1 & a^1 \quad a^2 & b^1 \quad c^2 \\ \hline
2 & c^1 \quad b^2 & d^1 \quad d^2 \\ \hline
\end{array}$$

Furthermore, the vectors q^K, $K = 1, 2$, will be represented by their first component, i.e. again with a transparent abuse of notation

$$q^K = \begin{pmatrix} q^K \\ 1 - q^K \end{pmatrix}.$$

We let σ^K be the symmetric difference $a^K - b^K - c^K + d^K$, and whenever $\sigma^L \neq 0$, define $p^K = (d^L - b^L)/\sigma^L$, $L \neq K$.

We are here interested in the case where both p^K exist and lie in $(0, 1)$. They constitute a mixed Nash equilibrium. It is a simple matter to see that if both σ^K have the same sign, there are in addition two pure Nash equilibria, diagonally opposite in the square $[0, 1] \times [0, 1]$, while if they are of opposite sign, this is the only Nash equilibrium.

The replicator dynamics read

$$\dot{q}^K = \sigma^K q^K (1 - q^K)(q^L - p^L), \qquad L = 3 - K.$$

Their behavior is characterized in part by the following result (see [5, 9] for a more detailed analysis)

Theorem 5. *If $\sigma^1 \sigma^2 < 0$, the point (p^1, p^2) is a center of the replicator dynamics, and all trajectories are periodical. If $\sigma^1 \sigma^2 > 0$, the point (p^1, p^2) is a saddle, and the two pure Nash equilibria are asymptotically stable.*

Proof. Consider the relative entropies:

$$U^K(q^K) = p^K \ln \frac{p^K}{q^K} + (1 - p^K) \ln \frac{1 - p^K}{1 - q^K}$$

and the function

$$V(q^1, q^2) = \sigma^2 U^1(q^1) - \sigma^1 U^2(q^2).$$

A straightforward computation shows that its Lagrangian derivative is null. The U^i's are convex, and go to infinity as p^K approaches 0 or 1. If $\sigma^1\sigma^2 < 0$, V is either convex or concave, diverging to plus or minus infinity as (q^1, q^2) approaches the boundary of the domain of definition. The trajectories are level curves of V, which in that case are the boundaries of convex level sets contained in the domain. If $\sigma^1\sigma^2 > 0$, the curve $V(q^1, q^2) = 0$ separates the attraction basins of the two pure Nash equilibria.

We provide an example in a "hawk and doves" type of games, but between two populations sharing the same preys, say wolves and lynxes hunting deer, but where contests occur only between individuals of different species. Each have two possible behaviors, H for "Hawkish" and D for "Dovish".

In that model, Lynxes are at a trophic level above that of wolves. In particular, if two aggressive individuals meet, the lynx is hurt, but the wolf is killed. We also assume that against a pacific (coward) wolf, an aggressive lynx gets less than 1 (the full benefit of the prey), because it has spent unnecessary time and effort chasing a competitor who would have fled anyhow.

The result is the following bi-matrix of rewards:

$L\backslash W$	D	H
	$1-\lambda$	1
D	λ	0
	0	$-\theta$
H	$1-\mu$	$1-\nu$

with $\lambda + \mu > 1 > \nu > 0$ and $\theta > 0$. In that game, we have $\sigma_1 = \lambda + \mu - \nu$ and $\sigma_2 = -\lambda - \theta$, and the Nash strategies $p^1 = \theta/(\lambda + \theta)$, $p^2 = (1 - \nu)/(\lambda + \mu - \nu)$. Figure 1 shows a typical trajectory, computed in the case $\lambda = \nu = 1/2$, $\theta = 2\mu = 1.5$, and initial state $(0.2, 0.2)$.

3.2.3 Joint Interest and Bifurcation

Maynard-Smith [11] considers the case where both populations share a joint foe, say Man. Then, there is a benefit for each population in maintaining the other one, as it contributes to keep that foe away. This is in effect an equivalent in population games of the concept of inclusive fitness in E.S.S. See [8] and the references therein. We address a side-topic on these lines, staying with our simple Wolves and Lynxes game.

Several ways of taking into account that joint interest have been proposed in the literature, depending on the detailed mechanisms at work. One possibility is by

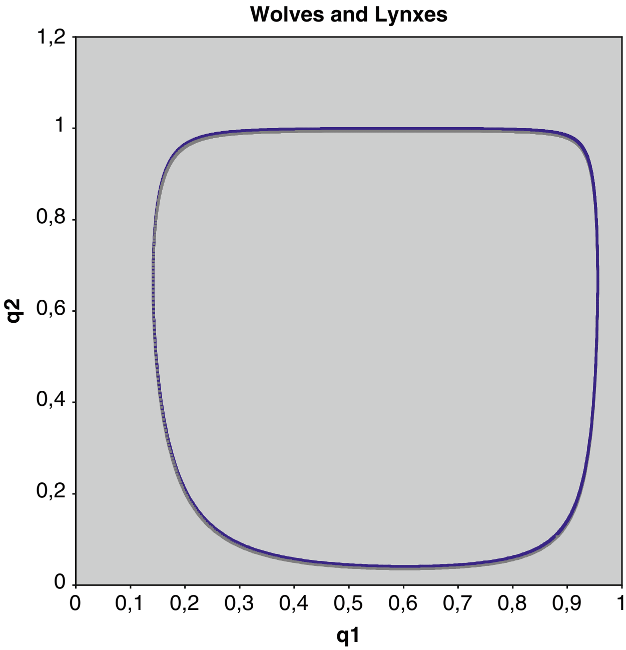


Fig. 1 Population dynamics for Wolves and Lynxes, $\lambda = \nu = 1/2$, $\theta = 2\mu = 1.5$, time span: 40 units

saying that each population K has a fitness $F_\alpha^K(q) = (1-\alpha)F^K(q) + \alpha F^L(q)$ for some coefficient $\alpha \in [0, 1/2]$. Then, we get

$$F_\alpha^K(q) = (1-\alpha)\langle q^K, H^K q^L\rangle + \alpha\langle q^L, H^L q^K\rangle = \langle q^K, H_\alpha^K q^L\rangle$$

with $H_\alpha^K = (1-\alpha)H^K + \alpha(H^L)^t$.

Now, the behavior of the replicator equation obviously depends on the coefficient α measuring the amount of joint interest. As an example, we work out the Wolves and Lynxes problem with $\lambda = \nu = 1/2$, $\theta = 2\mu = 3/2$.

Figure 2 summarizes the bifurcation that appears at $\alpha = 1/6$. (Third sketch.)

The conclusion is that a small variation in the parameter r may cause a dramatic difference in the qualitative behavior of the dynamics. That parameter is, at best, difficult to estimate. But a reverse use of this theory can be made, deriving from the observed behavior bounds for possible values of α if this model is to be used.

Notice that in the last case, α large, there is a single Nash equilibrium at $(1, 0)$. We are not in the case $(p_1, p_2) \in (0, 1)^2$, where pure Nash equilibria come in pairs.

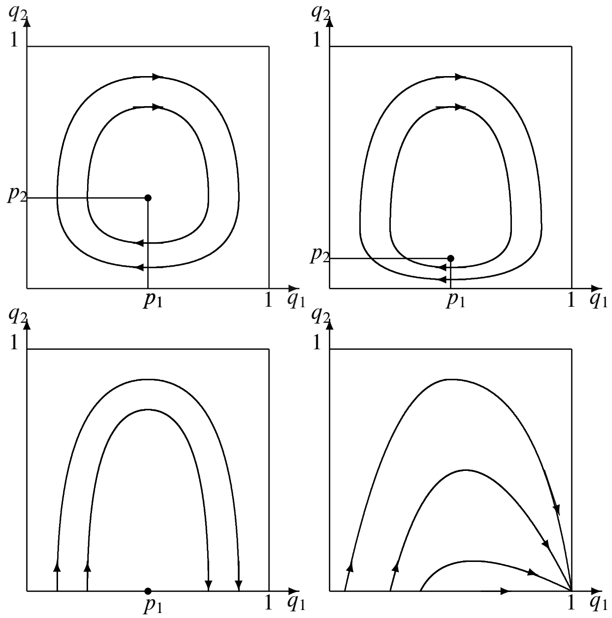


Fig. 2 The bifurcation as the joint interest raises, left to right and top to bottom, from 0 to 1/2

4 Infinite Dimensional Strategy Space

We now venture into an example in the realm of Wardrop equilibria with infinite dimensional trait spaces. Many examples can be found in the literature, see e.g. [8]. Examples in the realm of routing in massively dense communication network can be found in [2] and [1].

4.1 A General Resource Allocation Problem

In this example, remotely inspired by [15], individuals of a population may use two strategies, say $x \in \{0, 1\}$. However, they have to make that choice at each instant of time $t \in [0, T]$ over a fixed horizon T. Hence, the strategy set is the set of (measurable) functions $x(\cdot) : [0, T] \to \{0, 1\}$. We shall let $q(t)$ be the proportion of individuals using $x = 1$ at time t, and denote $p(\cdot)$ a Wardrop equilibrium.

4.1.1 Several Resources

The collective behavior of the "players" has an influence on the "state of the world", say environmental quantities such as amount of resources that they may be depleting, habitat quality, etc. This is modelized as a vector dynamical system in $\mathbb{R}^m$:

$$\dot{y} = f(y, q), \qquad y(0) = y_0, \tag{8}$$

where f is assumed of class C^1, and such that this differential equation has a solution over $[0, T]$ for all $q(\cdot)$ in the set Δ of measurable functions from $[0, T]$ to $[0, 1]$.

The increase of the time rate of fitness acquisition by using $x = 1$ over $x = 0$ is a scalar function $g(y)$, also of class C^1. As a result, the cumulative excess fitness at the end of the season is $G(x(\cdot), q(\cdot)) = \int_0^T x(t)g(y(t))\,\mathrm{d}t$. Let

$$A(y, q) := \mathrm{D}_1 f(y, q) \in \mathbb{R}^{n\times n}, \quad b(y, q) := \mathrm{D}_2 f(y, q) \in \mathbb{R}^n$$
$$c(y) := \mathrm{D}g(y) \in (\mathbb{R}^n)^\star.$$

Having a resource depletion problem in mind, we assume that

- If all individuals use the available resources, $g(y)$ decreases, and if none does, the environment regenerates, and $g(y)$ increases:

$$c(y)f(y, 1) < 0, \qquad c(y)f(y, 0) > 0, \tag{9}$$

- for all y (at least in a neighborhood of $\{y \mid g(y) = 0\}$),

$$c(y)b(y, q) < 0. \tag{10}$$

As long as $g(y(t)) > 0$, the clear optimum for every individuals is to use $x = 1$, so that a Wardrop equilibrium, and a fortiori an E.S.S., must have $q(t) = 1$. Conversely, as long as $g(y(t)) < 0$, we end up with $q(t) = 0$. The only possible mixed Wardrop equilibrium is therefore with $g(y) = 0$, hence at pairs $(y(\cdot), q(\cdot))$ satisfying

$$c(y)f(y, q) = 0. \tag{11}$$

Thanks to hypotheses (9) and (10), equation (11) defines a unique smooth implicit function $q = \varphi_0(y) \in (0, 1)$. Altogether, this defines a closed loop Wardrop equilibrium $q(t) = \varphi(y(t))$ with

$$\varphi(y) = \begin{cases} 1 & \text{if } g(y) > 0, \\ \varphi_0(y) & \text{if } g(y) = 0, \\ 0 & \text{if } g(y) < 0. \end{cases}$$

This will cause the y trajectory to reach $\{g(y) = 0\}$ at a time t_0 with $p = 1$ if $g(y_0) > 0$, and with $p = 0$ if $g(y_0) < 0$, and then to follow a path in the manifold $g(y) = 0$. We denote by $p(t)$ the strategy history thus defined, and by $z(\cdot)$ the trajectory generated. The ensuing collective fitness is $F(p, p) = \int_0^{t_0}[g(z(t))]^+\,\mathrm{d}t$.

We now investigate the second-order condition to check whether that Wardrop equilibrium is also an E.S.S. Since using a mixed strategy while $g(y) \neq 0$

is clearly nonoptimal, we shall only consider the time interval $[t_0, T]$ during which $g(z(t)) = 0$. To apply Theorem 2, introduce the transition matrix $\Phi(t,s)$ of the matrix function $t \mapsto A(z(t), p(t))$, and the function $h(t,s) = c(z(t))\Phi(t,s)b(z(s), p(s))$. A simple calculation shows that the second-order E.S.S. condition reads (for measurable $q(\cdot)$'s):

$$\forall q(\cdot) \in \Delta, \quad \iint_{0\le s\le t\le T} (q(t) - p(t))h(t,s)(q(s) - p(s))\, \mathrm{d}s\, \mathrm{d}t < 0. \tag{12}$$

Notice that thanks to hypothesis (10), $h(t,t) = c(z(t))b(z(t), p(t)) < 0$ for all $t \in [0, T]$. Keeping this in mind, we have the following result.

Theorem 6. *A necessary condition for the Wardrop equilibrium described above to be an E.S.S. is that,*

$$\forall s < t \in [t_0, T] : p(s) \text{ or } p(t) \in (0,1), \quad h^2(t,s) - h(t,t)h(s,s) \le 0. \tag{13}$$

Proof. Select $s < t$ with $p(s)$ or $p(t)$ in $(0,1)$ and an arbitrary real number ξ. It is possible to find a small positive time interval $\eta \in (0, (t-s)/2)$ and a number, $\alpha \neq 0$ such that $q(\cdot)$ below is admissible:

$$q(\tau) = \begin{cases} p(\tau) + \alpha & \text{for } \tau \in (s, s+\eta), \\ p(\tau) + \alpha\xi & \text{for } \tau \in (t-\eta, t), \\ p(\tau) & \text{otherwise.} \end{cases}$$

Placing this q in (12), we see that as η goes to zero, $2\langle p - q, H_2(p-q)\rangle \to \alpha^2\eta[\xi^2 h(t,t) - 2\xi h(t,s) + h(s,s)]$, which must be nonpositive for all $\xi \in \mathbb{R}$.

4.1.2 Single Resource

In the scalar case, $y \in \mathbb{R}$, the condition $g(z) = 0$ generically specifies z completely (if, e.g. g is assumed to be strictly increasing), and $c(z)f(z,p) = 0$ specifies p. Then, A, b and c are just scalar constants. As a consequence, $(y(t), q(t)) = (z, p)$ is a solution of (8) and $\mathrm{d}q/\mathrm{d}t = 0$. Then, it is natural to investigate the equivalent of a "shortsighted replicator equation":

$$\varepsilon\dot{q} = q(1-q)g(y) \tag{14}$$

for some $\varepsilon > 0$, – since the time scale of this adaptive equation is a priori not related to the time of our model.

Proposition 3. *In the scalar case, under condition (10),*

- *The condition (13) is equivalent to $A \le 0$, which is equivalent to the Lyapunov stability of the variational system associated with (8)(14) at (z, p).*

- *Strengthening condition (13) to strict inequalities is equivalent to* $A < 0$*, which in turn insures local stability of* (z, p) *in (8)(14).*

Proof. Here, A is now a scalar constant. Hence $\Phi(t,s) = \exp(A(t-s))$, and the condition (13) reads $b^2c^2\,(\exp[2A(t-s)]-1) \le 0$, which is equivalent to $A \le 0$.

The variational system associated with (8)(14) is just

$$\begin{pmatrix} \dot{y} \\ \dot{q} \end{pmatrix} = \begin{pmatrix} A & b \\ c & 0 \end{pmatrix} \begin{pmatrix} y \\ q \end{pmatrix},$$

and the characteristic polynomial of that matrix is $\lambda^2 - A\lambda - bc$. (And $bc < 0$.)

4.2 The Tragedy of the Commons

The following is a dynamic form of the famous "Tragedy of the Commons" game, of which we provide two variants. Here, y is scalar and measures the amount of the common resource available, say the pasture shared by shepherds. We assume dynamics of the form

$$\dot{y} = a(y)y + bq + r$$

with $a(y) > 0$, a natural regeneration rate, $b < 0$, a constant depletion rate if all the local flocks graze in that same common pasture, and $r > 0$, a constant zero-resource regeneration rate[10]. The efficiency of sheep grazing is $g(y) = c(y - \gamma)$ for some fixed level γ, which is also the equilibrium z in the Wardrop equilibrium, and c is a positive constant.

We first assume malthusian dynamics, i.e. a a constant positive rate. According to the above proposition, the Wardrop equilibrium is not an ESS.

In a second variant, we replace the linear growth by a more realistic logistic growth, i.e., introduce a positive *carrying capacity* K, and in the above dynamics, set $a = \alpha(1 - y/K)$, with $b < -\alpha K/4$. Now, $A(x, p) = \alpha(1 - 2\gamma/K)$. If $\gamma \in (K/2, K)$, $A < 0$ and the conclusion is reversed: we have indeed an E.S.S.

If γ is thought of as a tax for using the common resource, this may provide a hint at how to choose it.

5 Conclusion

We have opened this leisurely walk through Wardrop equilibria, E.S.S. and replicator dynamics by Wardrop's invention in road engineering and ended it by an example in an environmentally inspired problem. This is a tribute to its origins in road traffic

[10] We need to restrict $y(0)$ to be less than $-(b+r)/a$ to insure hypothesis (9).

analysis, as well as to the fact that theoretical biology is the field that has brought more research into evolutionary games lately and made it a full-fledged theory in itself.

As far as we know, the question of convergence of the replicator dynamics toward an E.R.S. in the infinite case is not settled at this time. This is only one of the open questions left. But our feeling is that many interesting problems lie in particular applications, whether finite or infinite, possibly even with infinite dimensional strategy space as in the last example we provided.

6 Appendix: E.S.S. and E.R.S.

We recall the definition

Definition 4. A Wardrop equilibrium p is said *regular* if there exists a $\Delta(X_1)$-neighborhood of p such that over that neighborhood $q \mapsto H_1(x,q)$ is Lipschitz continuous for all $x \in X$ and the restriction of the quadratic form $\langle r, H_1(q)r\rangle$ to measures r orthogonal to $\mathbf{1}$ in $\mathscr{M}(X_1)$ is negative definite, both uniformly in q.

We prove here the theorem

Theorem 3.

(a) All E.R.S. are E.S.S.
(b) In the finite case, all regular Wardrop equilibria[11] are E.R.S.

Proof.

(a) Applying the E.R.S. condition to $q_\varepsilon = p + \varepsilon(q-p)$ and using the linearity of F in its first argument, we get condition (2).

(b) The proof of this point requires that there exist a compact set $K \subset \Delta(X)$ such that, for any positive $\bar\varepsilon$, the set $\{q_\varepsilon = p + \varepsilon(q-p),\ q \in K,\ \varepsilon \le \bar\varepsilon\}$ be a neighborhood of p. (In particular, $p \notin K$.) It is why it is restricted to the finite case, where $\Delta(X)$ is isomorphic to the finite dimensional simplex.
Let p be an E.S.S., $B(p) \in \Delta(X)$ the set of best responses q to p, i.e. such that $F(q,p) = F(p,p)$, $X_1 \subset X$ its support, G_1 the restriction of G to $X_1 \times \Delta(X_1)$. Let also q_ε and K be as just said above. We recall the definition of the negative score function $\mathscr{E}(\varepsilon,q) = F(q-p,q_\varepsilon)$ and that $B(p) = \{q \mid \mathscr{E}(0,q) = 0\}$, and that the derivative of $\mathscr{E}$ w.r.t. ε is $\mathrm{D}_1\mathscr{E}(\varepsilon,q) = \langle q-p, \mathrm{D}G(q_\varepsilon)(q-p)\rangle$.
Consider first the case where $q \in B(p)$. Then $\mathscr{E}(0,q) = 0$. Notice that $\mathrm{D}_1\mathscr{E}(\varepsilon,q) = \langle q-p, \mathrm{D}G_1(q_\varepsilon)(q-p)\rangle$ where p and q are in $\Delta(X_1) = \Sigma_{n_1}$, and the scalar product is accordingly that of $\mathbb{R}^{n_1}$. As a consequence of the hypothesis of regularity, this derivative is negative for all $q \in K_1 := K \cap B(p)$. Since K is compact, and therefore also K_1, $\mathrm{D}_1\mathscr{E}(0,q)$ is uniformly bounded

[11] and all E.S.S. in the linear case, but this is proved in [9, Theorem 6.4.1].

away from 0 on K_1. Because the derivative is locally Lipschitz continuous w.r.t. ε, it follows that there exists a $\hat{\varepsilon} > 0$ such that $\mathscr{E}(\varepsilon, q) < 0$ for all $\varepsilon \in (0, \hat{\varepsilon})$, and because the Lipschitz constant is assumed to be uniform, this same $\hat{\varepsilon}$ is valid for every $q \in K_1$. Let us summarize:

$$\exists \hat{\varepsilon} > 0 : \forall \varepsilon \in (0, \hat{\varepsilon}),\ \forall q \in K \cap B(p), \quad \mathscr{E}(\varepsilon, q) < 0. \tag{15}$$

This can be stated with the help of the following concept:

Definition 5. Define the *invasion barrier* $\varepsilon_0(q)$ as

$$\varepsilon_0(q) = \sup\{\varepsilon_1 \in [0,1] \mid \forall \varepsilon \in (0, \varepsilon_1), \mathscr{E}(\varepsilon, q) < 0\}.$$

Then (15) reads $\forall q \in K_1, \varepsilon_0(q) \geq \hat{\varepsilon} > 0$.

We claim the following:

Lemma 1. *The function* $\min\{\hat{\varepsilon}, \varepsilon_0(\cdot)\}$ *is lower semicontinuous over* K.

Proof. (of the lemma) Let $\delta \in (0, \varepsilon_0(q))$ be given. We want to prove that for every $q \in K$, there exists a neighborhood $\mathscr{N}(q)$ such that for $q' \in \mathscr{N}(q)$, $\varepsilon_0(q') \geq \min\{\hat{\varepsilon}, \varepsilon_0(q)\} - \delta$, hence that

$$\forall q' \in \mathscr{N}(q),\ \forall \varepsilon \in (0, \min\{\hat{\varepsilon}, \varepsilon_0(q)\} - \delta), \mathscr{E}(\varepsilon, q') < 0. \tag{16}$$

Assume first that q and q' are in $B(p)$. We are in the situation above, $\varepsilon_0(q') \geq \hat{\varepsilon}$, which proves the inequality.

If $q' \notin B(p)$, which is always possible if $q \notin K_1$, then $\mathscr{E}(0, q') < 0$, and by continuity and compacity, $\exists \varepsilon_1 > 0$ such that $\forall \varepsilon \in [0, \varepsilon_1]$, $\mathscr{E}(\varepsilon, q') < 0$. Also, $\exists \eta > 0$: $\forall \varepsilon \in [\varepsilon_1, \varepsilon_0(q) - \delta], \mathscr{E}(\varepsilon, q) < -\eta$. Then, by continuity of $\mathscr{E}$ w.r.t. q, uniform w.r.t. ε, for $\mathscr{N}(q)$ small enough and $q' \in \mathscr{N}(q)$, $\mathscr{E}(\varepsilon, q') < 0$. Hence, we get the desired result (16).

As a consequence of the lemma, $\varepsilon_0(q)$ has a positive minimum, say $\bar{\varepsilon}$ over K. It follows that $\forall \varepsilon < \bar{\varepsilon},\ \forall q \in K,\ F(q, q_\varepsilon) < F(p, q_\varepsilon)$, and hence by linearity of F w.r.t. its first argument

$$\forall \varepsilon \in (0, \bar{\varepsilon}),\ \forall q \in K, \quad F(q_\varepsilon, q_\varepsilon) - F(p, q_\varepsilon) = \varepsilon[F(q, q_\varepsilon) - F(p, q_\varepsilon)] < 0.$$

By definition of K, this set of q_ε's is a neighborhood of p.

This proves the theorem.

References

1. Altman, E., Bernhard, P., Debbah, M., Silva, A.: Continuum Equilibria for Routing in Dense Ad-hoc Networks. Computer Networks, to appear (2009)

2. Altman, E., Bernhard, P., Silva, A.: The mathematics of routing in massively dense ad-hoc networks. In: International Conference on Ad-Hoc Networks and Wireless: “AdHoc-NOW 2008”. Sophia Antipolis (2008)
3. Apaloo, J.: Revisiting matrix games: The concept of neighborhood invader strategies. Theoretical Population Biology **69**, 235–242 (2006)
4. Beckman, M.: A continuous model of transportation. Econometrica **20**, 643–660 (1952)
5. Bernhard, P., Hamelin, F.: Two-by-two static, evolutionary, and dynamic games. In: Y. Bertot, G. Huet, J.J. Lévy, G. Plotkin (eds.) From Semantics to Computer Science: Essays in Honor of Gilles Kahn, pp. 452–474. Cambridge University Press, Cambridge, U.K. (2008)
6. Bishop, D.T., Cannings, C.: A generalized war of attrition. Journal of Theoretical Biology **70**, 85–125 (1978)
7. Cressman, R.: Evolutionary Dynamics and Games in Extensive form. MIT, Cambridge, MA (2003)
8. Day, T., Taylor, P.D.: A generalization of pontryagin’s maximum principle for dynamic evolutionary games among relatives. Theoretical Population Biology **57**, 339–356 (2000)
9. Hofbauer, J., Sigmund, K.: Evolutionary Games and Population Dynamics. Cambridge University Press, Cambridge, U.K. (1998)
10. Maynard Smith, J.: The theory of games and the evolution of animal conflicts. Journal of Theoretical Biology **47**, 209–221 (1974)
11. Maynard Smith, J.: Evolution and the Theory of Games. Cambridge University Press, Cambridge, U.K. (1982)
12. Maynard Smith, J., Price, G.R.: The logic of animal conflict. Nature **246**, 15–18 (1973)
13. von Neumann, J.: Zur theorie der gesellschaftspiele. Mathematische Annalen **100**, 295–320 (1928)
14. von Neumann, J., Morgenstern, O.: The Theory of Games and Economic Behavior. Princeton University Press, Princeton, NJ (1944)
15. Romagny, B., Lobry, C., Canalis-Durand, M.: Tragedy of the commons and prisoner’s dilemma. Tech. Rep. 97A20, GREQAM (1997)
16. Samuelson, L.: Evolutionary Games and Equilibrium Selection. MIT, Cambridge, MA (1997)
17. Sandholm, W.: Population Games. MIT, Cambridge, MA, to appear (2009)
18. Shaiju, A., Bernhard, P.: Evolutionarily stable strategies: Two nontrivial examples and a theorem. In: Bernhard, P., Gaitsgory, V., and Pourtallier, O. ed.: Advances in Dynamic Games and their Applications (also in 12th International Symposium on Dynamic Games and Applications, Sophia Antipolis, France, 2006), 377–398, Birkhauser (2009)
19. Taylor, P., Jonker, L.: Evolutionarily stable strategies and game dynamics. Mathematical Bioscience **40**, 145–156 (1978)
20. Vincent, T.L., Brown, J.S.: Evolutionary Game Theory, Natural Selection and Darwinian Dynamics. Cambridge University Press, Cambridge, U.K. (2006)
21. Wardrop, J.G.: Some theoretical aspects of road traffic research. Proceedings of the Institution of Civil Engineers, pp. 325–378 (1952)
22. Weibull, J.: Evolutionary Game Theory. MIT, Cambridge (1995)

A Markov Decision Evolutionary Game for Individual Energy Management

Yezekael Hayel, Hamidou Tembine, Eitan Altman, and Rachid El-Azouzi

Abstract We study in this paper a noncooperative game with an infinite number of players that are involved in many local interactions, each involving a randomly selected pair of players. Each player has to choose between an aggressive or a nonaggressive action. The expected lifetime of an individual as well as its expected total fitness during its lifetime (given as the total amount of packets it transmits during the lifetime) depend on the level of aggressiveness (power level) of all actions it takes during its life. The instantaneous reward of each player depends on the level of aggressiveness of his action as well as on that of his opponent. We model this as a Markov Decision Evolutionary Game which is an extension of the evolutionary game paradigm introduced in 1972 by Maynard Smith, and study the structure of equilibrium policies.

1 Introduction

The evolutionary games formalism is a central mathematical tool developed by biologists for predicting population dynamics in the context of *interactions between populations*. This formalism identifies and studies two concepts: The *Evolutionary Stability*, and the *Evolutionary Game Dynamics* [8].

One of the evolutionary stability concepts is the evolutionarily stable strategy (ESS). The ESS is characterized by a property of robustness against invaders (mutations). More specifically,

- If an ESS is reached, then the proportions of each population do not change in time.

Y. Hayel (✉), H. Tembine, and R. El-Azouzi
Laboratoire Informatique d'Avignon/CERI, Université d'Avignon, 84911, Avignon, France
e-mail: yezekael.hayel@univ-avignon.fr; tembineh@yahoo.fr; rachid.elazouzi@univ-avignon.fr

E. Altman
INRIA Sophia Antipolis Méditerranée, 2004, route des lucioles, 06902, Sophia Antipolis, France
e-mail: eitan.altman@sophia.inria.fr

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_16,

- At ESS, the populations are immune from being invaded by other small populations. This notion is stronger than Nash equilibrium in which it is only requested that a single player would not benefit by a change (mutation) of its behavior.

ESS has first been defined in 1972 by J. Maynard Smith [10], who further developed it in his seminal text *Evolution and the Theory of Games* [11], followed shortly by Axelrod's famous work [4]. Another element of evolutionary games is the *evolutionary game dynamics*. Some particular dynamics has been used for describing the evolution of road traffic congestion in which the fitness is determined by the strategies chosen by all drivers [14]. Others have also been studied in the context of the association problem in hybrid wireless communications [17].

Although evolutionary games were defined in the context of biological systems, they may be highly relevant to engineering as well (see [19]). In the biological context, evolutionary dynamics is mainly used to model and explain change of the size of the population(s) as a function of the chosen actions, whereas in engineering, we can go beyond characterizing and modeling existing evolution. In particular, in networking we may set guidelines or rules for upgrading systems, services, protocols, etc. so as to avoid undesirable behavior (such as oscillations or chaotic behavior), see [3]. In order to apply evolutionary game ideas to communication networks, we had to extend several basic concepts in evolutionary game theory. We had to extend the interactions between players to allow for an arbitrary number (possibly random) of players to be involved in each local interaction. We also had to consider non reciprocal interactions [18]. A further important extension is to consider each local interaction as a one step in a stochastic game rather than a matrix game. More precisely, players may have (random) states, and their action when interacting with other players determine not only their immediate payoff but also the probability to move to the other states. We call these "Markov Decision Evolutionary Games" [2]. Instead of maximizing a fitness related to a one-shot game, an individual maximizes the total expected fitness accumulated during its lifetime.

In this paper, we study such a Markov Decision Evolutionary game in a control problem faced by each individual within a large population of mobile terminals. Not only is the evolutionary game paradigm borrowed from biology, but also the nature of the control problem we consider: that of energy management. Indeed, the long-term animal survival is directly related to its energy strategies (competition over food, etc), and a population of animals that have good strategies for avoiding starvation is more fit and is expected to survive [9, 12]. By analogy, we may expect mobile terminals that adopt efficient energy strategies to see their population increase. Unlike the biological model [12] where the state evolution of each individual depends also on the action of the other one involved in a local interaction, the energy level (state) of a mobile depends only on its own actions (transmission power level). The contribution of this work can be summarized in different points:

- The Markov decision evolutionary game presented in this paper has a more complex structure than a standard evolutionary game [11]. In particular, the fitness criterion that is maximized by every user is equal to the sum of fitness obtained during all the opportunities in the player's lifetime. Restricting the game to stationary mixed policies, we obtain the characterization of Evolutionary stable strategies (ESS) under state-independent strategies assumption. We also provide some condition about the existence and uniqueness of ESS.

- For more general state-dependent strategies, we show that the nonaggressive action, which is weakly dominated in the one-shot game, is not necessarily dominated in the long-term dynamic game. Moreover, the ESS is still more aggressive than the global optimum and best response policies are threshold-based policies.
- Finally, we propose a replicator dynamics that extends the standard one and study its convergence to ESS.

The structure of the paper is as follows. The next section gives an overview of basic notion of evolutionary game theory. The model of interactions between players in a single population is presented in Sect. 3. The computation of ESS is proposed in Sect. 4. For simplicity, in Sect. 4.1 the model is reduced to state-independent mixed strategies. Then in Sect. 4.2 we consider the more general model with state-dependent strategies. Finally, Sect. 5 provides numerical illustrations and Sect. 6 concludes the paper. All proofs are in the technical report [6].

2 Background on Evolutionary Games

Consider a large population of players with finite number of choices K. We define by $J(p,q)$ the expected immediate payoff for a player if it uses a strategy (also called policy) p when meeting another player who adopts the strategy q. This payoff is called "fitness" and strategies with larger fitness are expected to propagate faster in a population. p and q belong to a set K of available strategies.

In the standard framework for evolutionary games, there are a finite number of so-called "pure strategies", and a mixed strategy of a player is a probability distribution over the pure strategies. An equivalent interpretation of strategies is obtained by assuming that players choose pure strategies and then the probability distribution represents the fraction of players in the population that choose each strategy.

Suppose that the whole population uses a strategy q and that a small fraction ϵ (called "mutations") adopts another strategy p. Evolutionary forces are expected to select against p if

$$J(q, \epsilon p + (1-\epsilon)q) > J(p, \epsilon p + (1-\epsilon)q) \quad (1)$$

Definition 1. A strategy $q \in \Delta(K)$ is said to be ESS if for every $p \neq q$ there exists some $\bar{\epsilon}_p > 0$ such that (1) holds for all $\epsilon \in (0, \bar{\epsilon}_p)$.

In fact, we expect that if

$$\text{for all } p \neq q, \quad J(q,q) > J(p,q), \quad (2)$$

then the mutations fraction in the population will tend to decrease (as it has a lower reward, meaning a lower growth rate). q is then immune to mutations. If it does not but if still the following holds

$$\text{for all } p \neq q, \quad J(q,q) = J(p,q) \text{ and } J(q,p) > J(p,p), \tag{3}$$

then a population using q is "weakly" immune against a mutation using p since if the mutant's population grows, then we shall frequently have players with strategy q competing with mutants; in such cases, the condition $J(q,p) > J(p,p)$ ensures that the growth rate of the original population exceeds that of the mutants. We shall need the following characterization:

Theorem 1. see [20, Proposition 2.1] or [7, Theorem 6.4.1, page 63] *A strategy q is ESS if and only if it satisfies (2) or (3).*

Corollary 1. *Equation (2) is a sufficient condition for q to be an ESS. A necessary condition for it to be an ESS is for all $p \neq q, \quad J(q,q) \geq J(p,q)$.*

The conditions on ESS can be related to and interpreted in terms of Nash equilibrium in a matrix game. The situation in which a player, say player 1, is faced with a member of a population in which a fraction p chooses strategy A is then translated to playing the matrix game against a second player who uses mixed strategies (randomizes) with probabilities p and $1-p$, resp.

The following theorem gives another necessary and sufficient condition for a strategy to be an ESS for evolutionary games with bilinear payoff functions. A proof can be found in [5].

Theorem 2. *Let q be a symmetric (Nash) equilibrium for the matrix game with payoffs (J, J^t), where J^t is the transposition matrix of $J = (J(i,j)_{i,j})$ and $BRP(q)$ be the pure best response to q i.e $BRP(q) = \{j \in K \mid \sum_k J(j,k)q_k = \max_{l \in K} \sum_{k' \in K} J(l,k')q_{k'}\}$. Define $\bar{q}$ as*

$$\bar{q}_j = \begin{cases} q_j & \text{if } j \in BRP(q) \\ 0 & \text{otherwise} \end{cases}.$$

Let $\tilde{J}$ the submatrix obtained from J by taking only the index $i,j \in BRP(q)$. Then q is an evolutionarily stable strategy if and only if $\sum_{k \in BRP(q)}(p_k - \bar{q}_k) \sum_{j \in BRP(q)} \tilde{J}_{k,j}(p_j - \bar{q}_j) < 0$ for all $p \neq q$.

Note that the last condition is equivalent to $\forall y \in Y, \sum_{k,j \in BRP(q)} y_k y_j \tilde{J}_{k,j} < 0$, where $Y := \left\{z \in \mathbb{R}^{|BRP(q)|} \setminus \{0\}, \sum_{j \in BRP(q)} z_j = 0, \text{ and } \bar{q}_j = 0 \Longrightarrow z_j \geq 0\right\}$ and $|BRP(q)|$ is the cardinal of the finite set $BRP(q)$.

3 Model of Individual Behavior and Resource Competition

We consider in this paper a theoretical generalization of the model of energy management proposed by Altman and Hayel in [1] for a distributed aloha network. The authors consider noncooperative interactions among a large population of mobile terminals. A terminal attempts transmissions during a finite horizon of times which depends on the state of its battery energy. At each time, each mobile has to take a decision on the transmission power, based on his/her own battery state. A transmission is successful if one of the two following conditions holds:

- No other user transmits during the time slot.
- The mobile transmits with a power which is bigger than the power of all other transmitting mobiles at the same base station.

In this paper, we assume that each mobile terminal is an individual or a player of a global population and the interaction for transmission is interpreted as a fight for food (as in [12]). It is important to note that the Markov process of each player is controlled by himself, whereas in [12] the interaction determines the evolution of the individual process. Each player starts his/her life in the highest energy state or level, say state n. Starting from state n, each player will visit the different levels in decreasing order until reaching the last state 0. The state 0 corresponds to the state *Empty* and the state n is the *Full* state of the battery. The other states $1, \ldots, n-1$ are intermediary states of the battery or energy. When the system is at the state 0 there is no energy and the player has to charge his battery or to buy a new battery. We will call *terminal absorbing fitness* the expected cost of charge or price of a new battery. At each time, there should be an interaction with another player and then an action should be taken by each player in the population. The actions taken by a player determine not only the immediate reward but also the transition probabilities to its next individual state (energy level). The instantaneous reward depends also on the level of aggressiveness of the other players. Then, the evolution of the level of energy depending on the player's action can be modeled as a discrete time, homogeneous Markov process as in [1, 16].

3.1 Individual Sequential Decision

Each player has to choose between two possible actions in each state (different from the *Empty* state): high(h) and low(l). The action h is considered as an aggressive action and the action l as passive. Since there are only two strategies h and l, the population aggressiveness can be described by a process $(\alpha_t)_{t\geq 1}$, where α_t is the fraction of the population using the aggressive action, that is the action h, at time t. The stochastic process α_t may depend on the history with length t. The main issue of this paper is to study how a player will manage its remaining energy (corresponding to its actual state) during his lifetime to optimize its total reward.

A pure strategy of a player at time t is a map $\alpha_t : H_t \longrightarrow \{h, l\}$, where $H_t = All^{t-1} \times S$ is the set of histories at time t with $All = \{(s, a) \mid s \in \{1, \ldots, n\}\ a \in \{h, l\}\}$. The collection $\sigma = (\sigma_t)_{t \geq 1}$ of pure strategies at each time constitutes a pure strategy of the population. We denote by Σ the set of all pure strategies.

Let σ be a strategy profile, and $z_t = (s_1, a_1, \ldots, s_{t-1}, a_{t-1}, s_t)$ be a history of length t. The continuation strategy $\overrightarrow{\sigma}_{z_t}$ given the history z_t is defined as follows

$$\overrightarrow{\sigma}_{z_t}(s'_1, a'_1, \ldots, a'_{t'-1}, s'_{t'}) = \sigma(s_1, a_1, \ldots, s_{t-1}, a_{t-1}, s_t, s'_1, a'_1, \ldots, a'_{t'-1}, s'_{t'}).$$

The fitness $v = r(\overrightarrow{\sigma}_{z_t})$, obtained after considering the strategy $\overrightarrow{\sigma}$, is called *continuation fitness* after z_t. If σ is a stationary strategy, then v depends only on the state s_t.

For each individual state $s \neq 0$, the action space of each player becomes the interval $[0, 1]$. A player *policy* is defined as the set of actions taken during his/her life in each time slot. We denote a policy u as the sequence $(u_1, u_2, \ldots)$, where u_t is the probability of choosing action h at time slot t.

For the remainder we shall consider only stationary policies where the probabilities u_t is the same for all time t. We denote by $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$ the player policy, where β_s is the probability of choosing action h in state s. The pure aggressive (resp. nonaggressive) policy is denoted by $\overline{h} = (1, \ldots, 1)$ (resp. $\overline{l} = (0, \ldots, 0)$).

3.2 Individual State

The individual state at time t is a Markovian random process $(X_t)_{t \geq 1}$, where transitions depend on individual action. Then, the energy management problem is a Markov decision process [13] dependent on the population profile.

The Markov decision process (X_t) has the following transition law

$$\forall s \neq 0, \forall a \in \{h, l\},\ q(X_{t+1} = s' | X_t = s, a) = \begin{cases} Q_s(a) & \text{if } s' = s \\ 1 - Q_s(a) & \text{if } s' = s - 1 \\ 0 & \text{otherwise} \end{cases}$$

and $q(0|0, a) = 1$; where $Q_s(a)$ is the remaining probability in state s using action a.

We assume two main characteristics of the transition probabilities.

- First, $\forall s \neq 0,\ Q_s(h) < Q_s(l) < 1$ reflects that using action h, the probability to remain in any given energy level is lower than using action l.
- Second, $\forall a$, the discrete function $s \longmapsto Q_s(a)$ is nondecreasing. This assumption means that the lower is the energy level, the lower is the probability to remain in the same level for each action.

Note that each player controls the transition state of its own battery, i.e. probabilities Q are independent of the other player's decisions.

3.3 Binary Reward

We define a binary reward of a player in an interaction as follows: The player wins the resource if (a) he/she has no opponents for the same resource during the time slot, or (b) he/she uses the most aggressive action and its opponents are passive. Otherwise, he/she gets zero. Note that the resource is not shared when interaction occurs between two players with the same action, as it is the case in the Hawk and Dove game studied in [12]. We extend the reward function to mixed strategies by randomizing the binary reward.

We denote by $1-p$ the probability for any player of having an opponent at the same time slot. The probability of success (probability to win to the resource) of a player using the action h when its energy level is at the state $s>0$ is given by $r(s,h,\alpha) = p + (1-p)[\alpha \times 0 + (1-\alpha) \times 1] = p + (1-p)(1-\alpha)$, where α is the fraction of the population who uses the action h at any given time. Note that α is stationary does not mean that the actions of each player are fixed in time. It only reflects a situation in which the system attains a stationary regime due to the averaging effect over a very large population, and the fact that all players choose an action in a given individual state using the same probability law.

Similarly using action l, the probability of success in each state $s \neq 0$ is given by: $r(s,l,\alpha) = p$.

The expected reward of a player in state $s > 0$ is then given by

$$r(s,\beta_s,\alpha) = \beta_s r(s,h,\alpha) + (1-\beta_s) r(s,l,\alpha) = p + \beta_s (1-p)(1-\alpha),$$

where β_s is the probability of choosing action h in state s. Considering this reward function, a player has more chance to win the resource with an aggressive action than a passive one, but a passive action saves energy for future.

Let $T(\beta)$ the expected time for an individual to reach the state 0 under the stationary policy β. The total reward $V_\beta(n,\alpha)$ of a player depends on his/her birth state which is state n for all users, his/her stationary policy β, and the global population profile α. This function corresponds to the expected fitness of an individual and it is given by

$$\begin{aligned} V_\beta(n,\alpha) &= \mathbb{E}^{\beta,\alpha} \sum_{t=1}^{T(\beta)} r(s_t, \beta_{s_t}, \alpha_t) \\ &= r(n,\beta_n,\alpha) + \mathbb{E}^{\beta,\alpha} \left(\sum_{t=2}^{T(\beta)} r(s_t, \beta_{s_t}, \alpha) \,|s_{t_1} = n \right). \end{aligned}$$

Remark 1. Consider a pairwise one-shot interaction between a player i which is in state $s \neq 0$ and a player j which is in state $s' \neq 0$. Define the action set of each player as $\{h,l\}$ and the payoff function as the binary reward defined above. It is not difficult to see that the instantaneous reward $r(s,\beta_s,\alpha)$ increases with β_s. Then, the one-shot game is a degenerate game with an infinite number of Nash equilibria. We will see later that the nonaggressive action which is weakly dominated in the one-shot game is not necessarily dominated in the long-term dynamic game that we will describe later.

3.4 Computing Fitness using Dynamic Programming

We define the dynamic programming operator $Y(s,a,\alpha,v)$ to be the total expected fitness of a player starting at state s if

- It takes action a at time 1.
- At time 2, the total sum of expected fitness from time 2 onward is v.
- At each time slot, the probability that another player uses action h, given that there is an interaction between individual players, is α.

We denote by v the continuation fitness function which depends on the state s and the global population profile α. We have $Y(s,a,\alpha,v) = r(s,a,\alpha) + Q_s(a)v(s,\alpha) + (1 - Q_s(a))v(s-1,\alpha), \forall a \in \{h,l\}$.

The function v is a fixed point of the operator ϕ defined by $v \longrightarrow (\phi v) := \beta_s Y(s,h,\alpha,v) + (1-\beta_s)Y(s,l,\alpha,v)$.

Proposition 1. *The expected fitness or total reward for a player during its lifetime is given by:*

$$V_\beta(n,\alpha) = \sum_{i=1}^{n} \frac{r(i,\beta_i,\alpha)}{1-Q_i(\beta_i)} - C.$$

Proof. We compute the fitness by using the recursive formula defined by the operator ϕ. First, we have that $v(0,\alpha) = 0$ and then

$$\begin{aligned} v(1,\alpha) = \phi(v) = \;& \beta_1[r(1,h,\alpha) + Q_1(h)v(1,\alpha) + 0] \\ & +(1-\beta_1)[r(s,l,\alpha) + Q_s(l)v(1,\alpha) + 0], \\ & \quad = r(1,\beta_1,\alpha) + [\beta_1 Q_1(h) + (1-\beta_1)Q_1(l)]v(1,\alpha). \end{aligned}$$

The fixed point $v(1,\alpha)$ is given by

$$V_\beta(1,\alpha) = \frac{r(1,\beta_1,\alpha)}{1-[\beta_1 Q_1(h) + (1-\beta_1)Q_1(l)]} = \frac{p + \beta_1(1-p)(1-\alpha)}{1-[\beta_1 Q_1(h) + (1-\beta_1)Q_1(l)]}.$$

For every state $s > 1$, we have

$$\begin{aligned} v(s,\alpha) &= \beta_s[r(s,h,\alpha) + Q_s(h)v(s,\alpha) + (1-Q_s(h))v(s-1,\alpha)] \\ &\quad +(1-\beta_s)[r(s,l,\alpha) + Q_s(l)v(s,\alpha) + (1-Q_s(l))v(s-1,\alpha)], \\ &= r(s,\beta_s,\alpha) + [\beta_s Q_s(h) + (1-\beta_s)Q_s(l)]v(s,\alpha) \\ &\quad +[\beta_s(1-Q_s(h)) + (1-\beta_s)(1-Q_s(l))]v(s-1,\alpha), \\ &= r(s,\beta_s,\alpha) + Q_s(\beta_s)v(s,\alpha) + (1-Q_s(\beta_s))v(s-1,\alpha), \end{aligned}$$

where $Q_s(\beta_s) := \beta_s Q_s(h) + (1-\beta_s)Q_s(l)$. Since $Q_s(\beta_s) < 1$ one has $v(s,\alpha) = \frac{r(s,\beta_s,\alpha)}{1-Q_s(\beta_s)} + v(s-1,\alpha)$. By rewriting $v(s,\alpha) = \sum_{i=1}^{s}[v(i,\alpha) - v(i-1,\alpha)]$, we obtain that

$$v(s,\alpha) = \sum_{i=1}^{s} \frac{r(i,\beta_i,\alpha)}{1-Q_i(\beta_i)}.$$

Hence, the expected fitness obtained by starting from the *Full* state n to the next new battery is then given by $V_\beta(n,\alpha) - C$. □

We denote by $T_s(\beta_s)$ the sojourn time of a player at the state s with the policy β_s. The sojourn time satisfies the following fixed point formula:

$$T_s(\beta_s) = 1 + \sum_{a\in\{h,l\}} Q_s(a)\beta_s(a)T_s(\beta_s).$$

Hence, $T_s(\beta_s) = \frac{1}{1-Q_s(\beta_s)}$ and the fraction of times which any player chooses the aggressive action h in state s is given by $\hat{\alpha}_s(\beta) = \beta_s \frac{T_s(\beta_s)}{\sum_{i=1}^{n} T_i(\beta_i)}$. Thus, the probability for any player to meet another player which is using action h is given by

$$\hat{\alpha}(\beta) := \sum_{s=1}^{n} \hat{\alpha}_s(\beta) = \frac{\sum_{s=1}^{n} \beta_s T_s(\beta_s)}{\sum_{i=1}^{n} T_i(\beta_i)} \leq 1. \tag{4}$$

The function $\widehat{\alpha}$ gives a mapping between the action vector β of an anonymous player and the global population profile α.

4 Evolutionary Stable State

As already mentioned, our game is different and has a more complex structure than a standard evolutionary game. In particular, the fitness that is maximized is not the outcome of a single interaction but of the sum of fitnesses obtained during all the opportunities in the player's lifetime. In spite of this difference, we shall still use the definition (1) or the equivalent conditions (2) or (3) for the ESS but with the following changes:

- β replaces the action q in the initial definition.
- β' replaces the action p in the initial definition.
- We use for $J(q,p)$ the total expected fitness $V_\beta(n,\alpha)$, where $\alpha = \widehat{\alpha}(\beta')$ is given in (4).

Thus, we obtain the following characterization of Evolutionarily Stable Strategies (ESS).

Corollary 2. *A necessary condition for β^* to be an ESS is*

$$\textit{for all } \beta' \neq \beta^*, \quad V_{\beta^*}(n,\widehat{\alpha}(\beta^*)) \geq V_{\beta'}(n,\widehat{\alpha}(\beta^*)) \tag{5}$$

A sufficient condition for β^* to be an ESS is that (5) holds with strict inequality for all $\beta' \neq \beta^*$.

Using this corollary and this simple necessary condition to be an ESS, we are now able to determine the existence of an ESS.

First of all, we need the following lemma in numerous proofs.

Lemma 1. *Given $\hat{\alpha}(\beta')$, the function $Z : \beta_s \longmapsto \frac{p+(1-p)(1-\hat{\alpha}(\beta'))\beta_s}{1-Q_s(\beta_s)}$ is monotone.*

Proof. Since Z is a continuously differentiable function, Z is strictly decreasing if

$$(1-p)(1-\hat{\alpha}(\beta'))(1-Q_s(l)) < (Q_s(l) - Q_s(h))p,$$

constant if $(1-p)(1-\hat{\alpha}(\beta'))(1-Q_s(l)) = (Q_s(l)-Q_s(h))p$ and strictly increasing if $(1-p)(1-\hat{\alpha}(\beta'))(1-Q_s(l)) > (Q_s(l)-Q_s(h))p$. Thus, given p, Q, the function Z is monotone. □

4.1 State-Independent Strategies

We assume in this section, for simplicity, that the probability of choosing action h does not depend on the state, that is $\forall s \ \beta_s = \rho$. We first consider only $\bar{h}$ the fully aggressive policy (the player uses always the aggressive action h in every state) and $\bar{l}$ the fully passive policy (the player uses always the nonaggressive action) policies.

Proposition 2. *If $p \neq 0$, then the strategy $\bar{h}$ which consists of playing always h ($\rho = 1$) cannot be an ESS. Otherwise, if $p = 0$ then $\bar{h}$ is an evolutionarily stable strategy.*

Proof. Suppose that $p \neq 0$. Then, for all stationary mixed policy β' we have

$$V_{\bar{l}}(n, \alpha(\beta')) = p \sum_s \frac{1}{1 - Q_s(l)}, \quad \text{and} \quad V_{\bar{h}}(n, \alpha(\bar{h})) = p \sum_s \frac{1}{1 - Q_s(h)}.$$

Thus taking as β' the fully aggressive policy $\bar{h}$ we obtain $V_{\bar{h}}(n, \alpha(\bar{h})) < V_{\bar{l}}(n, \alpha(\bar{h}))$. Then, it is better to use the action l if the opponent uses action h (the payoff is p and the lifetime is increased). Hence, the strategy $\bar{h}$ is not a best reply to itself. This means that $\bar{h}$ is not a Nash equilibrium and then $\bar{h}$ is not an ESS.

Suppose now that $p = 0$. Then the policy $\bar{l}$ is weakly dominated, and $V_{\bar{h}}(n, \alpha(\bar{h})) = 0 = V_{\bar{l}}(n, \alpha(\beta'))$, $\forall \beta'$. Moreover, for all $\beta' \neq \bar{h}$ we have $V_{\beta'}(n, \alpha(\bar{h})) = 0 = V_{\bar{h}}(n, \alpha(\bar{h}))$, and

$$\begin{aligned} V_{\bar{h}}(n, \alpha(\beta')) &= (1 - \alpha(\beta')) \sum_s \frac{1}{1 - Q_s(h)} \\ &> V_{\beta'}(n, \alpha(\beta')) = (1 - \alpha(\beta')) \sum_s \frac{\beta'_s}{1 - Q_s(\beta'_s)}. \end{aligned}$$

This induces policy $\bar{h}$ is an ESS and that completes the proof. □

This result is somehow logical because if, at each time slot, any player is always in interaction with another player ($p = 0$), then the only action which permits to win the fight is to be aggressive. We have the same kind of result for the nonaggressive action.

Proposition 3. *The strategy $\bar{l}$ which consists of playing always l ($\rho = 0$) is an ESS if and only if the subset of states $I_1 := \{s,\ p < \frac{1-Q_s(l)}{1-Q_s(h)}\}$ is empty.*

Proof. We have

$$V_0(n,\alpha(\beta')) = p\sum_{s=1}^{n}\frac{1}{1-Q_s(l)}, \quad \text{and} \quad V_\beta(n,\alpha(0)) = \sum_{s=1}^{n}\frac{p+(1-p)\beta_s}{1-Q_s(\beta)}.$$

Thus,

$$\begin{aligned}\max_\beta V_\beta(n,\alpha(0)) = \sum_{s\in I_1}\frac{1}{1-Q_s(h)} + p\sum_{s\in I_0}\frac{1}{1-Q_s(l)} \\ + \sum_{s\in I_*}\frac{1}{1-Q_s(h)} < V_0(n,\alpha(0)),\end{aligned} \tag{6}$$

where $I_1 := \{s,\ p < \frac{1-Q_s(l)}{1-Q_s(h)}\}$, $I_0 := \{s,\ p > \frac{1-Q_s(l)}{1-Q_s(h)}\}$, and $I_* := \{s,\ p = \frac{1-Q_s(l)}{1-Q_s(h)}\}$.

The subsets I_0, I_1, I_* constitute a partition of S. I_* has at most one element. Using Lemma 1 and (6), $\bar{l}$ is a best response to itself if $\sum_{s\in I_1}\frac{1}{1-Q_s(h)} < p\sum_{s\in I_1}\frac{1}{1-Q_s(l)}$. This inequality does not hold if I_1 is not empty. □

We consider now stationary mixed strategies where each player determines the same probability ρ of choosing the aggressive action in every state.

Proposition 4. *If the same level of aggressiveness $\beta_s = \rho \in (0,1)$, $\forall s$ is an ESS then it must satisfy*

$$\frac{p\sum_{s=1}^{n}\frac{(Q_s(l)-Q_s(h))}{(1-Q_s(\rho))^2}}{(1-p)\sum_{s=1}^{n}\frac{(1-Q_s(l))}{(1-Q_s(\rho))^2}} = 1-\rho. \tag{7}$$

If p, Q and $0 \le \rho \le 1$ satisfy (7), then ρ is an ESS.

Proof. The action ρ is a best reply to the action β' if it satisfies $V_\rho(n,\alpha(\beta')) \ge V_\mu(n,\alpha(\beta'))$, $\forall\mu \in [0,1]^n$. Then, ρ is a best response to itself if $V_\rho(n,\alpha(\rho)) \ge V_\mu(n,\alpha(\rho))$, $\forall\mu \in [0,1]^n$. The function $h(x) = \sum_{s=1}^{n}\frac{p+(1-\rho)(1-p)x}{1-Q_s(x)}$ has an interior extremum if

$$h'(x) = 0 \Longleftrightarrow 1 - \rho = \frac{p \sum_{s=1}^{n} \frac{(Q_s(l) - Q_s(h))}{(1-Q_s(x))^2}}{(1-p)\sum_s \frac{(1-Q_s(l))}{(1-Q_s(x))^2}}.$$

Note that $\alpha(\rho) = \rho$.

We now show that for every p, Q such that $\frac{p \sum_{s=1}^{n} \frac{(Q_s(l) - Q_s(h))}{(1-Q_s(\rho))^2}}{(1-p)\sum_s \frac{(1-Q_s(l))}{(1-Q_s(\rho))^2}} \leq 1$, every action $(\rho, 1-\rho)$ satisfying

$$1 - \rho = \frac{p \sum_{s=1}^{n} \frac{(Q_s(l) - Q_s(h))}{(1-Q_s(\rho))^2}}{(1-p)\sum_s \frac{(1-Q_s(l))}{(1-Q_s(\rho))^2}}$$

is an ESS. Let β' be a mixed action such that $V_\rho(n, \rho) = V_{\beta'}(n, \rho)$. We denote by $C' := V_\rho(n, \alpha(\beta')) - V_{\beta'}(n, \alpha(\beta'))$. We now prove that C' is strictly positive for all $\beta' \neq \rho$.

The condition $C' > 0$ is satisfied if $V_\rho(n, \alpha(\beta')) > V_{\beta'}(n, \alpha(\beta'))$ which is equivalent to

$$\sum_{s=1}^{n} \frac{p + (1-p)(1-\alpha(\beta'))\rho}{1 - Q_s(\rho)} > \sum_{s=1}^{n} \frac{p + (1-p)(1-\alpha(\beta'))\beta'}{1 - Q_s(\beta')}.$$

Since $V_\rho(n, \rho) = V_{\beta'}(n, \rho)$, one has

$$[p + (1-p)(1-\rho)\rho] \sum_s \frac{1}{1 - Q_s(\rho)} = [p + (1-p)(1-\rho)\beta'] \sum_s \frac{1}{1 - Q_s(\beta')}.$$

This implies that $\sum_s \frac{1}{1-Q_s(\beta')} = \frac{p+(1-p)(1-\rho)\rho}{p+(1-p)(1-\rho)\beta'} \sum_s \frac{1}{1-Q_s(\rho)}$.

By restituting this value in C', we obtain $C' > 0$ if

$$\frac{p + (1-p)(1-\alpha(\beta')}{p + (1-p)(1-\alpha(\beta'))\beta'} > \frac{p + (1-p)(1-\rho)\rho}{p + (1-p)(1-\rho)\beta'}.$$

By developing and simplifying the last expression, we obtain that $C' > 0$ is equivalent to $[\beta' - \rho][\alpha(\beta') - \rho] = [\beta' - \rho]^2 > 0$. This completes the proof. □

We are now able to determine explicitly, when it exists, an ESS of our Markov Decision Evolutionary Game. It is now interesting to provide conditions of existence of that ESS.

Proposition 5. *A sufficient condition of existence of an interior state-independent evolutionarily stable strategy is given by*

$$p < p_0 := \frac{\sum_{s=1}^{n} \frac{1}{1-Q_s(l)}}{\sum_{s=1}^{n} \frac{(1-Q_s(h))}{(1-Q_s(l))^2}}. \tag{8}$$

Proof. The function

$$\xi(p,\rho) = p\sum_{s=1}^{n}\frac{(Q_s(l)-Q_s(h))}{(1-Q_s(\rho))^2} - (1-p)(1-\rho)\sum_{s=1}^{n}\frac{(1-Q_s(l))}{(1-Q_s(\rho))^2},$$

is continuous on $(0,1)$ and $\xi(p,1) > 0,\ \forall p \in (0,1)$. Thus, if (8) is satisfied then $\xi(p,0) < 0$ and 0 is in the image of $(0,1)$ by $\xi(p,.)$. This implies that there exists $\rho \in (0,1)$ such that $\xi(p,\rho) = 0$. □

Proposition 6. *The game has at most one ESS in state-independent strategies.*

Proof. (a) Uniqueness of pure ESS: We have to examine the two strategies: $\bar{h}$ and $\bar{l}$. From Proposition 3, $\bar{l}$ is an ESS if and only $p > \max_{s\neq 0}\left(\frac{1-Q_s(l)}{1-Q_s(h)}\right)$ and from Proposition 2, $\bar{h}$ is an ESS if and only $p = 0$. Since, $\max_{s\neq 0}\left(\frac{1-Q_s(l)}{1-Q_s(h)}\right) > 0$, we conclude that only one of two strategies can be an ESS.

(b) Strictly mixed ESS (or interior ESS): A necessary condition for existence of an interior ESS is given by (7). Let ρ and ρ' in $(0,1)$ be two solutions of (7) i.e $\xi(p,\rho)=\xi(p,\rho')=0$. From Proposition 4, one has $V_{\rho'}(n,\alpha(\rho')) \geq V_\rho(n,\alpha(\rho')) > V_{\rho'}(n,\alpha(\rho'))$, which is a contradiction. We conclude that if an ESS exists in state-independent strategies, it is unique. □

Note that the inequality (8) is satisfied if

$$0 < p < \min_{s\neq 0}\left(\frac{1-Q_s(l)}{1-Q_s(h)}\right). \tag{9}$$

Moreover, the following result holds:

Lemma 2.

$$\min_{s\neq 0}\left(\frac{1-Q_s(l)}{1-Q_s(h)}\right) < p_0 < \max_{s\neq 0}\left(\frac{1-Q_s(l)}{1-Q_s(h)}\right).$$

Proof. Let $a = \min_{s\neq 0}\left(\frac{1-Q_s(l)}{1-Q_s(h)}\right)$, $b = \max_{s\neq 0}\left(\frac{1-Q_s(l)}{1-Q_s(h)}\right)$. Then,

$$\begin{array}{lcr}
\forall s, & b > \frac{1-Q_s(l)}{1-Q_s(h)} > a & \Longleftrightarrow \\
\forall s, & b(1-Q_s(h)) > 1-Q_s(l) > a(1-Q_s(h)) & \Longleftrightarrow \\
\forall s, & b(1-Q_s(h))-(1-Q_s(l)) > 0 > a(1-Q_s(h))-(1-Q_s(l)) & \Longleftrightarrow \\
\forall s, & \frac{b(1-Q_s(h))-(1-Q_s(l))}{(1-Q_s(l))^2} > 0 > \frac{a(1-Q_s(h))-(1-Q_s(l))}{(1-Q_s(l))^2} &
\end{array}$$

By taking the sum over s from one to n, one has,

$$b\sum_{s=1}^{n}\frac{1-Q_s(h)}{(1-Q_s(l))^2} - \sum_{s=1}^{n}\frac{1}{1-Q_s(l)} > 0 > a\sum_{s=1}^{n}\frac{1-Q_s(h)}{(1-Q_s(l))^2} - \sum_{s=1}^{n}\frac{1}{1-Q_s(l)}.$$

This means that $\xi(b,0) > \xi(p_0,0) > \xi(a,0)$. Since, the function $p \longmapsto \xi(p,\rho)$ is strictly increasing, the last inequality implies that $b > p_0 > a$. This completes the proof. □

The relation (9) gives another sufficient condition of existence of a stationary mixed ESS. By proposition 3, there is a pure nonaggressive ESS ($\bar{l}$) if and only if $p > \max_{s\neq 0}(\frac{1-Q_s(l)}{1-Q_s(h)})$. Then in some particular cases of parameters p and transition probabilities, there is no pure state-independent stationary ESS nor mixed, when

$$p_0 < p < \max_{s\neq 0}(\frac{1-Q_s(l)}{1-Q_s(h)}).$$

Let now compare the ESS, when it exists, to the global optimum solution. Somehow, we are interested in the price of anarchy of this system. A cooperative optimal strategy for players is denoted by $\widetilde{\rho}$ and is characterized by: $\widetilde{\rho} \in \arg\max_\beta V_\beta(n,\beta)$. This strategy gives the global optimum solution of the centralized system of our energy management model. We are interested in comparing the aggressiveness of the ESS to the global optimum solution.

Proposition 7. *Let ρ^* be the ESS in stationary strategies of the MDEG and $\widetilde{\rho}$ a global optimum solution. Then we have*

$$\widetilde{\rho} \le \min\left\{\frac{1}{2}, \rho^*\right\}.$$

Proof. The function

$$\beta \longmapsto V_\beta(n,\beta) = \sum_{s=1}^{n} \frac{p + (1-p)(1-\beta)\beta}{1 - Q_s(l) + \beta(Q_s(l) - Q_s(h))}$$

is continuous and strictly decreasing on $(\frac{1}{2}, 1)$. Thus, the function has a global maximizer on [0, 1] and the global maximizer is lower than $\frac{1}{2}$.

Let ρ^* be an ESS. Suppose that $\widetilde{\rho} > \rho^*$. Since ρ^* is an ESS, ρ^* satisfies $V_{\rho^*}(n,\rho^*) \ge V_{\widetilde{\rho}}(n,\rho^*)$ and $V_{\rho^*}(n,\widetilde{\rho}) > V_{\widetilde{\rho}}(n,\widetilde{\rho})$ if $V_{\rho^*}(n,\rho^*) = V_{\widetilde{\rho}}(n,\rho^*)$. Given a strategy β, the function $\alpha \longmapsto V_\beta(n,\alpha)$ is strictly decreasing. Hence, $V_{\rho^*}(n,\rho^*) > V_{\rho^*}(n,\widetilde{\rho}) > V_{\widetilde{\rho}}(n,\widetilde{\rho})$. The inequality $V_{\rho^*}(n,\rho^*) > V_{\widetilde{\rho}}(n,\widetilde{\rho})$ is a contradiction with the definition of $\widetilde{\rho}$. Hence, $\widetilde{\rho} \le \rho^*$. We conclude that a global maximizer is lower than $\frac{1}{2}$ and $\widetilde{\rho}$ coincides with ρ^* (it is the case if the ESS is a pure strategy) or $\widetilde{\rho} < \rho^*$. □

The main important result here is that the ESS is more aggressive than the global optimum solution, i.e. $\rho^* \ge \widetilde{\rho}$. This seems relatively intuitive because in a context of an evolutionary game, every player is somehow afraid to meet another player using an aggressive strategy, then the aggressive strategy is more used in the population.

4.2 State Dependent Strategies

In this section, we consider that the action taken by each player depends on his/her energy level (called state). We compute explicitly the best response correspondence $\mathrm{BR}: S\backslash\{0\} \longrightarrow 2^I$, where I is the compact interval [0, 1].

Proposition 8. *In stationary strategies, the best response to the strategy $\alpha(\beta')$ is determined by $(\beta_1, \ldots, \beta_n)$ such that for all $s \in \{1, \ldots, n\}$,*

$$\beta_s(\alpha(\beta')) = \begin{cases} 1 & \textit{if } \alpha(\beta') < 1 - \frac{p(Q_s(l)-Q_s(h))}{(1-p)(1-Q_s(l))} \\ 0 & \textit{if } \alpha(\beta') > 1 - \frac{p(Q_s(l)-Q_s(h))}{(1-p)(1-Q_s(l))} \\ \textit{any strategy } \eta_s \in [0,1] & \textit{if } \alpha(\beta') = 1 - \frac{p(Q_s(l)-Q_s(h))}{(1-p)(1-Q_s(l))} \end{cases}$$

Thus, the optimal stationary strategy $\beta = (\beta_1, \ldots, \beta_n)$ is characterized by

$$\beta_s = \begin{cases} 1 \ \textit{if } s \in I_\alpha := \left\{j,\ \frac{Q_j(l)-Q_j(h)}{(1-Q_j(l))} < (1-\alpha)\left(-1+\frac{1}{p}\right)\right\} \\ 0 \ \textit{if } s \in J_\alpha := \left\{j,\ \frac{Q_j(l)-Q_j(h)}{(1-Q_j(l))} > (1-\alpha)\left(-1+\frac{1}{p}\right)\right\} \end{cases}$$

Proof. The best reply β_s to $\alpha(\beta')$ maximizes the function $Z(\beta_s)=\frac{p+(1-p)(1-\hat{\alpha}(\beta'))\beta_s}{1-Q_s(\beta_s)}$ defined in Lemma 1. From Lemma 1, Z is monotone. The maximizer is one if

$$(1-p)(1-\hat{\alpha}(\beta'))(1-Q_s(l)) > (Q_s(l)-Q_s(h))p,$$

zero if $(1-p)(1-\hat{\alpha}(\beta'))(1-Q_s(l)) < (Q_s(l)-Q_s(h))p$ and any strategy in [0, 1] is maximizer of Z if the equality

$$(1-p)(1-\hat{\alpha}(\beta'))(1-Q_s(l)) = (Q_s(l)-Q_s(h))p$$

holds. We conclude that

$$\beta_s(\alpha(\beta')) = \begin{cases} 1 & \text{if } \alpha(\beta') < 1 - \frac{p(Q_s(l)-Q_s(h))}{(1-p)(1-Q_s(l))} \\ 0 & \text{if } \alpha(\beta') > 1 - \frac{p(Q_s(l)-Q_s(h))}{(1-p)(1-Q_s(l))} \\ \text{any strategy } \eta_s \in [0,1] & \text{if } \alpha(\beta') = 1 - \frac{p(Q_s(l)-Q_s(h))}{(1-p)(1-Q_s(l))}. \end{cases}$$

□

Let

$$f(s) = \frac{T_s(l)}{T_s(h)} = \frac{1-Q_s(h)}{1-Q_s(l)} \tag{10}$$

be the ratio between the sojourn time with the action l and the sojourn time with h in each state s. By taking assumption on the transition probabilities, we have that for every state s, $f(s) > 1$. This ratio can be interpreted also as the proportion of time obtained taking the nonaggressive action l in state s compared to the aggressive one h The inequality $\alpha(\beta') < 1 - \frac{p(Q_s(l)-Q_s(h))}{(1-p)(1-Q_s(l))} = 1 - \frac{p}{1-p}(f(s)-1)$ is equivalent to

$$(1-\alpha(\beta'))\frac{1-p}{p} > f(s) - 1, \tag{11}$$

where $f(s) - 1$ is exactly the gain in mean sojourn time in state s using action l instead of action h. Equation (11) can be rewritten: $\frac{p}{1-Q_s(l)} < \frac{1-(1-p(1-\alpha(\beta')))}{1-Q_s(h)}$, where the left side term is the average reward for a player during the time it is in state s and it uses action l, and the right term is the average reward during the time it is in state s and it uses action h. Then, the threshold policies are based on the comparison between the average instantaneous reward in each state depending on the action taken by each player which determines the instantaneous reward and also the remaining time in this state.

Definition 2 (Threshold Policies). We define the two following pure threshold policies u_1 and u_2:

- u_1 is a risky policy : there exists a state s such that $\forall s' > s, u_1(s') = L$ and $\forall s'' < s, u_1(s'') = H$. This policy is called a control limit policy in [12].
- u_2 is a careful policy : there exists a state s such that $\forall s' > s, u_2(s') = H$ and $\forall s'' < s, u_2(s'') = L$.

That kind of threshold policies, keeping the same action until the energy or state is low, has been also obtained in [12] in a context of a dynamic Hawk and Dove game. We have the following result showing that the best response strategies are u_1 or u_2 depending on the structure of the Markov Decision Process.

Proposition 9. *If the ratio $f(s)$ is increasing (resp. decreasing), then the best response policy of any player is the policy u_1 (resp. u_2).*

Proof. We assume that the function f defined in (10) is decreasing in s. If $\frac{1-Q_n(h)}{1-Q_n(l)} < (1-\alpha)(-1+\frac{1}{p})$, then

$$\beta_s = \begin{cases} \text{plays } h & \text{if } s \geq j(\alpha, p) \\ \text{plays } l & \text{if } s \leq j(\alpha, p) - 2 \end{cases}$$

else if $\frac{1-Q_n(h)}{1-Q_n(l)} > (1-\alpha)(-1+\frac{1}{p})$ then β_s consists of playing always l, where $j(\alpha, p) := \min\{j, \frac{1-Q_j(h)}{1-Q_j(l)} < (1-\alpha)(-1+\frac{1}{p})\}$. Note that one has at most one state s such that $\frac{1-Q_j(s)}{1-Q_s(l)} = (1-\alpha)(-1+\frac{1}{p})$. We have the inverse relations if $f(s)$ is increasing. □

Theorem 3 (Partially mixed equilibrium). *If there exists a state s such that $\beta_s = \frac{\alpha^* + (1-Q_s(l))(\kappa_2\alpha^* - \kappa_1)}{1-(Q_s(l)-Q_s(h))(\kappa_2\alpha^* - \kappa_1)} \in (0,1)$ where*

$$\kappa_1 = \sum_{j \in I} \frac{1}{1 - Q_j(h)},\ \kappa_2 = \kappa_1 + \sum_{j \notin I,\ j \neq s} \frac{1}{1 - Q_j(l)},\ \alpha^* = 1 - \frac{p(Q_s(l) - Q_s(h))}{(1-p)(1 - Q_s(l))},$$

$$I = \{j \neq s,\ f(s) > f(j)\}$$

and $\beta_j,\ j \neq s$ *given by*

$$\beta_j = \left\{ \begin{array}{ll} \textit{plays } h \textit{ if } \frac{(Q_s(l) - Q_s(h))}{(1 - Q_s(l))} > \frac{(Q_j(l) - Q_j(h))}{(1 - Q_j(l))} & \textit{i.e } f(s) > f(j) \\ \textit{plays } l & \textit{otherwise} \end{array} \right\},$$

then $\beta = (\beta_1, \ldots, \beta_n)$ *is an equilibrium.*

Proof. By Proposition 8, the best reply has the form

$$BR(\beta_s) = \begin{cases} \text{play action } h & \text{if } \alpha(\beta_s) < 1 - \frac{p(Q_s(l) - Q_s(h))}{(1-p)(1-Q_s(l))}, \\ \text{play action } l & \text{if } \alpha(\beta_s) > 1 - \frac{p(Q_s(l) - Q_s(h))}{(1-p)(1-Q_s(l))}, \\ \text{any strategy } h, l \text{ or mixed} & \text{if } \alpha(\beta_s) = 1 - \frac{p(Q_s(l) - Q_s(h))}{(1-p)(1-Q_s(l))}. \end{cases}$$

The fixed point equation gives $\beta \in BR(\beta)$ with $\alpha(\beta) = \frac{\sum_{s=1}^{n} \frac{\beta_s}{1 - Q_s(\beta_s)}}{\sum_{s=1}^{n} \frac{1}{1 - Q_s(\beta_s)}}$. Since f is monotone and $Q_s(l) < Q_s(h)$, there exist at most one state s_0 such that

$$\frac{1 - Q_{s_0}(h)}{1 - Q_{s_0}(l)} = (1 - \alpha(\beta)) \left(-1 + \frac{1}{p} \right) \tag{12}$$

and a mixed equilibrium is characterized by

$$\beta'_s = \begin{cases} \text{play action } h & \text{if } s' > s_0, \\ \text{play action } l & \text{if } s' < s_0, \\ \text{any strategy } h, l \text{ or mixed} & \text{if } s' = s_0, \end{cases}$$

where β_{s_0} satisfies (12). After some basic calculations, (12) has a unique solution given by

$$\beta_{s_0} = \frac{\alpha^* + (1 - Q_s(l))(\kappa_2 \alpha^* - \kappa_1)}{1 - (Q_s(l) - Q_s(h))(\kappa_2 \alpha^* - \kappa_1)} \in (0, 1)$$

with $\kappa_1 = \sum_{j \in I} \frac{1}{1 - Q_j(h)}$, $\kappa_2 = \kappa_1 + \sum_{j \notin I,\ j \neq s} \frac{1}{1 - Q_j(l)}$, $\alpha^* = 1 - \frac{p(Q_s(l) - Q_s(h))}{(1-p)(1-Q_s(l))}$, $I = \{j \neq s,\ f(s) > f(j)\}$. □

Note that a necessary condition of existence such s is that $p > \frac{(1-c)(1-Q_s(l))}{(1 - Q_s(h) - c(1 - Q_s(l)))}$ with $c = \frac{1 + \kappa_1(1 - Q_s(h))}{1 + \kappa_2(1 - Qs(h))} \in (0, 1)$. We then obtain several equilibria by permutation on the state satisfying $\beta_s = \frac{\alpha^* + (1 - Q_s(l))(\kappa_2 \alpha^* - \kappa_1)}{1 - (Q_s(l) - Q_s(h))(\kappa_2 \alpha^* - \kappa_1)} \in (0, 1)$.

Lemma 3. *For all state s, the probability of choosing action h at the previous mixed equilibrium point, namely β_s, is decreasing in p.*

Proof. It suffices to proof in the unique state s in which $\beta_s \in (0,1)$. The fraction of players with the aggressive action in state s is then given by

$$\beta_s = \frac{\alpha^* + (1 - Q_s(l))(\kappa_2\alpha^* - \kappa_1)}{1 - (Q_s(l) - Q_s(h))(\kappa_2\alpha^* - \kappa_1)} \in (0,1),$$

where $\kappa_1 = \sum_{j\in I} \frac{1}{1-Q_j(h)}$, $\kappa_2 = \kappa_1 + \sum_{j\notin I,\ j\neq s} \frac{1}{1-Q_j(l)}$, $\alpha^* = 1 - \frac{p(Q_s(l)-Q_s(h))}{(1-p)(1-Q_s(l))}$, i $= \{j \neq s,\ f(s) > f(j)\}$. Since the function $p \longmapsto \alpha^*(p) = 1 - \frac{p(Q_s(l)-Q_s(h))}{(1-p)(1-Q_s(l))}$ is strictly decreasing, the denominator is nondecreasing in p and the numerator is strictly decreasing in p. Thus, β_s is decreasing in p. This completes the proof. □

4.3 Dynamics

By proposition 9 we know that, when the function f is monotone, the best response strategies have a given structure defined as threshold policies u_1 or u_2. To numerically observe the equilibrium policies in the population, we construct the replicator dynamics which has properties of convergence to equilibrium and ESS. A replicator dynamic is a differential equation that describes the way strategies change in time as a function of the fitness. Roughly speaking, they are based on the idea that the average growth rate per player that uses a given action is proportional to the excess of fitness of that action with respect to the average fitness.

When f is monotone increasing (respectively decreasing) the set of policies are $u_1^1, \ldots, u_1^{n+1}$ (respectively $u_2^1, \ldots, u_2^{n+1}$) where policy u_1^i (resp. u_2^i) consists of taking action h (resp. l) in the $i-1$ first states. For all $i = 1, \ldots, n+1$, the proportion of the population playing u_1^i is denoted by σ_i. The replicator dynamics describes the evolution of the different policies in the population in time and is given by the following differential equation:

$$\begin{aligned}\dot{\sigma}_i(t) &= \sigma_i(t)\left[\sum_{j=1}^{n+1}\sigma_j(t)V_{u_1^i}(n,u_1^j) - \sum_{k=1}^{n+1}\sum_{j=1}^{n+1}\sigma_k(t)\sigma_j(t)V_{u_1^k}(n,u_1^j)\right] \\ &=: f_i(\sigma(t)). \end{aligned} \tag{13}$$

The replicator dynamics satisfies the so-called *positive correlation* condition [15]:

$$f(\sigma) \neq 0 \Longrightarrow \sum_j G_{u_1^j}(\sigma) f_j(\sigma) > 0.$$

It follows from [8, 15] that any equilibrium σ^* (in the $n-$simplex of the $(n+1)-$dimensional Euclidean space $\mathbb{R}^{n+1}$) of the evolutionary game is a rest point of the replicator dynamics. To describe the long-term behavior of the dynamics, we shall say that a stationary point (or rest point) σ^* is *stable* under (13) if for every neighborhood $\mathcal{N}$ of σ^* there exists a neighborhood $\mathcal{N}' \subseteqq \mathcal{N}$ of σ^* such that $\sigma^* \in \mathcal{N}' \Longrightarrow \sigma(t) \in \mathcal{N},\ \forall t \geq 0$. If σ^* is a stable rest point of (13), then σ^* is an equilibrium.

5 Numerical Illustrations

We consider an evolutionary game where each player has $n = 5$ states. Each player starts his/her lifetime full of energy (state 5) and has to decide between an aggressive or a nonaggressive action at each pairwise interaction with another player. We assume that all players are in the same population that means that all players have the same transition probabilities Q.

We consider the following transition probabilities

$$Q = \begin{pmatrix} 0.1 & 0.05 \\ 0.2 & 0.1 \\ 0.5 & 0.4 \\ 0.7 & 0.6 \\ 0.8 & 0.7 \end{pmatrix},$$

where for $i = 1, \ldots, 5$, Q_{i1} (resp. Q_{i2}) is the probability to remain in state i using action l (resp. h).

5.1 State-Independent Strategy

Each player chooses action h with a probability that does not depend on his/her state. We denote this probability by ρ. First, using the transition probability matrix Q, the threshold p_0 given by (8) is $p_0 = 0.75$. Then from proposition 5 we know that if the probability p is less than 0.75, there exists an ESS. Moreover, this ESS is the unique solution of the system defined in (7).

Second, we have $M = \max_s \frac{1-Q_s(l)}{1-Q_s(h)} = 0.9474$, then using proposition 3 the subset of states I_1 is empty. Then, we have that if $p > 0.9474$ the unique ESS is to take always action l, that is $\rho = 0$.

We observe also on Fig. 1 that taking the pure aggressive action h is only an ESS when $p = 0$. It is somehow intuitive because if the system is such that every player meets always another player, the best strategy to survive is to be aggressive. Related

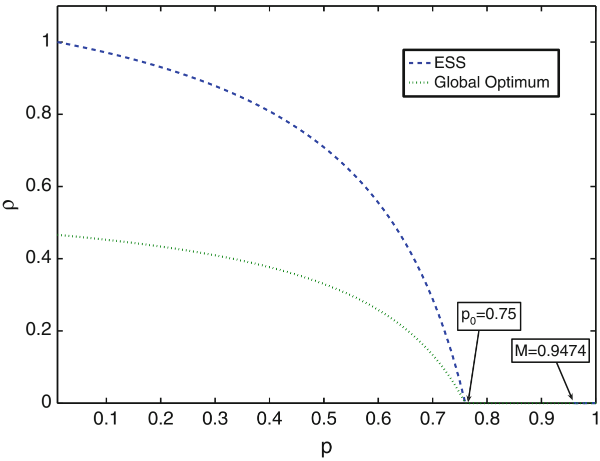


Fig. 1 ESS ρ with state-independent actions

to this comment, we observe also that the probability to choose an aggressive action is monotone decreasing with p.

Finally, we also verify the result of the proposition 7 comparing the ESS with the global optimum solution.

5.2 State-Dependent Strategies

We consider that each player decides to take an aggressive action depending on his/her state. Using the transition probability matrix Q, the function f defined by (10), is strictly increasing in s. Then the best response policy is u_1. We describe the set of policies $u_1^1, u_1^2, \ldots, u_1^6$ where the policy u_1^i has 5 components which are the action in each state s from 1 to 5. For example, $u_1^3 = (hhhll)$ which means the user takes action h in states 1, 2, and 3; and action l in states 4 and 5.

In Fig. 2, the replicator dynamics converge when $p = 0.5$ to the pure policy u_1^2, i.e. each player takes action l only in the full energy level (state 5). In Fig. 3, the replicator dynamics converge to a mixed strategy between policies u_1^1 and u_1^2 where 87.41% of the population uses policy u_1^1 and 12.59% policy u_1^2. When $p = 0.9$, we observe in Fig. 4 the replicator dynamics converge to the pure policy u_1^5 which consists of taking action h only in state 1. Moreover, we observe that every rest point of the replicator dynamics in the three cases is stable and then there is an equilibrium.

Finally, with those examples, we have observed that when the probability to meet another player decreases (means p increases), players become less aggressive and the equilibrium tends to nonaggressive policies.

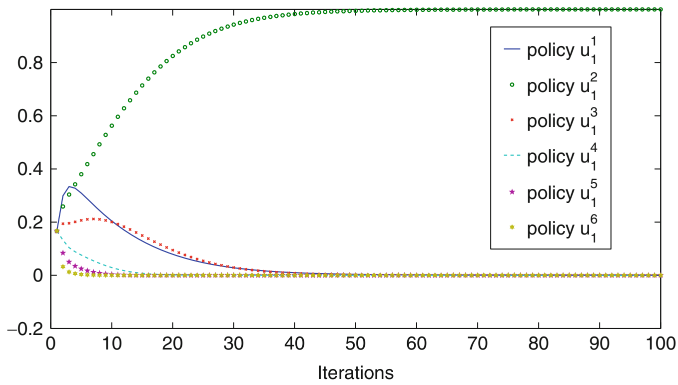


Fig. 2 Replicator dynamic with $p = 0.5$

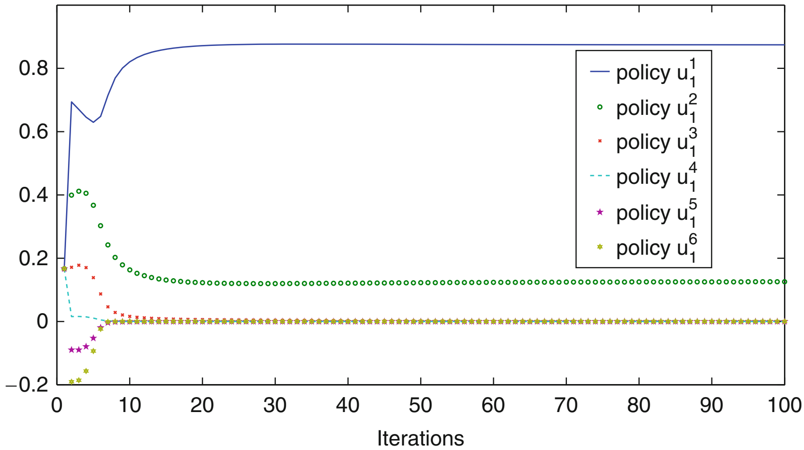


Fig. 3 Replicator dynamic with $p = 0.1$

6 Conclusion

In the paper, an energy management in noncooperative population game under an evolutionary Markov game framework has been studied. A problem considering the stochastic evolutionary games where each player can be in different state during his/her life and has the possibility to take several actions in each individual state has been presented. Those actions have an impact not only on the instantaneous fitness of the player but also on its future individual's state. Restricting the game to stationary mixed policies, we have determined explicitly the ESS of this Markov Decision Evolutionary Game (MDEG). Considering more general state-dependent policies, a threshold structure of the best response policies has been obtained. The convergence of the replicator dynamics to an ESS has been studied numerically.

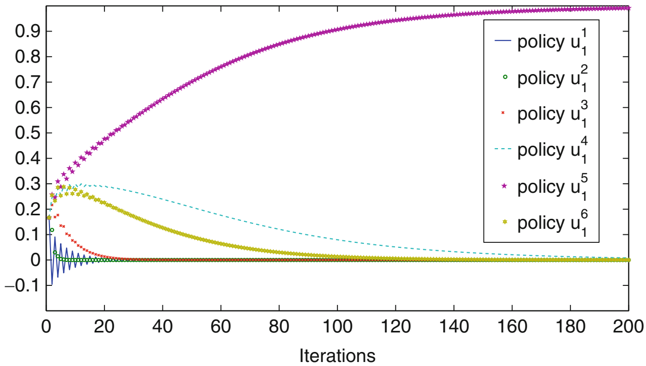


Fig. 4 Replicator dynamic with $p = 0.9$

Acknowledgements This work has been partially supported by the European Commission within the framework of the BIONETS project, IST-FET-SAC-FP6-027748, www.bionets.eu. It was also supported in part by the Arc Popeye collaborative project of INRIA.

References

1. Altman, E., Hayel, Y.: A Stochastic Evolutionary Game of Energy Management in a Distributed Aloha Network. In: Proceedings of IEEE INFOCOM (2008)
2. Altman, E., Hayel, Y., Tembine, H., El Azouzi, R.: Markov Decision Evolutionary Games with Expected Average Fitness. In: Brown, J.S., Vincent, T.L.S. (eds.) Evolutionary Ecology Research Journal, Special issue in honour of Thomas Vincent (2009)
3. Altman, E., El Azouzi, R., Hayel, Y., Tembine, H.: The Evolution of Transport Protocols: An Evolutionary Game Perspective. Computer Networks **53**10, 1751–1759 (2009)
4. Axelrod, R.: The Evolution of Cooperation (Revised ed.). Basic Books, New York, NY (2006)
5. Haigh, J.: Game theory and evolution. Advances in Applied Probability **7**(1), 8–11 (1975)
6. Hayel, Y., Tembine, H., Altman, E., El Azouzi, R.: A Markov Decision Evolutionary Game for Individual Energy Management. Technical Report, CERI (2009)
7. Hofbauer, J., Sigmund, K.: Evolutionary Games and Population Dynamics. Cambridge University Press, Cambridge, UK (1998)
8. Hofbauer, J., Sigmund, K.: Evolutionary game dynamics. American Mathematical Society **40**(4) (2003)
9. Houston, A.I., McNamara, J.M.: Evolutionarily stable strategies in the repeated hawk-dove game. Behavioral Ecology **2**(3), 219–227 (1991)
10. Maynard Smith, J.: Game Theory and the Evolution of Fighting. In: Maynard Smith, J. (ed) On Evolution, pp. 8–28. Edinburgh University Press, Edinburgh (1972)
11. Maynard Smith, J.: Evolution and the Theory of Games. Cambridge University Press, Cambridge, UK (1982)
12. McNamara, J.M., Merad, S., Collins, E.J.: The Hawk-Dove Game as an Average-Cost Problem. Adv. Appl. Probab. **23**(4), 667–682 (1984)
13. Puterman, M.: Markov Decision Processes: Discrete Stochastic Dynamic Programming. Wiley, New York, NY (1994)
14. Sandholm, W.H.: Potential games with continuous player sets. Journal of Economic Theory **97**(1), 81–108 (2001)

15. Sandholm, W.H.: Population Games and Evolutionary Dynamics. MIT, Cambridge, MA (to be published)
16. Susu, A., Acquaviva, A., Atienza, D., De Micheli, G.: Stochastic Modeling and Analysis for Environmentally Powered Wireless Sensor Nodes. Proceedings of WiOpt (2008)
17. Tembine, H., Altman, E., El Azouzi, R., Sandholm, W.H.: Evolutionary game dynamics with migration for hybrid power control in wireless communications. 47th IEEE Conference on Decision and Control, pp. 4479–4484, Cancun, Mexico (2008)
18. Tembine, H., Altman, e., El Azouzi, R., Hayel, Y.: Evolutionary Games in Wireless Networks. IEEE Transactions on Systems, Man, and Cybernetics: Part B, special issue on Game Theory
19. Vincent, T.L., Vincent, T.L.S.: Evolution and control system design. IEEE Control Systems Magazine **20**(5), 20–35 (2000)
20. Weibull, J.W.: Evolutionary Game Theory. MIT, Cambridge, MA (1995)

Mutual Mate Choice with Multiple Criteria

David M. Ramsey

Abstract This article presents a model of mutual mate search based on two trait measures. One measure describes the attractiveness of an individual and preferences are common according to this measure i.e., each female prefers highly attractive males and all females agree as to which males are attractive. Preferences are homotypic with respect to the second measure, referred to as character i.e., all individuals prefer mates of a similar character. It is assumed that attractiveness is easy to measure, but to observe the character of a prospective partner, it is necessary to court. Hence, on meeting a prospective partner an individual must decide whether to try and court the other. Courtship only occurs by mutual consent. During courtship, individuals observe the character of the prospective partner and then decide whether to mate or not. Mutual acceptance is required for mating to occur. This paper presents the model and outlines a procedure for finding a Nash equilibrium which satisfies a set of criteria based on the concept of subgame perfection. Two examples are presented and it is shown that multiple equilibria may exist.

1 Introduction

Kalick and Hamilton [19] noted that there is a large correlation between the perceived attractiveness of human partners. It has been proposed that this is due to individuals preferring prospective partners of similar attractiveness, i.e., they have homotypic preferences. However, such assortative mating occurs when all individuals prefer more attractive mates, i.e., they have common preferences (see, for example, Alpern and Reyniers [2], Johnstone [18], McNamara and Collins [22], Parker [24], Real [27], Simão and Todd [32] and [33]). Common preferences have often been observed in the field. Bakker and Milinski [5] observe that female sticklebacks prefer brightly colored males. Monaghan et al. [23] note that male finches

D.M. Ramsey (✉)
Department of Mathematics and Statistics, University of Limerick, Limerick, Ireland
e-mail: david.ramsey@ul.ie

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_17,

prefer fecund females. However, humans tend to choose partners with similar religious and political views. Such preferences must be regarded as homotypic. Cooke and Davies [10] observe that snow geese prefer partners who are similarly colored to their parents. Robertson [28] notes that female frogs benefit from mating with a male of appropriate size to theirs. In support of this, Ryan et al. [29] note that female cricket frogs are more likely to approach loudspeakers emitting the call of a similar sized male. Although many models of mate choice have been considered in which individuals have common preferences and some work has been done on mate choice with homotypic preferences, little consideration has been given to models in which mate choice involves both types of preference.

Janetos [17] was the first to present a model of mate choice with common preferences. He assumed that only females are choosy and the value of a male to a female comes from a distribution known to the females. There is a fixed cost for observing each prospective mate, but there is no limit on the number of males a female can observe. Real [26] developed these ideas.

In many species both sexes are choosy and such problems are game theoretic. Parker [24] presents a model in which both sexes prefer mates of high value. He concludes that assortative mating should occur with individuals being divided into classes. Class i males are paired with class i females and there may be one class of males or females who do not mate. Unlike the models of Janetos [17] and Real [26], Parker's model did not assume that individuals observe a sequence of prospective mates. In the mathematics and economics literature such problems are formulated as marriage problems or job search problems. McNamara and Collins [22] consider a job search game in which job seekers observe a sequence of job offers and, correspondingly, employers observe a sequence of candidates. Both groups have a fixed cost of observing a candidate or employer, as appropriate. Their conclusions are similar to those of Parker [24]. Real [27] developed these ideas within the framework of mate choice problems. For similar problems in the economics literature see e.g., Shimer and Smith [31] and Smith [34].

In the above models, it is assumed that the distribution of the value of prospective partners has reached a steady state. There may be a mating season and as it progresses the distribution of the value of available partners changes. Collins and McNamara [8] were the first to formulate such a model as a one-sided job search problem with continuous time. Ramsey [25] considers a similar problem with discrete time. Johnstone [18] presents numerical results for a discrete time, two-sided mate choice problem with a finite horizon. Alpern and Reyniers [2] and Alpern and Katrantzi [3] use a more analytic approach to similar mate choice problems. Burdett and Coles [7] consider a dynamic model in which the outflow resulting from partnership formation is balanced by job seekers and employers coming into the employment market.

These models all assume that individuals know their own attractiveness and the distribution of the attractiveness of prospective mates. Intrinsic knowledge of the distribution of mates' attractiveness is reasonable if this distribution does not change over time, as individuals' search rules evolve to adapt to this distribution. However, in reality, this distribution varies in time and/or space due to variability in, e.g., the climate and availability of resources. Mazalov et al. [21] and Collins et al. [9] consider models in which only females are choosy and the distribution of the value of males is unknown. When mate choice is two sided, it is important whether an individual knows its own quality or not, since the strategy an individual uses should be dependent on its quality. Fawcett and Bleay [13] consider a model of two-sided mate choice in which an individual implicitly learns about its own quality from the reactions of prospective partners. Several such models and results of simulations have been presented in the psychology and artificial intelligence literature. Kalick and Hamilton [19] used simulations assuming mate choice is based on attractiveness to generate strategies that match patterns of mating in the human population. Their model was discussed and developed by Simão and Todd [32] and [33].

Alpern and Reyniers [1] consider a model in which individuals have homotypic preferences. Mate choice is based on a numeric trait x with the same distribution in both sexes. The cost of mating is assumed to be the absolute difference between the traits of the partners. The time horizon is finite and there is a sufficiently large cost for not mating to ensure that in the last search period an individual prefers mating with the most dissimilar partner to not mating at all.

It has been observed that mate choice can be based on multiple criteria. Kodric-Brown [20] notes that female choice is based on the dominance status of a male, together with his coloration and courtship behavior. Iwasa and Pomiankowski [16] model the evolution of female preferences for multiple sexual ornaments. It has been argued that traits can be combined to define a one-dimensional measure of attractiveness. However, it may be easier to observe some traits than others and one trait may be a better indicator of quality than another. Fawcett and Johnstone [12] consider a model in which only females are choosy and can use two signals of male quality. The costs of observing these two signals, as well as a signal's reliability differ. If signals are not very reliable (but also not very unreliable), a female should first observe the signal that is cheapest to evaluate. If this signal indicates that a male is of high quality, then the female should observe the second signal. It seems that such a procedure is followed by female fiddler crabs. If a male is sufficiently large, then they will inspect the burrow he has prepared (see Backwell and Passmore [4]).

This paper presents a model of mate choice involving both common and homotypic preferences. All individuals prefer mates of high attractiveness who have a similar character to themselves. It is assumed that individuals know their own attractiveness and character. The distributions of character and attractiveness are discrete with finite supports and constant over time. Together, the attractiveness, character, and sex of an individual determine its type. Individuals can observe the attractiveness and character of prospective mates perfectly. Attractiveness can be observed quickly, but to measure character, a courtship period is required. It is assumed that courtship is costly and search costs incorporate the costs of finding a prospective mate and of observing his/her attractiveness. Individuals decide whether to court on the basis of attractiveness. Courtship occurs only by mutual consent. After courtship, a couple form a breeding pair only by mutual consent, based on both attractiveness and character. At equilibrium each individual uses a strategy appropriate to its type. This strategy defines both the set of prospective partners an individual is willing to court and the set of prospective partners an individual is willing to pair with after courtship. The set of strategies corresponding to such an equilibrium is called an equilibrium strategy profile.

Section 2 presents the model, together with a set of criteria that we wish an equilibrium to satisfy based on subgame perfection.

Section 3 describes a general method for calculating the expected rewards of individuals under a given strategy profile. Section 4 considers the courtship subgame (when courting individuals decide whether to mate or not) and the offer/acceptance subgame (when individuals decide whether to court or not).

Section 5 considers so-called quasi-symmetric games, in which the distributions of attractiveness and character, as well as the search and dating costs, do not depend on sex. An algorithm for deriving a symmetric equilibrium is presented. A simple example is presented.

Section 6 presents results for a larger scale problem and shows that there may be multiple equilibria. Section 7 suggests directions for further research.

2 The Model

It is assumed that mate choice is based on two traits. The first is referred to as attractiveness and the second as character. We consider a steady-state model in which the distributions of attractiveness and character according to sex (denoted $X_{1,s}$, $X_{2,s}$, where $s \in \{m, f\}$) do not change over time. We suppose that the $X_{k,s}$, $k \in \{1, 2\}$ are discrete random variables with finite supports. Attractiveness is independent of character. The type of an individual is defined by its attractiveness, character, and sex. The type of a male is denoted $\mathbf{x}_m = [x_{1,m}, x_{2,m}]$. The type of a female is denoted $\mathbf{x}_f = [x_{1,f}, x_{2,f}]$. Each individual is assumed to know its own type.

Preferences are common with respect to attractiveness, i.e., all individuals prefer mates of high attractiveness. Preferences are homotypic with respect to character, i.e., an individual prefers mates of a similar character. Suppose the reward obtained by a type $\mathbf{x}_m$ male from mating with a type $\mathbf{x}_f$ female is $g(x_{2,m}, \mathbf{x}_f)$, where g is strictly increasing with respect to $x_{1,f}$ and strictly decreasing with respect to $|x_{2,m} - x_{2,f}|$. The reward obtained by a type $\mathbf{x}_f$ female from mating with a type $\mathbf{x}_m$ male is $h(\mathbf{x}_m, x_{2,f})$.

The population is assumed to be large. At each moment n ($n = 1, 2, \ldots$), each unmated individual is presented with a prospective mate picked at random. Thus, it is implicitly assumed that the operational sex ratio, r, (the number of searching males divided by the number of searching females) is one and individuals can observe as many prospective mates as they wish. Suppose $r > 1$. This can be modeled by assuming that a proportion $\frac{r-1}{r}$ of females give a reward $-\infty$ (in reality males paired with such a female meet no female). An individual cannot return to a prospective mate found earlier.

Suppose that individuals can observe the type of a prospective mate perfectly. Furthermore, attractiveness can be measured almost instantaneously, but courtship is required to observe character. The search costs incurred at each moment by males and females are $c_{1,m}$ and $c_{1,f}$, respectively. The costs of courting are $c_{2,m}$ and $c_{2,f}$ to males and females, respectively. Seymour and Sozou [30] and Sozou [35] model how courtship procedures can help to gain information. For a review of empirical studies in this field, see Vahed [37].

On meeting a female, a male must decide whether to court her, based on her attractiveness. If the male wishes to court, the female then decides whether to proceed with courtship, based on his attractiveness. This means that the game played between the male and female is essentially asymmetric (the importance of this is considered later in the section on the offer/acceptance subgame). Courtship occurs only by mutual consent. It might pay individuals to immediately accept a prospective partner as a mate without courtship. However, to keep the strategy space as simple as possible, it is assumed that individuals must court before forming a pair.

During courtship, an individual observes the character of its prospective partner. Each individual then decides whether to accept the other as a partner. If acceptance is mutual, a mating pair is formed. Otherwise, both continue searching. Since at this stage both individuals have perfect information on the type of the other, it may be assumed that these decisions are made simultaneously.

The total reward is assumed to be the reward gained from mating minus the search costs incurred. Hence, the total reward of a male of type $\mathbf{x}_m$ from mating with a female of type $\mathbf{x}_f$ after searching for n_1 moments and courting n_2 females is $g(x_{2,m}, \mathbf{x}_f) - n_1 c_{1,m} - n_2 c_{2,m}$. The total reward of a female is defined analogously.

The search game played by the population as a whole is called the supergame and denoted Γ. In Γ, each individual plays a sequence of games with prospective mates until he/she finds a mate. This supergame depends on the distributions of attractiveness and character in both sexes and the search and courting costs.

Assume that the population follow some strategy profile, denoted π. Such a profile defines the strategy to be used in Γ by each individual according to their type (assumed to be a pure strategy). An individual's strategy must define the following:

1. Which levels of attractiveness induce the offer or acceptance of courtship (as appropriate).
2. Which types of prospective partners an individual is prepared to mate with after courtship (this description should consider all types of the opposite sex, even those with which courtship would not occur).

The game played by a male and female on meeting can be split into two subgames. The first is referred to as the offer/acceptance subgame, where they decide whether to court. The second is called the courting game, where both decide whether to accept the other as a mate. These two subgames together define the game played when a male of type $\mathbf{x}_m$ meets a female of type $\mathbf{x}_f$, denoted $G(\mathbf{x}_m, \mathbf{x}_f; \pi)$.

When the population play according to a Nash equilibrium profile π^*, then no individual can gain by using a different strategy to the one defined by π^*. We look for a profile π^* which satisfies the following criteria:

Condition 1 In the courting game, an individual accepts a prospective mate if and only if the reward from such a pairing is at least the expected reward from future search.

Condition 2 An individual offers (accepts, as appropriate) courtship, if the expected reward from the resulting courting subgame (including courtship costs) is at least the individual's expected reward from future search.

Condition 3 The decisions made by an individual do not depend on the moment at which the decision is made.

Conditions 1 and 2 are necessary and sufficient conditions to ensure that when the population use the profile π^*, each individual plays according to a subgame perfect equilibrium in the appropriately defined offer/acceptance and courting subgames.

The most preferred mates of an individual are those of maximum attractiveness who have the most similar character. Condition 1 states that in the courting game an individual will always accept his/her most preferred mate. In addition, if an individual accepts a prospective partner who gives him/her a payoff of k, then he/she would accept any other partner who gives at least k.

Condition 3 states that the Nash equilibrium profile should be stationary. This reflects the following facts: (1) An individual starting to search at moment i faces the same problem as one starting at moment 1. (2) Since the search costs are linear, after searching for i moments and not finding a mate, an individual maximizes his/her expected reward from search simply by maximizing the expected reward from future search.

Many strategy profiles lead to the same pattern of courting and mate formation. For example, suppose that there are three levels of attractiveness and character. Consider π_1, the strategy profile according to which:

a. Individuals are willing to court those of at least the same attractiveness.
b. In the courting game individuals of the two extreme characters accept prospective partners of either the same character or the central character, individuals of the central character only accept prospective partners of the same character.

Under π_1 only individuals of the same attractiveness court and matings only occur between individuals of the same type. Suppose under π_2 individuals only court prospective partners of the same attractiveness and are prepared to mate with those of the same type. The pattern of courting and mating observed under these two strategy profiles is the same.

We might be interested in Nash equilibrium profiles that satisfy the following conditions:

Condition 4 A male of type $\mathbf{x}_m$ is willing to court any female of attractiveness $\geq t_m(\mathbf{x}_m)$, i.e., males use threshold rules to initiate courtship.

Condition 5 A female of type $\mathbf{x}_f$ is willing to be courted by any male of attractiveness $\geq t_f(\mathbf{x}_f)$.

In general, such equilibrium profiles would not satisfy Conditions 1–3. Under a profile satisfying Condition 5, a female of low attractiveness would be willing to court a male of high attractiveness, although he would never accept her in the courting game. However, we expect that under the forces of natural selection, individuals do not court prospective mates they are sure to reject in the courting game. Thus, a female of low attractiveness would not have to choose whether to enter into courtship with a highly attractive male. Nash equilibria that satisfy Conditions 4 and 5 might be desirable, since under such a profile a male will never offer courtship to one female and not offer courtship to a more attractive female (i.e., expected to be a better partner since the male does not have any information on her character).

It may pay males of low attractiveness not to seek courtship with highly attractive females to avoid unnecessary courtship costs or when there is active competition for mates between individuals of the same sex (see Fawcett and Johnstone [11] and Härdling and Kokko [15]).

3 Deriving the Expected Payoffs Under a Given Strategy Profile

Given the strategy profile used, we can define which pairs of types of individuals court and which form mating pairs. We can thus calculate the expected length of search and the expected number of courtships for each individual. Let $p(\mathbf{x}_m)$ be the probability that a male is of type $\mathbf{x}_m$ and $q(\mathbf{x}_f)$ be the probability that a female is of type $\mathbf{x}_f$. Let $F_1(\mathbf{x}_m;\pi)$ be the set of types of female that a male of type $\mathbf{x}_m$ will court under the strategy profile π and the requirement of mutual consent. Define $F_2(\mathbf{x}_m;\pi)$ to be the set of types of female that pair with a male of type $\mathbf{x}_m$. The sets $M_1(\mathbf{x}_f;\pi)$ and $M_2(\mathbf{x}_f;\pi)$ are defined analogously. By definition, $F_2(\mathbf{x}_m;\pi) \subseteq F_1(\mathbf{x}_m;\pi)$, $M_2(\mathbf{x}_f;\pi) \subseteq M_1(\mathbf{x}_f;\pi)$.

Theorem 1. *The expected length of search of a male of type $\mathbf{x}_m$, $L_m(\mathbf{x}_m;\pi)$, is the reciprocal of the probability of finding a female in $F_2(\mathbf{x}_m;\pi)$ at each stage. The expected number of courtships of this male, $D_m(\mathbf{x}_m;\pi)$, is the expected length of search times the probability of courting at each stage. Hence,*

$$L_m(\mathbf{x}_m;\pi) = \frac{1}{\sum_{\mathbf{x}_f \in F_2(\mathbf{x}_m;\pi)} q(\mathbf{x}_f)}; \qquad D_m(\mathbf{x}_m;\pi) = \frac{\sum_{\mathbf{x}_f \in F_1(\mathbf{x}_m;\pi)} q(\mathbf{x}_f)}{\sum_{\mathbf{x}_f \in F_2(\mathbf{x}_m;\pi)} q(\mathbf{x}_f)}.$$

The expected length of search and the expected number of courtships of a female can be calculated analogously.

The expected reward from mating of a type $\mathbf{x}_m$ male under the profile π, $R_m(\mathbf{x}_m;\pi)$, is the expected reward from mating given that the female's type is in $F_2(\mathbf{x}_m;\pi)$. Hence,

$$R_m(\mathbf{x}_m;\pi) = \frac{\sum_{\mathbf{x}_f \in F_2(\mathbf{x}_m;\pi)} q(\mathbf{x}_f) g(x_{2,m},\mathbf{x}_f)}{\sum_{\mathbf{x}_f \in F_2(\mathbf{x}_m;\pi)} q(\mathbf{x}_f)} - c_{1,m} L_m(\mathbf{x}_m;\pi) - c_{2,m} D_m(\mathbf{x}_m;\pi).$$

The expected reward of a female can be calculated analogously.

Definition 1. A value of Γ is given by the set of expected rewards of each individual according to type at a Nash equilibrium profile satisfying Criteria 1–3, i.e., $\{R_m(\mathbf{x}_m;\pi^*)\}_{\mathbf{x}_m \in M}$ and $\{R_f(\mathbf{x}_f;\pi^*)\}_{\mathbf{x}_f \in F}$, where M and F are the sets of types of males and females, respectively. A value of the game to an individual of a given type is given by the appropriate element from this set.

4 The Courting and Offer/Acceptance Subgames

4.1 The Courting Subgame

Assume that the population are following profile π. The male and female both have two possible actions: accept (i.e., mate), denoted a, or reject denoted r. We may assume that these decisions are taken simultaneously. We ignore the costs already incurred by both individuals, as they are subtracted from all the payoffs in the matrix and hence do not affect the equilibria in this subgame.

Suppose the male is of type $\mathbf{x}_m$ and the female is of type $\mathbf{x}_f$. The payoff matrix is

$$\begin{array}{r c} & \begin{array}{cc} \text{Female: } a \qquad\qquad & \qquad\qquad \text{Female: } r \end{array} \\ \begin{array}{r} \text{Male: } a \\ \text{Male: } r \end{array} & \begin{pmatrix} [g(x_{2,m},\mathbf{x}_f), h(\mathbf{x}_m, x_{2,f})] & [R_m(\mathbf{x}_m;\pi), R_f(\mathbf{x}_f;\pi)] \\ [R_m(\mathbf{x}_m;\pi), R_f(\mathbf{x}_f;\pi)] & [R_m(\mathbf{x}_m;\pi), R_f(\mathbf{x}_f;\pi)] \end{pmatrix}. \end{array}$$

The appropriate Nash equilibrium of this subgame is for the male to accept the female if and only if $g(x_{2,m},\mathbf{x}_f) \geq R_m(\mathbf{x}_m;\pi)$ and the female to accept the male if and only if $h(\mathbf{x}_m, x_{2,f}) \geq R_f(\mathbf{x}_f;\pi)$. Note that for convenience it is assumed that when an individual is indifferent between rejecting and accepting, then he/she accepts. Using the rule given above, a searcher will take the optimal action whenever the prospective partner "mistakenly" accepts.

Let $\mathbf{v}(\mathbf{x}_m,\mathbf{x}_f;\pi) = [v_m(\mathbf{x}_m,\mathbf{x}_f;\pi), v_f(\mathbf{x}_m,\mathbf{x}_f;\pi)]$ denote the value of this game, where $v_m(\mathbf{x}_m,\mathbf{x}_f;\pi)$ and $v_f(\mathbf{x}_m,\mathbf{x}_f;\pi)$ are the values of the game to the male and female, respectively. If $g(x_{2,m},\mathbf{x}_f) \geq R_m(\mathbf{x}_m;\pi)$ and $h(\mathbf{x}_m, x_{2,f}) \geq R_f(\mathbf{x}_f;\pi)$, then $v_m(\mathbf{x}_m,\mathbf{x}_f;\pi) = g(x_{2,m},\mathbf{x}_f)$ and $v_f(\mathbf{x}_m,\mathbf{x}_f;\pi) = h(\mathbf{x}_m, x_{2,f})$. Otherwise, $v_m(\mathbf{x}_m,\mathbf{x}_f;\pi) = R_m(\mathbf{x}_m;\pi)$ and $v_f(\mathbf{x}_m,\mathbf{x}_f;\pi) = R_f(\mathbf{x}_f;\pi)$.

At an equilibrium profile of Γ, the appropriate Nash equilibrium must be played in all the possible courting subgames (even if such courtship games cannot occur).

4.2 The Offer/Acceptance Subgame

Once the courting subgame has been solved, we may solve the offer/acceptance subgame and hence the game $G(\mathbf{x}_m,\mathbf{x}_f;\pi)$. The possible actions of a male are n – do not try to court and o – try to court. The possible actions of a female

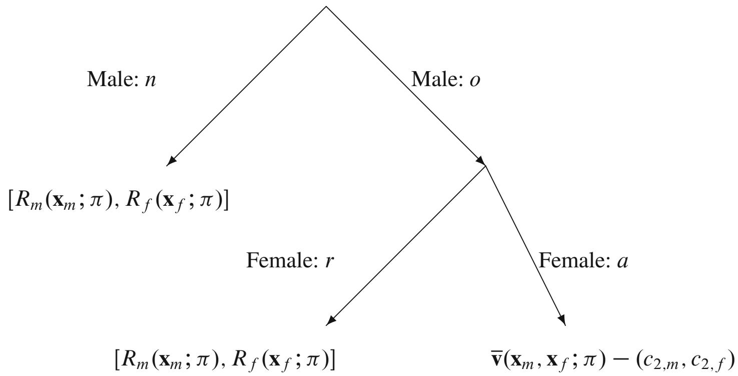


Fig. 1 Extensive form of the offer/acceptance game

are a – accept courtship and r – reject courtship. Assume that the population is following strategy profile π. As before, we may ignore previously incurred costs. Since the order in which actions are taken is important, we consider the extensive form of this game, given in Fig. 1.

Here, $\overline{\mathbf{v}}(\mathbf{x}_m, \mathbf{x}_f;\pi) = [\overline{v}_m(\mathbf{x}_m, x_{1,f};\pi), \overline{v}_f(x_{1,m}, \mathbf{x}_f;\pi)]$ denotes the expected value of the courting game given the strategy profile used, the measures of attractiveness of the pair, and the fact that courtship followed. This calculation and the calculation of expected rewards when one individual deviates from this profile are considered in Sect. 5.1.

When a male makes his decision he has no information on the female's character. However, the fact that a male wishes to initiate courtship may give the female information regarding his character. Suppose that the decision of a male of attractiveness $x_{1,m}$ on whether to initiate courtship does not depend on his character. The conditional distribution of the male's character given that courtship was initiated is simply the marginal distribution of male character. In this case, we say that the decision of the male is nonrevealing (with respect to his character). If the population has evolved to some equilibrium, the evolutionary process will have implicitly taught males which females will enter into courtship. Hence, we assume that males implicitly know the conditional distribution of the female's character given that courtship occurs. A strategy profile π is said to be nonrevealing if the decision of an individual on whether to court is only dependent on that individual's attractiveness.

The offer/acceptance game must be solved by recursion. The female only has to make a decision in the case when the male wishes to initiate courtship. She should court if and only if her expected reward from courting is at least as great as the expected reward from future search, i.e., $\overline{v}_f(x_{1,m}, \mathbf{x}_f;\pi) - c_{2,f} \geq R_f(\mathbf{x}_f;\pi)$.

Similarly, the male should initiate courtship if his expected reward from courtship is at least as great as his expected reward from future search. This is considered in more detail in Sect. 5.1.

At a Nash equilibrium profile π^* of Γ, each individual must adopt the appropriate Nash equilibrium strategy in any possible offer/acceptance subgame.

5 An Algorithm to Derive a Subgame Perfect Equilibrium for Quasisymmetric Games

In quasisymmetric games, the distribution of types, as well as the search and courting costs, are independent of sex. Let c_1 and c_2 be the search costs and the cost of courting, respectively. The reward of a type $[i, j]$ individual from mating with a type $[k, l]$ individual is defined to be $g(k, |l - j|)$. Let A and C be the sets of attractiveness and character levels, respectively. The set of types, T, is $T = A \times C$. An equilibrium is symmetric when the following conditions are satisfied:

1. For all $[i, j] \in T$, if a type $[i, j]$ male wishes to court a female of attractiveness k, then a type $[i, j]$ female accepts courtship from males of attractiveness k.
2. For all $[i, j] \in T$, if a type $[i, j]$ male wishes to mate with a type $[k, l]$ female in the courting game, then a type $[i, j]$ female wishes to mate with a type $[k, l]$ male.

At such an equilibrium the conditional distribution of the character of a female of attractiveness i courting a male of attractiveness j equals the conditional distribution of the character of a male of attractiveness i courting a female of attractiveness j. The value of Γ to an individual is independent of sex. Let $R_m(\mathbf{x}; \pi^*) = R_f(\mathbf{x}; \pi^*) = R(\mathbf{x}; \pi^*)$.

The algorithm starts by assuming individuals only court those of the same attractiveness and pair with those of the same type. Call this profile π_0. This is a Nash equilibrium profile when the costs are sufficiently small. This follows from the following argument: For sufficiently small costs, individuals of maximum attractiveness only wish to mate with their preferred type of partner (i.e., their own type). For the remaining players, the game reduces to one in which individuals of maximum attractiveness are not present. Individuals of the second highest level of attractiveness only wish to mate with their preferred partner in this reduced game (i.e., their own type) and so on.

We then iteratively improve the payoffs of individuals of maximum attractiveness by changing the patterns of courting and mating as follows:

1. Extend the set of acceptable mates of those of maximum attractiveness, $x_{1,\max}$. We start each stage with the type with the highest expected reward under the present strategy profile and finish with the type with the least reward. The set of acceptable mates is extended by either a) including any type of individual courted but not paired with who gives the maximum possible reward or b) decreasing the minimum level of attractiveness required for initiating courtship by 1, in this case individuals of the same character and any type which gives a reward at least as great as any type already in the set of acceptable mates should be included in the extension. This choice is made to maximize the expected reward from search. At this stage, it is assumed that individuals always wish to court those of higher attractiveness.

2. Step 1 is repeated until no improvements in the expected payoffs of individuals of attractiveness $x_{1,\max}$ are possible.
3. Suppose some but not all individuals of attractiveness $x_{1,\max}$ are willing to court prospective mates of attractiveness x_1. Steps 1 and 2 are repeated under the assumption that individuals of attractiveness x_1 do not accept the offer of courtship from an individual of attractiveness $x_{1,\max}$ if there is no chance that mating will follow under the present strategy profile.
4. Step 3 is repeated until the "optimal response" of players of lower attractiveness is the same at the end of Step 3 as it was at the beginning. It should be noted that if an individual of attractiveness $x_{1,\max}$ is willing to court and pair with an individual of type $[i, j]$, where $i < x_{1,\max}$, it is assumed that acceptance is mutual.

This procedure is then repeated for individuals of successively lower levels of attractiveness. It assumed that initially such individuals only pair with individuals of the same type and individuals of a higher level of attractiveness who are willing to pair with them. Since the algorithm attempts to maximize the expected reward of an individual given the behavior of individuals of greater attractiveness, the pattern of courting and mating that results from this procedure will be very similar to the pattern of courting and mating that results from a Nash equilibrium profile.

This procedure only defines a pattern of courting and mating (i.e., what behavior is observed). Once the algorithm has converged, we use policy iteration to check whether the proposed pattern of courting and mating corresponds to a Nash equilibrium of the required form and, if so, fully define the profile (i.e., define the behavior that occurs both on and away from the equilibrium path).

5.1 Example

Suppose there are 2 levels of attractiveness, $X_1 \in \{2, 3\}$, and three levels of character $X_2 \in \{0, 1, 2\}$. Each of the six possible types are equally likely. The search costs, c_1, are 0.3 and the courting costs, c_2, are 0.25. The reward obtained by a type $[i, j]$ individual from mating with a type $[k, l]$ individual is $k - |j - l|$.

Assuming individuals only court those of the same attractiveness and mate with those of the same type, the expected number of prospective mates seen and of courtships are six and three, respectively. The reward obtained from mating is the attractiveness of the partner. It follows that for $x = 2, 3$, $y = 0, 1, 2$:

$$R_{\pi_0}(x, y) = x - 0.3 \times 6 - 0.25 \times 3 = x - 2.55.$$

We now expand the set of types acceptable to an individual of attractiveness 3. It is of greater benefit for such an individual to pair with other individuals of attractiveness 3 with a neighboring level of character rather than with individuals of attractiveness 2 of the same character. Hence, we extend the set of types acceptable to individuals of attractiveness 3 to include those of neighboring character.

The expected payoffs of individuals of attractiveness 3 under such a profile π_1 are given by

$$R_{\pi_1}(3,0) = R_{\pi_1}(3,2) = \frac{1}{2}(2+3) - 0.3 \times 3 - 0.25 \times \frac{3}{2} = 1.225$$
$$R_{\pi_1}(3,1) = \frac{1}{3}(2+3+2) - 0.3 \times 2 - 0.25 \approx 1.4833.$$

Type $(3,0)$ individuals cannot gain by pairing with $(3,2)$ individuals as the reward obtained from such a mating is less than $R_{\pi_1}(3,0)$. We now check whether type $(3,0)$ individuals can gain by pairing with type $(2,0)$ individuals and type $(3,1)$ individuals can gain by pairing with $(2,1)$ individuals. Under such a profile, π_2, individuals of attractiveness 3 would be willing to court any prospective mate. We have

$$R_{\pi_2}(3,0) = R_{\pi_2}(3,2) = \frac{1}{3}(2+3+2) - 0.3 \times 2 - 0.25 \times 2 \approx 1.2333$$
$$R_{\pi_2}(3,1) = \frac{1}{4}(2+3+2+2) - 0.3 \times \frac{3}{2} - 0.25 \times \frac{3}{2} = 1.425.$$

Hence, type $(3,0)$ individuals should pair with type $(2,0)$ individuals [by symmetry type $(3,2)$ individuals should mate with those of type $(2,2)$], but type $(3,1)$ individuals should not pair with type $(2,1)$ individuals. Let π_3 be the corresponding profile.

Since the expected payoffs of individuals of attractiveness 3 are greater than can be obtained by pairing with any other type of individual, Step 1 is concluded. We now consider whether individuals of attractiveness 2 should be willing to court prospective mates of attractiveness 3. Since individuals of attractiveness 2 are only courted by individuals of type $(3,0)$ and $(3,2)$, who will not pair with a type $(2,1)$ individual, it follows that individuals of type $(2,1)$ should reject courtship with those of attractiveness 3. Denote this new profile by π_4. Under such a profile, when an individual of attractiveness 3 courts one of attractiveness 2, the conditional distribution of the character of either is as follows: 0 with probability $\frac{1}{2}$ and 2 with probability $\frac{1}{2}$.

Now consider the best response of individuals of attractiveness 3. Those of type $(3,1)$ should not accept individuals of type $(2,0)$ and $(2,2)$ and individuals of type $(3,0)$ should still court those of attractiveness 2, as it is more likely under π_4 than under π_3 that the prospective mate will be of the only acceptable type. We have $R_{\pi_4}(3,1) = R_{\pi_1}(3,1) \approx 1.4833$ and

$$R_{\pi_4}(3,0) = R_{\pi_4}(3,2) = \frac{1}{3}(2+3+2) - 0.3 \times 2 - 0.25 \times \frac{5}{3} \approx 1.3167.$$

The optimal response of individuals of attractiveness 2 to π_4 is the same as their optimal response to π_3. Hence, we now consider extending the sets of acceptable mates of individuals of attractiveness 2. Under π_4, individuals of type $(2,0)$ court those of type $(3,0)$, $(3,2)$ and of attractiveness 2 and pair with those of type $(2,0)$

or $(3,0)$. Individuals of type $(2,1)$ only court those of attractiveness 2 and pair with those of type $(2,1)$. It follows that $R_{\pi_4}(2,1) = R_{\pi_0}(2,1) = -0.55$ and

$$R_{\pi_4}(2,0) = R_{\pi_4}(2,2) = \frac{1}{2}(2+3) - 0.3 \times 3 - 0.25 \times \frac{5}{2} = 0.975.$$

Since the expected payoffs of all these types are less than those gained from pairing with an individual of attractiveness 2 with a neighboring level of character, the sets of acceptable mates should be extended. Denote the resulting profile by π_5. We have

$$R_{\pi_5}(2,0) = R_{\pi_5}(2,2) = \frac{1}{3}(2+3+1) - 0.3 \times 2 - 0.25 \times \frac{5}{3} = 0.9833$$
$$R_{\pi_5}(2,1) = \frac{1}{3}(1+2+1) - 0.3 \times 2 - 0.25 \approx 0.4833.$$

Extending the set of acceptable mates will not increase the expected reward of any of these types. Hence, π_5 is our candidate for a courting and mating profile which corresponds to a Nash equilibrium.

We now use policy iteration to check whether this courting and mating profile corresponds to a strategy profile satisfying the appropriate criteria. First, consider the courting game. Each individual should accept a prospective mate when the reward obtained from such a pairing is at least as great as the expected reward from future search. It follows that

1. Type $(3,0)$ individuals should pair with individuals of type $(3,0)$, $(3,1)$ or $(2,0)$.
2. Type $(3,1)$ individuals should pair with individuals of type $(3,0)$, $(3,1)$, $(3,2)$ or $(2,1)$.
3. Type $(2,0)$ individuals should pair with individuals of type $(2,0)$, $(2,1)$, $(3,0)$, $(3,1)$ or $(3,2)$. Acceptance is not mutual in the case of the two final types.
4. Type $(2,1)$ individuals should pair with any prospective mate, but acceptance is not mutual when the prospective mate is of type $(3,0)$ or $(3,2)$.

By symmetry if an individual of type $(i,0)$ accepts a prospective partner of type (j,k) at any stage, then an individual of type $(i,2)$ accepts a prospective partner of type $(j,2-k)$.

We now consider the offer/acceptance game. First, consider the response of a female to an offer of courtship from a male of attractiveness i, where $i \in \{2,3\}$. If under a strategy profile π a female of attractiveness j is offered courtship by some males of attractiveness i, then it is assumed that the distribution of the male's character comes from the conditional distribution of character given the offer of courtship. Suppose no male of attractiveness i is willing to court a female of attractiveness j under π. It is assumed that the distribution of the male's character is simply the marginal distribution of the character of males (i.e., the probability of an individual making a "mistake" is $o(1)$ and independent of his/her type and the prospective partner's type). Analogous assumptions are made in the calculation of whether a male should be willing to initiate courtship or not.

Suppose a type $(3,0)$ female is offered courtship by a male of attractiveness 3. Since all males of attractiveness 3 are willing to court such a female, the expected reward from entering into courtship is

$$\overline{v}_f([3,0],3) = \frac{1}{3}[3 + 2 + R_{\pi_5}(3,0)] - 0.25 > R_{\pi_5}(3,0).$$

Hence, a type $(3,0)$ female should accept courtship.

Using a similar procedure, it can be shown that females of types $(3,0)$ and $(3,2)$ should accept courtship with any male, but type $(3,1)$ females should only accept courtship with males of attractiveness 3. Type $(2,0)$ and type $(2,2)$ females should accept courtship with any male. Since no male of attractiveness 3 who wishes to court a female of attractiveness 2 will pair with a type $(2,1)$ female, such females should reject courtship with a male of attractiveness 3.

Now consider whether a type $(3,0)$ male should offer courtship to a female of attractiveness 2. Type $(2,1)$ females will reject such an offer, no courting costs are incurred and the future expected reward of the male is $R_{\pi_5}(3,0)$. Females of type $(2,0)$ and $(2,2)$ will accept such an offer and the future expected rewards of the male in these cases are $2 - c_2 = 1.75$ and $R_{\pi_5}(3,0) - c_2$, respectively. It follows that the expected reward obtained from offering courtship in this case is $\frac{1}{3}\left[1.75 + 2R_{\pi_5}(3,0) - 0.25\right] \approx 1.3778 > R_{\pi_5}(3,0)$. Hence, a type $(3,0)$ male should offer courtship to a female of attractiveness 2. Arguing similarly, males of types $(3,0)$ and $(3,2)$ should offer courtship to any female, but type $(3,1)$ males should only try to court females of attractiveness 3. In a similar way, it can be shown that males of type $(2,0)$ or $(2,2)$ should offer courtship to any female and type $(2,1)$ males should only offer courtship to females of attractiveness 2.

Thus, the behavior observed under π_5 corresponds to a Nash equilibrium strategy profile satisfying Conditions 1–3. The description of this strategy profile is as follows [by symmetry if an individual of type $(i,0)$ accepts a prospective partner of type (j,k) at any stage, then an individual of type $(i,2)$ accepts a prospective partner of type $(j, 2-k)$]:

Type (3,0) individuals In the offer/acceptance game, court those of any attractiveness. In the courting game, pair with those of type (3,0), (3,1), and (2,0).

Type (3,1) individuals In the offer/acceptance game, court those of attractiveness 3. In the courting game, pair with those of type (3,0), (3,1), (3,2), and (2,1). Note, however, that (3,1) individuals do not court those of attractiveness 2.

Type (2,0) individuals In the offer/acceptance game, court those of any attractiveness. In the courting game, pair with those of any type except (2,2). In the cases of those of type (3,1) or (3,2), acceptance is not mutual.

Type (2,1) individuals In the offer/acceptance game, court only those of attractiveness 2. In the courting subgame, pair with those of any type.

One interesting aspect of this strategy profile is that type $(3,1)$ individuals would pair with type $(2,1)$ individuals in the courting game, but the marginal gain from such a pairing is not large enough to justify the costs of a type $(3,1)$ male courting a female of attractiveness 2.

The following argument indicates that this is the unique Nash equilibrium profile satisfying Conditions 1–3: If an individual of attractiveness 3 only mates with prospective mates of the same type, then his/her expected payoff is less than 2. Since all other prospective partners give a reward of 2 or less, the expected reward of an individual of attractiveness 3 must be less than 2 at any equilibrium. Hence, in the courting game such individuals must accept individuals of attractiveness 3 with neighboring character. Given this, type $(3, 1)$ individuals should not court those of attractiveness 2. Individuals of types $(3, 0)$ and $(3, 2)$ should court those of attractiveness 2 and mate with those of the same character (as derived above). The only assumption that was made in the derivation of the optimal response of type $(2, 0)$ and type $(2, 2)$ individuals was that they should pair with those of type $(3, 0)$ and $(3, 2)$, respectively. It is reasonably simple to show by considering all the possible courting and mating patterns that this is the case.

It should be noted that for a given level of attractiveness, individuals of central character may have a lower expected payoff than those of more extreme character. Alpern and Reyniers [1] found that when mate choice is based purely on homotypic preferences according to a trait with a symmetric distribution around x_0 whose density function is nonincreasing for values above x_0, then those of central character have a higher expected payoff than those of more extreme character. This apparent discrepancy may be explained by the fact that individuals of maximum attractiveness and extreme character are less choosy than those of maximum attractiveness and central character. This may lead to relatively better mating opportunities for individuals of less than maximum attractiveness and extreme character.

6 Results for a More Complex Problem

The algorithm was used to solve a problem in which there were 10 levels of attractiveness ($x_1 \in \{11, 12, \dots, 20\}$) and character ($x_2 \in \{1, 2, \dots, 10\}$). The search costs and courting costs are $c_1 = 0.1$ and $c_2 = 0.1$, respectively. The reward gained by a type (i, j) individual mating with a type (k, l) individual is $k - |j - l|$. It was assumed that the equilibrium profile is symmetric with respect to character (i.e., by taking $11 - j$ rather than j to be the character of an individual, the courting and mating pattern remain the same). The mating pattern proposed by the algorithm is described in Table 1. Analysis of the courting and offer/acceptance subgames confirmed that the proposed pattern corresponded to a Nash equilibrium profile.

It should be noted that there are other Nash equilibria which satisfy Conditions 1–3. At the equilibrium presented above, no individual of attractiveness 20 is willing to court prospective mates of attractiveness 17. It can be checked that when all individuals of attractiveness 17 are willing to court prospective mates of attractiveness 20, then it would not pay a type $(20, 1)$ or $(20, 10)$ individual to enter into such a courtship (these are the only types that might benefit). Suppose type $(17, 1)$ and $(17, 10)$ individuals are willing to court prospective mates of attractiveness 20, while other individuals of attractiveness 17 are unwilling. It can be shown that individuals

Table 1 A Nash equilibrium for a more complex problem

Type	Pairs with	Reward
$(20, j), j = 4, 5$	$(20, j \pm 2), (20, j \pm 1), (20, j), (19, j \pm 1), (19, j)$	17.1250
$(20, j), j = 3$	$(20, j \pm 2), (20, j \pm 1), (20, j), (19, j \pm 1), (19, j), (18, 3)$	17.1556
$(20, j), j = 2$	$(20, 4), (20, j \pm 1), (20, j), (19, j \pm 1), (19, j), (18, 2)$	17.0500
$(20, 1)$	$(20, 1), (20, 2), (20, 3), (19, 1), (19, 2), (18, 1), (18, 2)$	16.6750
$(19, j), j = 4, 5$	$(20, j \pm 1), (20, j), (19, j \pm 2), (19, j \pm 1), (19, j), (18, j \pm 1), (18, j)$	16.9091
$(19, j), j = 3$	$(20, 1), (20, j \pm 1), (20, j), (19, j \pm 2),$ $(19, j \pm 1), (19, j), (18, j \pm 1), (18, j)$	17.0000
$(19, j), j = 2$	$(20, j \pm 1), (20, j), (19, j \pm 1), (19, j), (19, 4)(18, j \pm 1), (18, j)$	16.9000
$(19, 1)$	$(20, 1), (20, 2), (19, 1), (19, 2), (19, 3), (18, 1), (18, 2)$	16.4286
$(18, j), j = 4, 5$	$(19, j \pm 1), (19, j), (18, j \pm 1), (18, j), (18, j + 2)(17, j \pm 1), (17, j)$	15.9000
$(18, j), j = 3$	$(20, j), (19, j \pm 1), (19, j), (18, j \pm 1), (18, j)$	16.3429
$(18, j), j = 2$	$(20, 1), (20, 2), (19, j \pm 1), (19, j), (18, j \pm 1), (18, j)$	16.6750
$(18, 1)$	$(20, 1), (19, 1), (19, 2), (18, 1), (18, 2), (17, 1), (17, 2)$	15.9143
$(17, j), j = 5$	$(18, j \pm 1), (18, j), (17, j \pm 2), (17, j \pm 1), (17, j), (16, j \pm 1), (16, j)$	14.9455
$(17, j), j = 4$	$(18, 4), (18, 5), (17, j \pm 2), (17, j \pm 1), (17, j), (16, j \pm 1), (16, j)$	14.7400
$(17, j), j = 3$	$(18, 4), (17, j \pm 2), (17, j \pm 1), (17, j), (16, j \pm 1), (16, j)$	14.3778
$(17, j), j = 2$	$(18, 1), (17, j \pm 1), (17, j), (17, 4), (16, j \pm 1)$	14.3000
$(17, 1)$	$(18, 1), (17, 1), (17, 2), (17, 3), (16, 1), (16, 2)$	14.0667
$(i, j), 12 \leq i \leq 16,$ $j = 3, 4, 5$	$(i + 1, j \pm 1), (i + 1, j), (i, j \pm 2), (i, j \pm 1),$ $(i, j), (i - 1, j \pm 1), (i - 1, j)$	$i - \frac{23}{11}$
$(i, j), j = 2$ $12 \leq i \leq 16,$	$(i + 1, j \pm 1), (i + 1, j), (i, j \pm 1), (i, j),$ $(i, 4), (i - 1, j \pm 1), (i - 1, j)$	$i - \frac{21}{10}$
$(i, 1), 12 \leq i \leq 16$	$(i + 1, 1), (i + 1, 2), (i, 1), (i, 2), (i, 3), (i - 1, 1), (i - 1, 2)$	$i - \frac{18}{7}$
$(11, j), j = 3, 4, 5$	$(12, j \pm 1), (12, j), (11, j \pm 2), (11, j \pm 1), (11, j)$	8.8750
$(11, j), j = 2$	$(12, j \pm 1), (12, j), (11, j \pm 1), (11, j), (11, 4)$	8.8571
$(11, 1)$	$(12, 1), (12, 2), (11, 1), (11, 2), (11, 3)$	8.200

of type $(20, 1)$ and type $(20, 10)$ should be willing to court such prospective mates and then mate if they have the same character. Other equilibria were also found.

It seems that there will be a very large set of Nash equilibria satisfying Conditions 1–3 and the problem of deriving all of them would seem to be very difficult. However, it seems that all these equilibria would be qualitatively similar in terms of the choosiness of the individuals and the level of association between the attractiveness and characters of partners.

7 Conclusion

This paper has presented a paper of mutual mate choice with both common and homotypic preferences. A method for finding a Nash equilibrium satisfying various criteria based on the concept of subgame perfection was described.

The use of this combination of preferences would seem to be logical in relation to human mate choice. Although there is no perfect correlation in individuals' assessment of the attractiveness of members of the opposite sex, there is normally a very high level of agreement, particularly among men (see Townsend and Wasserman [36]). Females tend to place more emphasis on the character of a male. It has been argued that approaches which adapt ideas from matching problems, where individuals have their own individual rankings of members of the opposite sex , should be adopted to mate choice problems (see Gale and Shapley [14] and Bergstrom and Real [6]). The approach used here seems to be a good compromise between the approaches used in matching and the assumptions of common preferences. These "mixed" preferences are reasonably tractable within the framework of searching for a mate in a large population and allow a general enough framework to model the preferences of individuals reasonably well (although modeling character as a one-dimensional variable is rather simplistic). Using a larger number of types, we can approximate continuous distributions of attractiveness and character.

It would be interesting to consider different ways in which information is gained during the search process. For example, in human mate choice some information about the character of a prospective mate may be readily available via the way an individual dresses or his/her musical tastes. An improved model would allow some information to be gained on both the attractiveness and the character of a prospective partner at each stage of the process.

In terms of the evolution of such strategies, the model assumes that the basic framework is given, i.e., mate choice is mutual and the various search and courtship costs are given. However, this model cannot explain why such a system has evolved, only the evolution of decisions within this framework. In human mate choice, individuals may lower their search costs by joining some internet or social group. Such methods can also lead to biasing the operational sex ratio or conditional distribution of the character of a prospective mate in a searcher's favor, e.g., a male interested in singing could join a choir.

As it seems there may be a large number of Nash equilibria satisfying the required conditions, it would be of interest to carry out simulations of how mate choice strategies evolve using replicator dynamics. Also, it would be useful to investigate how the payoff functions, together with the relative costs of searching and courting, affect the importance of attractiveness and character in the decision process. Using attractiveness as an initial filter in the decision process will lead to attractiveness becoming relatively more important than character, especially if the costs of courting are relatively high. It is intended that a future paper will investigate these issues in more detail.

Acknowledgements I would like to thank Prof. Steve Alpern for conversations that led me to develop this model and the anonymous reviewers for their suggestions.

References

1. Alpern, S., Reyniers, D.: Strategic mating with homotypic preferences. J. Theor. Biol. **198**, 71–88 (1999)
2. Alpern, S., Reyniers, D.: Strategic mating with common preferences. J. Theor. Biol. **237**, 337–354 (2005)
3. Alpern, S., Katrantzi, I.: Equilibria of two-sided matching games with common preferences. Eur. J. Oper. Res. **196**, 1214–1222 (2009)

4. Backwell, P.R.Y., Passmore, N.I.: Time constraints and multiple choice criteria in the sampling behavior and mate choice of the fiddler crab, *Uca annulipes*. Behav. Ecol. Sociobiol. **38**, 407–416 (1996)
5. Bakker, T.C.M., Milinski, M.: Sequential mate choice and the previous male effect in sticklebacks. Behav. Ecol. Sociobiol. **29**, 205–210 (1991)
6. Bergstrom, C.T., Real, L.A.: Towards a theory of mutual mate choice: lessons from two-sided matching. Evol. Ecol. Res. **2**, 493–508 (2000)
7. Burdett, K., Coles, M.G.: Long-term partnership formation: marriage and employment. Econ. J. **109**, 307–334 (1999)
8. Collins, E.J., McNamara, J.M.: The job-search problem with competition: an evolutionarily stable strategy. Adv. Appl. Prob. **25**, 314–333 (1993)
9. Collins, E.J., McNamara, J.M., Ramsey, D.M.: Learning rules for optimal selection in a varying environment: mate choice revisited. Behav. Ecol. **17**, 799–809 (2006)
10. Cooke, F., Davies, J.C.: Assortative mating, mate choice and repreductive fitness in Snow Geese. In: Bateson P. (ed.) Mate Choice, pp. 279–295. Cambridge University Press, Cambridge (1983)
11. Fawcett, T.W., Johnstone, R.A.: Mate choice in the face of costly competition. Behav. Ecol. **14**, 771–779 (2003)
12. Fawcett, T.W., Johnstone, R.A.: Optimal assessment of multiple cues. Proc. R. Soc. Lond., **270**, 1637–1643 (2003)
13. Fawcett, T.W., Bleay, C.: Previous experiences shape adaptive mate preferences. Behav. Ecol. **20**, 68–78 (2009)
14. Gale, D., Shapley, L.S.: College admissions and the stability of marriage. Am. Math. Monthly **69**, 9–15 (1962)
15. Härdling, R., Kokko, H.: The evolution of prudent choice. Evol. Ecol. Res. **7**, 697–715 (2005)
16. Iwasa, Y., Pomiankowski, A.: The evolution of mate preferences for multiple sexual ornaments. Evol. **48**, 853–867 (1994)
17. Janetos A. C.: Strategies of female mate choice: a theoretical analysis. Behav. Ecol. Sociobiol. **7**, 107–112 (1980)
18. Johnstone, R.A.: The tactics of mutual mate choice and competitive search. Behav. Ecol. Sociobiol. **40**, 51–59 (1997)
19. Kalick S. M. and Hamilton T. E.: The matching hypothesis reexamined. J. Personality Soc. Psychol. **51**, 673–682 (1986)
20. Kodric-Brown, A.: Female choice of multiple male criteria in guppies: interacting effects of dominance, colouration and courtship. Behav. Ecol. Sociobiol. **32**, 415–420 (1993)
21. Mazalov, V., Perrin, N., Dombrovsky, Y.: Adaptive search and information updating in sequential mate choice. Am. Nat. **148**, 123–137 (1996)
22. McNamara, J.M., Collins, E.J.: The job search problem as an employer-candidate game. J. Appl. Prob. **28**, 815–827 (1990)
23. Monaghan, P., Metcalfe, N., Houston, D.C.: Male finches selectively pair with fecund females. Proc. R. Soc. Lond. B **263**, 1183–1186 (1996)
24. Parker, G.A.: Mate quality and mating decisions. In: Bateson P. (ed.) Mate Choice, pp. 227–256. Cambridge University Press, Cambridge (1983)
25. Ramsey, D.M.: A large population job search game with discrete time. Eur. J. Oper. Res. **188**, 586–602 (2008)
26. Real, L.A.: Search theory and mate choice. I. Models of single-sex discriminination. Am. Nat. **136**, 376–404 (1990)
27. Real, L.A.: Search theory and mate choice. II. Mutual interaction, assortative mating, and equilibrium variation in male and female fitness. Am. Nat. **138**, 901–917 (1991)
28. Robertson, J.G.M.: Female choice increases fertilization success in the Australian frog, *Uperiola laevigata*. Anim. Behav. **39**, 639–645 (1990)
29. Ryan, M.J., Perrill, S.A., Wilczynski, W.: Auditory tuning and call frequency predict population based mating preferences in the cricket frog, *Acris crepitans*. Am. Nat. **139**, 1370–1383 (1992)

30. Seymour, R.M., Sozou, P.D.: Duration of courtship as a costly signal. J. Theor. Biol. **256**, 1–13 (2008)
31. Shimer, R., Smith, L.: Assortative matching and search. Econometrica **68**, 343–370 (2000)
32. Simão, J., Todd, P.M.: Modeling mate choice in monogamous mating systems with courtship. J. Adaptive Behav. **10**, 113–136 (2002)
33. Simão, J., Todd, P.M.: Emergent patterns of mate choice in human populations. Art. Life **9**, 403–417 (2003)
34. Smith, L.: The marriage model with search frictions. J. Pol. Econ. **114**, 1124–1144 (2006)
35. Sozou, P.D.: Costly but worthless gifts facilitate courtship. Proc. R. Soc. Lond. B **272**, 1877–1884 (2005)
36. Townsend, J.M., Wasserman, T.: The perception of sexual attractiveness: sex differences in variability. Arch. Sexual Behav. **26**, 243–268 (1997)
37. Vahed, K.: The function of nuptial feeding in insects: a review of empirical studies. Biol. Rev. **73**, 43–78 (1998)

Part IV
Cooperative Games

1. *Ekaterina Shevkoplyas*, The Shapley Value in Cooperative Differential Games with Random Duration
2. *David W.K. Yeung*, Dynamically Consistent Cooperative Solutions in Differential Games with Asynchronous Players' Horizons

The Shapley Value in Cooperative Differential Games with Random Duration

Ekaterina Shevkoplyas

Abstract The class of cooperative differential games with random duration is studied in this chapter. The problem of the Shapley Value calculation is examined. As a result, the Hamilton–Jacobi–Bellman equation for the problem with random duration is derived. The Shapley Value calculation method, which uses the obtained equation, is represented by an algorithm. An application of the theoretical results is illustrated with a model of nonrenewable resource extraction by n firms or countries.

1 Introduction

It is well known that differential games allow to model conflicts between players in continuous time. In this chapter, we have focused on the duration of these games. In differential games, it is common to consider games with prescribed duration (finite time horizon) or games with infinite time horizon. Obviously, many processes occurring in the real world end at a random moment in time. In particular, during the recent economic crisis, many financial contracts and agreements have been terminated prior to the completion of long-term projects. Hence, the game with random duration, which simulates conflict-controlled process between participants, reflects more adequately real-life processes.

In addition, in the modern civilized world many problems can be solved not in the face of fierce competition, but in conditions of cooperation and collaboration. That is why we consider the cooperative form of the game. In this chapter, we explore cooperative differential games with random duration. The class of cooperative differential games with random duration was introduced for the first time in the chapter [4]. Section 2 contains a formulation of such games.

E. Shevkoplyas (✉)
St. Petersburg State University, Faculty of Applied Mathematics, University pr., 35, Petrodvorets, 198 904 St. Petersburg, Russia
e-mail: ekaterina.shevkoplyas@gmail.com

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_18,

In cooperative differential games, players solve the optimal control problem of the total payoff maximization under a set of constraints. In particular, a differential equation which describes the evolution of the state of the game can be considered as such a constraint. One of the basic techniques for solving the optimal control problem is the Hamilton–Jacobi–Bellman equation [1].

However, there is a nonstandard dynamic programming problem for games with random duration because of the objective functional form (double integral). This is the reason why the usual Hamilton–Jacobi–Bellman equation is no longer appropriate for the solution of cooperative differential games with random duration. In Sect. 3, the Hamilton–Jacobi–Bellman equation for general problem with random duration, with arbitrary probability density function $f(t) = F'(t)$, is derived.

Section 4 presents a discussion of two methods for constructing the characteristic function, namely, the classical approach suggested by J. von Neumann and O. Morgenshtern, and the nonstandard Nash equilibrium approach presented by L. Petrosjan and G. Zaccour in the paper [3]. The main idea of the second approach is the assumption that if a subset of players forms a coalition S, then the left-out players stick to their feedback Nash strategies. Such a method is used for problems with environmental and other contexts where it can be assumed that players from the remaining set $N \setminus S$ behave nonaggressively. Notice that in general the characteristic function constructed by the method suggested in [3] is not superadditive.

Since we are considering the cooperative form of the game, we have to choose an optimality principle. One of the most popular optimality principles in cooperative theory is the Shapley Value. In Sect. 5, we introduce an algorithm for the computation of the Shapley Value based on Petrosjan and Zaccour's approach. The Hamilton–Jacobi–Bellman equation derived in the problem with random duration is used in this algorithm.

In Sect. 6, an application of our theoretical results is presented. We investigate one simple model of nonrenewable resource procurement made by n firms or countries [1] under the condition of random game duration. To construct the Shapley value of the game, the algorithm constructed in Sect. 5 is used.

2 Game Formulation

There are two common concepts in differential game theory about game duration. In the first approach, the game is considered on the fixed time interval $[t_0, T]$. The second approach is usually used for economic applications: if the final time instant of the game is very large, then it may be supposed $T = \infty$ (see for example [3]). A differential game with infinite time horizon has an infinite duration and never ends. To mitigate this fact, the player's utility function is usually discounted in time by an exponential function (although, of course, there are other ways), so that the chances for a player to earn something in the far future tend to zero. But it seems more realistic to consider conflict-controlled processes on the time interval $[t_0, T]$ with an assumption that the final time T is a random variable with a known distribution function.

Consider a n-person differential game $\Gamma(x_0)$ from the initial state x_0 with random duration $T - t_0$ [4]. Here, the random variable T with distribution function $F(t), t \in [t_0, \infty)$,

$$\int_{t_0}^{\infty} \mathrm{d}F(t) = 1,$$

is the time instant when the game $\Gamma(x_0)$ ends. The game starts at the moment t_0 from the position x_0.

Let the motion equations have the form

$$\dot{x} = g(x, u_1, \dots, u_n),\ x \in R^n, u_i \in U \subseteq \operatorname{comp} R^l,\ x(t_0) = x_0. \qquad (1)$$

The "instantaneous" payoff at the moment τ, $\tau \in [t_0, \infty)$ is defined as $h_i(x(\tau))$. Then the expected integral payoff of the player $i, i = 1, \dots, n$ is evaluated by the formula

$$K_i(x_0, u_1, \dots, u_n) = \int_{t_0}^{\infty} \int_{t_0}^{t} h_i(x(\tau))\mathrm{d}\tau \mathrm{d}F(t), \quad h_i \geq 0, \quad i = 1, \dots, n. \qquad (2)$$

Let $x^*(t)$ and $u^*(t) = (u_1^*(t), \dots, u_n^*(t))$ be the optimal trajectory and the corresponding n-tuple of optimal controls maximizing the joint expected payoff of players (we suppose that maximum is attained)

$$\begin{aligned}\max_u \sum_{i=1}^{n} K_i(x_0, u_1, \dots, u_n) &= \sum_{i=1}^{n} K_i(x_0, u_1^*, \dots, u_n^*) \\ &= \sum_{i=1}^{n} \int_{t_0}^{\infty} \int_{t_0}^{t} h_i(x^*(\tau))\mathrm{d}\tau \mathrm{d}F(t) = V(I, x_0).\end{aligned}$$

For simplicity, we shall suppose further that the optimal trajectory is unique.

For the set of subgames $\Gamma(x^*(\vartheta))$ occurring along an optimal trajectory $x^*(\vartheta)$, one can similarly define the expected total integral payoff in cooperative game $\bar{\Gamma}(x^*(\vartheta))$:

$$V(I, x^*(\vartheta)) = \sum_{i=1}^{n} \int_{\vartheta}^{\infty} \int_{\vartheta}^{t} h_i(x^*(\tau))\mathrm{d}\tau \, \mathrm{d}F_{\vartheta}(t). \qquad (3)$$

Here $F_{\vartheta}(t)$ is a conditional distribution function. In this paper we consider only stationary processes, so we have the following expression for $F_{\vartheta}(t)$:

$$F_{\vartheta}(t) = \frac{F(t) - F(\vartheta)}{1 - F(\vartheta)}, \quad t \in [\vartheta, \infty).$$

It is clear that $(1 - F(\vartheta))$ is the probability to start $\Gamma(x^*(\vartheta))$.

In the same way, we get the expression for conditional distribution in subgames $\Gamma(x^*(\vartheta + \Delta))$:

$$F_{\vartheta+\Delta}(t) = \frac{F_\vartheta(t) - F_\vartheta(\vartheta + \Delta)}{1 - F_\vartheta(\vartheta + \Delta)} = \frac{F(t) - F(\vartheta + \Delta)}{1 - F(\vartheta + \Delta)}.$$

Further, we assume the existence of a density function $f(t) = F'(t)$. As above, the formula for the conditional density function is:

$$f_\vartheta(t) = \frac{f(t)}{1 - F(\vartheta)},\ t \in [\vartheta, \infty);\ f_{\vartheta+\Delta}(t) = \frac{f(t)}{1 - F(\vartheta + \Delta)}, \qquad t \in [\vartheta + \Delta, \infty). \tag{4}$$

From (4), we obtain

$$f_\vartheta(t) = \frac{1 - F(\vartheta + \Delta)}{1 - F(\vartheta)} f_{\vartheta+\Delta}(t). \tag{5}$$

3 The Hamilton–Jacobi–Bellman Equation

The Hamilton–Jacobi–Bellman equation lies at the heart of the dynamic programming approach to optimal control problems. Let us remark that the functional (3) does not have the standard form of the dynamic programming problem. Thus, we first need to derive the Hamilton–Jacobi–Bellman equation that is appropriate for the problem with random duration.

Denote $H(x(t)) = \sum\limits_{i=1}^{n} h_i(x(t))$. In the general case, we consider $H(x, u)$.

Let $P(x, \vartheta)$ to be an optimization problem

$$\max_u \left[\int_\vartheta^\infty \int_\vartheta^t H(x(\tau), u(\tau)) f_\vartheta(t) \mathrm{d}\tau \mathrm{d}t \right], \tag{6}$$

subject to (1), $x(\vartheta) = x$.

Let $W(x, \vartheta)$ be the optimal value (or Bellman function) of the objective functional of problem $P(x, \vartheta)$ in (6):

$$W(x, \vartheta) = \max_u \left[\int_\vartheta^\infty \int_\vartheta^t H(x(\tau), u(\tau)) f_\vartheta(t) \mathrm{d}\tau \mathrm{d}t \right]. \tag{7}$$

We can see that the maximal total payoff in $\Gamma(x_0)$ is

$$V(I, x_0) = W(x_0, t_0).$$

In control theory, one usually makes the assumptions that functions g and H are sufficiently smooth and satisfy certain boundedness conditions to ensure that solutions to (7) are uniquely defined and that the integral in (7) makes sense. Here, we do not impose any strong restriction because we cannot easily assume any restrictive properties for the functions g and H, but we make an assumption that the objective functional W is well defined.

Obviously, if one behaves optimally from $t + \Delta$ onward, the total expected payoff is given by formula

$$W(x, \vartheta + \Delta) = \max_u \left[\int_{\vartheta+\Delta}^{\infty} \int_{\vartheta+\Delta}^{t} H(x(\tau), u(\tau)) f_{\vartheta+\Delta}(t) \mathrm{d}\tau \mathrm{d}t \right]. \tag{8}$$

Using (7), (5), (4), and (8), we get:

$$\begin{aligned} W(x, \vartheta) &= \max_u \int_{\vartheta}^{\vartheta+\Delta} \int_{\vartheta}^{t} H(x(\tau), u(\tau)) f_{\vartheta}(t) \mathrm{d}\tau \mathrm{d}t \\ &\quad + \int_{\vartheta+\Delta}^{\infty} \int_{\vartheta}^{\vartheta+\Delta} H(x(\tau), u(\tau)) f_{\vartheta}(t) \mathrm{d}t \mathrm{d}\tau \\ &\quad + \frac{1 - F(\vartheta + \Delta)}{1 - F(\vartheta)} \int_{\vartheta+\Delta}^{\infty} \int_{\vartheta+\Delta}^{t} H(x(\tau), u(\tau)) f_{\vartheta+\Delta}(t) \mathrm{d}\tau \mathrm{d}t \\ &= \max \Bigg(\int_{\vartheta}^{\vartheta+\Delta} \int_{\vartheta}^{t} H(x(\tau), u(\tau)) f_{\vartheta}(t) \mathrm{d}\tau \mathrm{d}t \\ &\qquad + \frac{1 - F(\vartheta + \Delta)}{1 - F(\vartheta)} \int_{\vartheta}^{\vartheta+\Delta} H(x(\tau), u(\tau)) \mathrm{d}\tau \\ &\qquad + \frac{1 - F(\vartheta + \Delta)}{1 - F(\vartheta)} W(x(\vartheta + \Delta), \vartheta + \Delta) \Bigg). \end{aligned} \tag{9}$$

Notice that

$$\frac{1 - F(\vartheta + \Delta)}{1 - F(\vartheta)} = 1 + \frac{F(\vartheta) - F(\vartheta + \Delta)}{1 - F(\vartheta)}.$$

Now subtract $W(x, \vartheta)$ from both sides of (9) and divide the resulting equation by Δ. This yields

$$\begin{aligned} 0 = \max_u \Bigg(&\frac{1}{\Delta} \int_{\vartheta}^{\vartheta+\Delta} \int_{\vartheta}^{t} H(x(\tau), u(\tau)) f_{\vartheta}(t) \mathrm{d}\tau \mathrm{d}t + \frac{1}{\Delta} \int_{\vartheta}^{\vartheta+\Delta} H(x(\tau), u(\tau)) \mathrm{d}\tau \\ &+ \frac{F(\vartheta) - F(\vartheta + \Delta)}{1 - F(\vartheta)} \frac{1}{\Delta} \int_{\vartheta}^{\vartheta+\Delta} H(x(\tau), u(\tau)) \mathrm{d}\tau \\ &+ \frac{W(x(\vartheta + \Delta), \vartheta + \Delta) - W(x(\vartheta), \vartheta)}{\Delta} \\ &+ \frac{F(\vartheta) - F(\vartheta + \Delta)}{1 - F(\vartheta)} \frac{1}{\Delta} W(x(\vartheta + \Delta), \vartheta + \Delta) \Bigg). \end{aligned} \tag{10}$$

Let $\Delta \to 0$. From the mean value theorem, we know that

$$\lim_{\Delta \to 0} \frac{1}{\Delta} \int_{\vartheta}^{\vartheta+\Delta} \int_{\vartheta}^{t} H(x(\tau), u(\tau)) f_{\vartheta}(t) \mathrm{d}\tau \mathrm{d}t = 0. \tag{11}$$

Moreover, we have

$$\begin{aligned}
&\lim_{\Delta \to 0} \quad \frac{1}{\Delta} \int_{\vartheta}^{\vartheta+\Delta} H(x(\tau), u(\tau)) \mathrm{d}\tau = H(x(\vartheta), u(\vartheta)); \\
&\lim_{\Delta \to 0} \frac{1}{\Delta} \frac{F(\vartheta) - F(\vartheta + \Delta)}{1 - F(\vartheta)} \int_{\vartheta}^{\vartheta+\Delta} H(x(\tau), u(\tau)) \mathrm{d}\tau = 0; \\
&\lim_{\Delta \to 0} \frac{1}{\Delta} \frac{F(\vartheta) - F(\vartheta + \Delta)}{1 - F(\vartheta)} = -\frac{F'(\vartheta)}{1 - F(\vartheta)} = -\frac{f(\vartheta)}{1 - F(\vartheta)}.
\end{aligned} \tag{12}$$

Combining (10), (11), and (12), we obtain

$$\begin{aligned}
0 = \max_{u} \Bigg(& H(x(\vartheta), u(\vartheta)) + \frac{\mathrm{d}}{\mathrm{d}\vartheta} W(x, \vartheta) \\
& + \lim_{\Delta \to 0} \left[\frac{1}{\Delta} \frac{F(\vartheta) - F(\vartheta + \Delta)}{1 - F(\vartheta)} W(x(\vartheta + \Delta), \vartheta + \Delta) \right] \Bigg).
\end{aligned}$$

So we have the Hamilton–Jacobi–Bellman equation:

$$\frac{f(\vartheta)}{1 - F(\vartheta)} W(x, \vartheta) = \frac{\partial W(x, \vartheta)}{\partial \vartheta} + \max_{u} \left[H(x(\vartheta), u(\vartheta)) + \frac{\partial W(x, \vartheta)}{\partial x} g(x, u) \right]. \tag{13}$$

Notice that the term $\frac{f(\vartheta)}{1-F(\vartheta)}$ in the left-hand side of (13) is a well-known function in mathematical reliability theory [2], called the Hazard function (or failure rate) with typical notation $\lambda(\vartheta)$ or $h(\vartheta)$. In the sequel, we use the notation $\lambda(t)$. In mathematical reliability theory, the Hazard function is a conditional density function for the random failure time of a technical system. Thus, we can see some parallels between the random time instant when the game ends and the failure time for system of elements.

Finally, we get the following form for the new Hamilton–Jacobi–Bellman equation:

$$\begin{aligned}
\lambda(\vartheta) W(x, \vartheta) &= \frac{\partial W(x, \vartheta)}{\partial \vartheta} + \max_{u} \left[H(x(\vartheta), u(\vartheta)) + \frac{\partial W(x, \vartheta)}{\partial x} g(x, u) \right]; \\
\lambda(\vartheta) &= \frac{f(\vartheta)}{1 - F(\vartheta)}.
\end{aligned} \tag{14}$$

In mathematical reliability theory, the Hazard function $\lambda(t)$ describing the life-cycle of a system usually has the following characteristics: it is decreasing during a “burn-in” period, nearly constant during “adult” period (or regime of normal exploitation), and increasing function during the "wear-out" period. Now we can apply this approach to game theory because we can consider the game as a system of interacting elements (players). One of the probability distribution commonly used to model a lifecycle of three periods is the Weibull Law [5]. Notice that Weibull Law is used in actuarial mathematics, gerontology, and for biological systems too. We choose the Weibull distribution to model the probability distribution for the random final time of the game, but of course this is a subject for further research.

The Weibull distribution has the following characteristics:

$$\begin{aligned} f(t) &= \lambda\delta(t-t_0)^{\delta-1}\mathrm{e}^{-\lambda(t-t_0)^{\delta}}; \\ \lambda(t) &= \lambda\delta(t-t_0)^{\delta-1}; \\ &t \geq t_0;\ \lambda > 0;\ \delta > 0. \end{aligned} \tag{15}$$

Here, λ and δ are two parameters. $\delta < 1$ corresponds to the "burn-in" period, $\delta = 1$ corresponds to the "adult" period and $\delta > 1$ corresponds to the "wear-out" period. It is interesting and a well-known fact in reliability theory that the Weibull distribution during the adult stage ($\delta = 1$, $\lambda(t) = \lambda = const$) is equivalent to the exponential distribution. Thus, if we use the exponential distribution for the random final time instant T, then we indeed are considering the game to be in "adult" stage or in normal exploitation regime.

Suppose that the final time instant T obeys to the exponential distribution:

$$\begin{aligned} f(t) = \lambda\mathrm{e}^{-\lambda(t-t_0)}, \quad & F(t) = 1-\mathrm{e}^{-\lambda(t-t_0)} \quad \text{if} \quad t \geq t_0, \\ & F(t) = f(t) = 0 \quad \text{if} \quad t < t_0. \end{aligned}$$

Then the Bellman function is as follows

$$\max_u \left[\int_{\vartheta}^{\infty} \int_{\vartheta}^{t} H(x(\tau), u(\tau))\lambda\mathrm{e}^{-\lambda(t-\vartheta)}\mathrm{d}\tau\mathrm{d}t \right].$$

Remark that for a problem with random duration $(T - t_0) \in [0, \infty)$, the first term on the right-hand side (13) is equal to zero ($\frac{\partial W(t,x)}{\partial t} = 0$) in the case of exponential distribution, but this is not the case for arbitrary distribution. The Hamilton–Jacobi–Bellman equation (13) then takes the form:

$$\lambda W(x,t) = \max_u \left\{ H(x(t), u(t)) + \frac{\partial W(x,t)}{\partial x} g(x,u) \right\}.$$

This equation for a problem with exponential distribution for random terminal time T is similar to the Hamilton–Jacobi–Bellman equation for an infinite time horizon problem with discount factor λ [1].

4 The Characteristic Function

The common way to define the characteristic function in $\Gamma(x_0)$ is as follows:

$$V(S, x_0) = \begin{cases} 0, & S = \emptyset; \\ \max\limits_{u_S} \min\limits_{u_{N\setminus S}} \sum_{i\in S} K_i(x_0, u), & S \subset N; \\ \max\limits_{u} \sum_{i=1}^{n} K_i(x_0, u), & S = N. \end{cases} \tag{16}$$

In that case, $V(S, x_0)$ (16) is superadditive.

However, this approach does not seem to be the best one in the context of environmental problems. It is unlikely that if a subset of players forms a coalition to tackle an environmental problem, then the remaining players would form an antagonistic anti-coalition. For environmental problems, an alternative construction of the characteristic function construction assumes that left-out players stick to their feedback Nash strategies. This approach was proposed in [3].

Then we have the following definition of the characteristic function:

$$V(S, x^*(\vartheta)) = \begin{cases} 0, \; S = \emptyset; \\ W_i(x^*(\vartheta), \vartheta), \quad i = 1, \ldots, n; \quad \{i\} \in I; \\ W_K(x^*(\vartheta), \vartheta), \; K \subseteq I. \end{cases} \tag{17}$$

where $W_i(x^*(\vartheta), \vartheta)$, $W_K(x^*(\vartheta), \vartheta)$ are the results of the corresponding Hamilton–Jacobi–Bellman equations. Remark that the function $V(S, x^*(\vartheta))$ (17) is not superadditive in general.

5 The Shapley Value

One of the main tasks in cooperative game theory is to answer the question of how to divide the joint payoff. The most popular principle is the Shapley Value [3]. Here, we introduce an algorithm to compute the Shapley Value based on their approach - which requires examining the superadditivity of the characteristic function.

(i) Maximize the total expected payoff of the grand coalition I.

$$W_I(x, \vartheta) = \max_{u_i \in U} \frac{1}{1 - F(\vartheta)} \sum_{i=1}^{n} \int_{\vartheta}^{\infty} \int_{\vartheta}^{t} h_i(x(\tau))\, f(t) \mathrm{d}\tau \mathrm{d}t,$$
$$x(\vartheta) = x.$$

Denote $\sum_{i=1}^{n} h_i(\cdot)$ by $H(\cdot)$. Then the Bellman function $W_I(x, \vartheta)$ satisfies the HJB equation (14). The results of this optimization are the optimal trajectory $x^*(t)$ and the optimal strategies $u^* = (u_1^*, \ldots, u_n^*)$.

(ii) Calculate a feedback Nash equilibrium.
Without cooperation, each player i seeks to maximize his/her expected payoff (2). Thus, player i solves a dynamic programming problem:

$$\begin{aligned} W_i(x,\vartheta) &= \max_{u_i \in U} \frac{1}{1-F(\vartheta)} \int_{\vartheta}^{\infty} \int_{\vartheta}^{t} h_i(x(\tau))\, f(t) \mathrm{d}\tau \mathrm{d}t, \\ x(\vartheta) &= x. \end{aligned}$$

Denote $h_i(\cdot)$ by $H(\cdot)$. By this notation $W_i(x,\vartheta)$ satisfies the HJB equations (14) for all $i \in I$.
Denote by $u^N(\cdot) = \{u_i^N(\cdot), \quad i = 1,\dots,n\}$ any feedback Nash equilibrium of this noncooperative game $\Gamma(x_0)$. Let the corresponding trajectory be $x^N(t)$.
Compute $W_i(x^*(\vartheta),\vartheta)$ under the condition that before time instant ϑ the players use their optimal strategies u_i^*.

(iii) Compute outcomes for all remaining possible coalitions.

$$\begin{aligned} W_K(x,\vartheta) &= \max_{u_i, i \in K} \frac{1}{1-F(\vartheta)} \sum_{i \in K} \int_{\vartheta}^{\infty} \int_{\vartheta}^{t} h_i(x(\tau))\, f(t) \mathrm{d}\tau \mathrm{d}t, \\ u_j &= u_j^N \quad \text{for} \quad j \in I \setminus K, \\ x(\vartheta) &= x. \end{aligned}$$

Here, we insert for the left-out players $i \in I \setminus K$ their Nash values (see Step 2).
By the notation $\sum_{i \in K} h_i(\cdot) = H(\cdot)$, the Bellman function $W_K(x,\vartheta)$ satisfies the corresponding HJB equation (14).

(iv) Define the characteristic function $V(S, x^*(\vartheta))$, $\forall S \subseteq I$ as

$$V(S, x^*(\vartheta)) = \begin{cases} 0, & S = \emptyset; \\ W_i(x^*(\vartheta),\vartheta), & i = 1,\dots,n; \{i\} \in I; \\ W_K(x^*(\vartheta),\vartheta), & K \subseteq I. \end{cases} \tag{18}$$

(v) Test of the superadditivity for characteristic function.
We need to check whether the following is satisfied:

$$V(x_0, S_1 \cup S_2) \geq V(x_0, S_1) + V(x_0, S_2).$$

If the superadditivity of the characteristic function is satisfied, then go to the next step. Otherwise, the Shapley Value cannot be used as a sharing principle for the players.

(vi) Calculation of the Shapley Value by formula:

$$Sh_i = \sum_{\substack{S \subset N \\ i \in S}} \frac{(n-s)!(s-1)!}{n!} [V(S) - V(S \setminus \{i\})], \quad i = 1,\dots,n.$$

6 An Example. A Game Theoretic Model of Nonrenewable Resource Extraction with Random Duration

Consider the simple model of common-property nonrenewable resource extraction published in [1].

There are n symmetric players (firms or countries) in the game of nonrenewable resource extraction. Let $x(t)$ and $c_i(t)$ denote, respectively, the stock of the nonrenewable resource, such as an oil field, and player i's rate of extraction at time t. We assume that $c_i(t) \geq 0$ and that, if $x(t) = 0$, then the only feasible rate of extraction is $c_i(t) = 0$. Let the transition equation take the form

$$\begin{aligned} \dot{x}(t) &= -\sum_{i=1}^{n} c_i(t), \quad i = 1, \ldots, n; \\ x(t_0) &= x_0. \end{aligned}$$

The game starts at $t_0 = 0$ from x_0. We suppose that the game ends at the random time instant T obeying Weibull distribution (15).

Each player i has a utility function $h(c_i)$, defined for all $c_i > 0$ by

$$h(c_i) = Aln(c_i) + B.$$

Here, A is positive and B is a nonrestricted constant. Without loss of generality, assume $A = 1$ and $B = 0$.

As in the general case we define the integral expected payoff

$$K_i(x_0, c_1, \ldots, c_n) = \int_0^{\infty} \int_0^{t} h(c_i(\tau)) f(t) \mathrm{d}\tau \mathrm{d}t, \quad i = 1, \ldots, n$$

and consider the total payoff in the cooperative form of the game:

$$\begin{aligned} \max_{\{c_i\}} \sum_{i=1}^{n} K_i(x_0, c_1, \ldots, c_n) &= \sum_{i=1}^{n} K_i(x_0, c_1^I, \ldots, c_n^I) \\ &= \int_0^{\infty} \int_0^{t} \sum_{i=1,\ldots,n} h(c_i^I) f(t) \mathrm{d}\tau \mathrm{d}t. \end{aligned}$$

Here, $f(t)$ is the density function for Weibull distribution (15). Now we apply the five steps of our algorithm to obtain the Shapley Value.

Step 1. Consider the grand coalition $I = \{1, \ldots, n\}$. Then the Bellman function is as follows:

$$W_I(x, \vartheta) = \max_{\{c_i, i \in I\}} \int_{\vartheta}^{\infty} \int_{\vartheta}^{t} (ln(c_i) + \sum_{j \neq i} ln(c_j)) f_{\vartheta}(t) \mathrm{d}\tau \mathrm{d}t. \qquad (19)$$

Define $\sum_{i=1}^{n} h_i(c_i(\cdot)) = H(c())$. Then we can use the Hamilton–Jacobi–Bellman equation (14) for Bellman function $W_I(x,t)$:

$$\lambda(\vartheta)W_I(x,\vartheta) = \frac{\partial W_I(x,\vartheta)}{\partial \vartheta} + \max_{\{c_i, i\in I\}}\left[H(c(\vartheta)) + \frac{\partial W_I(x,\vartheta)}{\partial x} g(x,u)\right]. \quad (20)$$

Combining (20) and (19), we obtain

$$\lambda(\vartheta)W_I(x,\vartheta) = \frac{\partial W_I(x,\vartheta)}{\partial \vartheta} + \max_{\{c_i, i\in I\}}\left(ln(c_i) + \sum_{j\neq i} ln(c_j) + \frac{\partial W_I(x,\vartheta)}{\partial x}(-c_i - \sum_{j\neq i} c_j)\right). \quad (21)$$

Suppose the Bellman function W_I has the form

$$W_I(x,t) = A_I(t)ln(x) + B_I(t). \quad (22)$$

Then from (22) we get

$$\frac{\partial W_I(x,t)}{\partial x} = \frac{A_I(t)}{x}; \qquad \frac{\partial W_I(x,t)}{\partial t} = \dot{A}_I(t)ln(x) + \dot{B}_I(t).$$

Differentiating the right-hand side of (21) with respect to c_i, we obtain the optimal strategies

$$c_i^I = \frac{1}{\frac{\partial W_I}{\partial x}}. \quad (23)$$

Substituting (22) and (23) in (21), we obtain a system of equations satisfied by coefficients $A_I(t)$, $B_I(t)$:

$$\dot{A}_I(t) - \lambda(t)A_I(t) + n = 0;$$
$$\dot{B}_I(t) - \lambda(t)B_I(t) - nln(B_I(t)) - n = 0.$$

Then we obtain the optimal extraction rule:

$$c_i^I = \frac{\mathrm{e}^{-\lambda(t)t}}{n\int_t^\infty \mathrm{e}^{-\lambda(s)s}\mathrm{d}s} x. \quad (24)$$

Notice that we represent the result for an arbitrary distribution with the Hazard function $\lambda(t)$ by formula (24). If we insert the Hazard function $\lambda(t)$ corresponding to distribution (15), then we get a very complicated formula for optimal strategies (24) and as a consequence for all follow-up calculations. But if we consider only the “adult” stage of the game, then we can obtain results in explicit form. In our example of nonrenewable resource extraction, "adult" stage means that the players (and, in particular, their equipments) stay in the regime of normal exploitation of resource. As mentioned above, the Weibull distribution for the "adult" period corresponds to the exponential distribution with Hazard function $\lambda(t) = \lambda$. Then from (24) we get the result for optimal extraction rule in the case of exponential distribution:

$$c_i^I = \frac{\lambda}{n}x, \qquad i = 1, \dots, n. \tag{25}$$

Finally, we have optimal trajectory and optimal controls

$$x^I(t) = x_o * \mathrm{e}^{-\lambda(t-t_0)};$$
$$c_i^I(t) = \frac{x_0\lambda}{n}\mathrm{e}^{-\lambda(t-t_0)},$$

and

$$\begin{aligned} V(I, x^I(\vartheta)) = W_I(x^I(\vartheta), \vartheta) &= \frac{n}{\lambda}ln(x^I) - \frac{n}{\lambda} - \frac{nln(n)}{\lambda} + \frac{nln(\lambda)}{\lambda} \\ &= \frac{n}{\lambda}ln(x_0) - n(\vartheta - t_0) - \frac{n}{\lambda} - \frac{nln(n)}{\lambda} + \frac{nln(\lambda)}{\lambda}. \end{aligned}$$

Notice that the optimal trajectory $x^I(t)$ satisfies Lyapunov stability condition. Let $\vartheta = t_0$. Then

$$V(I, x_0) = W_I(x_0, t_0) = \frac{n}{\lambda}ln(x_0) - \frac{n}{\lambda} - \frac{nln(n)}{\lambda} + \frac{nln(\lambda)}{\lambda}. \tag{26}$$

Step 2. We now find a feedback Nash equilibrium. The Bellman function for player i is as follows:

$$W_i(x, \vartheta) = \max_{c_i} \int_{\vartheta}^{\infty}\int_{\vartheta}^{t} ln(c_i) f_{\vartheta}(t)\mathrm{d}\tau\mathrm{d}t.$$

The initial state is

$$x(\vartheta) = x^I(\vartheta).$$

Now the HJB equation (14) has the form

$$\lambda(\vartheta)W_i(x, \vartheta) = \frac{\partial W_i(x, \vartheta)}{\partial \vartheta} + \max_{c_i}\left(ln(c_i) + \frac{\partial W_i(x, \vartheta)}{\partial x}(-c_i - \sum_{j \neq i} c_j)\right).$$

We find $W_i(x, t)$ in the form

$$W_i(x, t) = A_N(t)ln(x) + B_N(t).$$

As before we get extraction rule in the form of feedback strategies:

$$c_i^N = \frac{e^{-\lambda(t)t}}{\int\limits_t^\infty e^{-\lambda(s)s} ds} x, \qquad i = 1, \dots, n.$$

For the case of adult stage ($\lambda(t) = \lambda$) or exponential distribution of final game time, we get the results for feedback Nash equilibrium, trajectory, and value of characteristic function in explicit form:

$$\begin{aligned} c_i^N &= \lambda x, \qquad i = 1, \dots, n; \\ x^N(t) &= x^I(\vartheta) * e^{-n\lambda(t-\vartheta)}; \\ c_i^N(t) &= \lambda x^I(\vartheta) * e^{-n\lambda(t-\vartheta)}; \\ V(\{i\}, x^I(\vartheta)) &= W_i(x^I(\vartheta)) = \frac{ln(x^I(\vartheta))}{\lambda} - \frac{n}{\lambda} + \frac{ln(\lambda)}{\lambda}. \end{aligned} \tag{27}$$

Let $\vartheta = t_0$. Then

$$V(\{i\}, x_0) = W_i(x_0, t_0) = \frac{ln(x_0)}{\lambda} - \frac{n}{\lambda} + \frac{ln(\lambda)}{\lambda}.$$

The main results obtained by Steps 1 and 2 have been published in [1] for the case of discounted utility functions with infinite time horizon.

Step 3. Consider a coalition $K \subset I$, $|K| = k$, $|I \setminus K| = n - k$. For this case, we have the Bellman function:

$$W_K(x, \vartheta) = \max_{\{c_i, i \in K\}} \int_\vartheta^\infty \int_\vartheta^t \sum_{i \in K} ln(c_i) f_\vartheta(t) d\tau dt.$$

The initial state is

$$x(\vartheta) = x^I(\vartheta).$$

For the sake of simplicity, we again consider exponential distribution. Recall that the left-out players $i \in I \setminus K$ use feedback Nash strategies (27).

In the same way, we get

$$\begin{aligned} x^K(t) &= x^I(\vartheta) * e^{-(n-k+1)\lambda(t-\vartheta)}; \\ c_i^K(t) &= \frac{\lambda}{k} x^I(\vartheta) * e^{-(n-k+1)\lambda(t-\vartheta)}; \\ V(K, x^I(\vartheta)) &= W_K(x^I(\vartheta), \vartheta) \\ &= \frac{k}{\lambda} ln(x^I(\vartheta)) - \frac{k}{\lambda} - \frac{k(n-k)}{\lambda} - \frac{k}{\lambda} ln(k) + \frac{k ln(\lambda)}{\lambda}. \end{aligned}$$

Let $\vartheta = t_0$. Then

$$V(K, x_0) = W_K(x_0, t_0) = \frac{k}{\lambda} ln(x_0)) - \frac{k}{\lambda} - \frac{k(n-k)}{\lambda} - \frac{k}{\lambda} ln(k) + \frac{k ln(\lambda)}{\lambda}. \quad (28)$$

Step 4. We have constructed the characteristic function $V(K, x_0), K \subseteq I$ (see (26), (28)).

Step 5.

Proposition 1. *Suppose the characteristic function $V(K, x_0), K \subseteq I$ is given by (26),(28). Then $V(K, x_0)$ is superadditive.*

To prove this proposition, we need following lemma.

Lemma 1. *Let $s_1 \geq 1, s_2 \geq 1$. Then*

$$s_1 ln(s_1) + s_2 ln(s_2) + 2s_1 s_2 \geq (s_1 + s_2) ln(s_1 + s_2).$$

The proof of this lemma is straightforward; it is easy to show that the left-hand side increases at a faster rate than the right-hand side.

Proof of Proposition1 is then easily done by direct computation.

Step 6. Finally, we get the Shapley Value in our example:

$$Sh_i(x(t)) = \frac{V(I, x)}{n} = \frac{ln(x)}{\lambda} - \frac{1}{\lambda} - \frac{ln(n)}{\lambda} + \frac{ln(\lambda)}{\lambda}; \qquad i = 1, \ldots, n,$$
$$Sh_i(x_0) = \frac{V(I, x_0)}{n} = \frac{ln(x_0)}{\lambda} - \frac{1}{\lambda} - \frac{ln(n)}{\lambda} + \frac{ln(\lambda)}{\lambda}; \qquad i = 1, \ldots, n. \quad (29)$$

Thus, the allocation of the total expected payoff is represented by the Shapley Value (29) and the optimal behavior of the players (namely the rate of resource extraction) is determined by the formula (25).

7 Conclusion

In this chapter, we investigated a new model in differential game theory. A differential game has been considered on the time interval $[t_0, T]$, where T is a random variable. This model allows a more realistic representation of real-life processes. One of the main objects of the paper is the derivation of a Hamilton–Jacobi–Bellman equation suitable for a problem with random duration. On the basis of this derived equation, we suggest an algorithm for the computation of the Shapley Value for cooperative differential games with random duration, based on the approach suggested in [3]. To illustrate the application of this algorithm, we obtained the expression of the Shapley Value for a simple model of nonrenewable resource extraction under condition of a random game duration with exponential distribution.

Acknowledgements Many thanks to Professor Leon A. Petrosjan for the initialization and support of my research. Moreover, I wish to thank Professor George Zaccour and Mr. Dmitry Gromov for assistance at an early stage of acquaintance with the material and during the further work. I express my gratitude to Professor Jesús Marín-Solano for his valuable comments about dynamic programming problems. Finally, I thank the editor Professor Michèle Breton and two anonymous reviewers for very helpful comments on the previous draft version of the chapter.

References

1. Dockner, E.J., Jorgensen, S., van Long, N., Sorger, G.: Differential Games in Economics and Management Science. Cambridge University Press, Cambridge, U.K. (2000)
2. Henley, E.J., Kumamoto, H.: Reliability Engineering and Risk Assessment. Prentice Hall (1981)
3. Petrosjan, L.A., Zaccour, G.: Time-consistent Shapley Value Allocation of Pollution Cost Reduction. Journal of Economic Dynamics and Control **27**, 381–398 (2003)
4. Petrosjan, L.A., Shevkoplyas, E.V.: Cooperative Solutions for Games with Random Duration. Game Theory and Applications **IX**, Nova Science Publishers, 125–139 (2003)
5. Weibull, W.: A statistical distribution function of wide applicability. Journal of Applied Mechanics **18**, 293–297 (1951)

Dynamically Consistent Cooperative Solutions in Differential Games with Asynchronous Players' Horizons

David W.K. Yeung

Abstract This paper considers cooperative differential games in which players enter the game at different times and have diverse horizons. Moreover, the types of future players are not known with certainty. Dynamically consistent cooperative solutions and analytically tractable payoff distribution mechanisms leading to the realization of these solutions are derived. This analysis widens the application of cooperative differential game theory to problems where the players' game horizons are asynchronous and the types of future players are uncertain. It represents the first attempt to seek dynamically consistent solution for cooperative games with asynchronous players' horizons and uncertain types of future players.

1 Introduction

In many game situations, the players' time horizons differ. This may arise from different life spans, different entry and exit times in different markets, and the different duration for leases and contracts. Asynchronous horizon game situations occur frequently in economic and social activities. In this paper, we consider cooperative differential games in which players enter the game at different times and have diverse horizons. Moreover, the types of future players are not necessarily known with certainty.

Cooperative games suggest the possibility of socially optimal and group efficient solutions to decision problems involving strategic action. In dynamic cooperative games, a stringent condition for a dynamically stable solution is required: In the solution, the optimality principle must remain optimal throughout the game, at any instant of time along the optimal state trajectory determined at the outset. This condition is known as *dynamic stability* or *time consistency*. The question of dynamic

D.W.K. Yeung (✉)
Department of Business Administration, Hong Kong Shue Yan University, Hong Kong
and
Center of Game Theory, St. Petersburg State University, St. Petersburg, Russia
e-mail: dwkyeung@hksyu.edu

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_19,

stability in differential games has been rigorously explored in the past three decades. (see [3, 5, 6]). In the presence of stochastic elements, a more stringent condition – that of *subgame consistency* – is required for a dynamically stable cooperative solution. In particular, 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 behavior would remain optimal. In particular, dynamic consistency ensures that as the game proceeds players are guided by the same optimality principle at each instant of time, and hence do not possess incentives to deviate from the previously adopted optimal behavior. A rigorous framework for the study of subgame-consistent solutions in cooperative stochastic differential games was established in the work of Yeung and Petrosyan [8–10]. A generalized theorem was developed for the derivation of an analytically tractable payoff distribution procedure leading to dynamically consistent solutions.

In this paper, dynamically consistent cooperative solutions are derived for differential games with asynchronous players' horizons and uncertain types of future players. Analytically tractable payoff distribution mechanisms which lead to the realization of these solutions are derived. This analysis extends the application of cooperative differential game theory to problems where the players' game horizons are asynchronous and the types of future players are uncertain. The organization of the paper is as follows. Section 2 presents the game formulation and characterizes noncooperative outcomes. Dynamic cooperation among players coexisting in the same duration is examined in Sect. 3. Section 4 provides an analysis on payoff distribution procedures leading to dynamically consistent solutions in this asynchronous horizons scenario. An illustration in cooperative resource extraction is given in Sect. 5. Concluding remarks and model extensions are given in Sect. 6.

2 Game Formulation and Noncooperative Outcome

In this section, we first present a simple analytical framework of differential games with asynchronous players' horizons, and characterize its noncooperative outcome.

2.1 Game Formulation

For clarity in exposition and without loss of generality, we consider a general class of differential games, in which there are $\upsilon+1$ overlapping cohorts or generations of players. The game begins at time t_1 and terminates at time $t_{\upsilon+1}$. In the time interval $[t_1, t_2)$, there coexist a generation 0 player whose game horizon is $[t_1, t_2)$ and a generation 1 player whose game horizon is $[t_1, t_3)$. In the time interval $[t_k, t_{k+1})$ for $k \in \{2, 3, \cdots, \upsilon-1\}$, there coexist a generation $k-1$ player whose game horizon is $[t_{k-1}, t_{k+1})$ and a generation k player whose game horizon is $[t_k, t_{k+2})$. In the last time interval $[t_\upsilon, t_{\upsilon+1})$, there coexist a generation $\upsilon-1$ player and a generation υ player whose game horizon is just $[t_\upsilon, t_{\upsilon+1}]$.

For the sake of notational convenience in exposition, the players are of types $\omega_{a_k} \in \{\omega_1, \omega_2, \cdots, \omega_\varsigma\}$. When the game starts at initial time t_1, it is known that in the time interval $[t_1, t_2)$, there coexist a type ω_1 generation 0 player and a type ω_2 generation 1 player. At time t_1, it is also known that the probability of the generation k player being type $\omega_{a_k} \in \{\omega_1, \omega_2, \cdots, \omega_\varsigma\}$ is $\lambda_{a_k} \in \{\lambda_1, \lambda_2, \cdots, \lambda_\varsigma\}$, for $k \in \{2, 3, \cdots, \upsilon\}$. The type of generation k player will become known with certainty at time t_k.

The instantaneous payoff functions and terminal rewards of the type ω_{a_k} generation k player and the type $\omega_{a_{k-1}}$ generation $k-1$ player coexisting in the time interval $[t_k, t_{k+1})$ are, respectively:

$$g^{k-1(\omega_{k-1})}\left[s, x(s), u_{k-1}^{(\omega_{k-1},O)\omega_k}(s), u_k^{(\omega_k,Y)\omega_{k-1}}(s)\right]$$
$$\text{and } q^{k-1(\omega_{k-1})}\left[t_{k+1}, x(t_{k+1})\right], \text{and}$$
$$g^{k(\omega_k)}\left[s, x(s), u_{k-1}^{(\omega_{k-1},O)\omega_k}(s), u_k^{(\omega_k,Y)\omega_{k-1}}(s)\right] \text{ and } q^{k(\omega_k)}[t_{k+2}, x(t_{k+2})],$$

for $k \in \{1, 2, 3, \cdots, \upsilon\}$, where $u_{k-1}^{(w_{k-1},O)\omega_k}(s)$ is the vector of controls of the type $\omega_{a_{k-1}}$ generation $k-1$ player when he/she is in his/her last (old) life stage while the type ω_{a_k} generation k player is coexisting; and $u_k^{(w_k,Y)\omega_{k-1}}(s)$ is that of the type ω_{a_k} generation k player when he/she is in his/her first (young) life stage while the type $\omega_{a_{k-1}}$ generation $k-1$ player is coexisting.

Note that the superindex "O" in $u_{k-1}^{(w_{k-1},O)\omega_k}(s)$ denote Old and the superindex "Y" in $u_k^{(w_k,Y)\omega_{k-1}}(s)$ denote young. The state dynamics of the game is characterized by the vector-valued differential equations:

$$\dot{x}(s) = f[s, x(s), u_{k-1}^{(\omega_{k-1},O)\omega_k}(s), u_k^{(\omega_k,Y)\omega_{k-1}}(s)], \text{for } s \in [t_k, t_{k+1}), \tag{1}$$

if the type ω_{a_k} generation k player and the type $\omega_{a_{k-1}}$ generation a_{k-1} player coexisting in the time interval $[t_k, t_{k+1})$ for $k \in \{1, 2, 3, \cdots, \upsilon\}$, and $x(t_1) = x_0 \in X$.

In the game interval $[t_k, t_{k+1})$ for $k \in \{1, 2, 3, \cdots, \upsilon - 1\}$ with type ω_{k-1} generation $k-1$ player and type ω_k generation k player, the type ω_{k-1} generation $k-1$ player seeks to maximize:

$$\int_{t_k}^{t_{k+1}} g^{k-1(\omega_{k-1})}\left[s, x(s), u_{k-1}^{(\omega_{k-1},O)\omega_k}(s), u_k^{(\omega_k,Y)\omega_{k-1}}(s)\right] \mathrm{e}^{-r(s-t_k)} \mathrm{d}s$$
$$+\mathrm{e}^{-r(t_{k+1}-t_k)} q^{k-1(\omega_{k-1})}[t_{k+1}, x(t_{k+1})]$$

and the type ω_k generation k player seeks to maximize:

$$\int_{t_k}^{t_{k+1}} g^{k(\omega_k)}\left[s, x(s), u_{k-1}^{(\omega_{k-1},O)\omega_k}(s), u_k^{(\omega_k,Y)\omega_{k-1}}(s)\right] \mathrm{e}^{-r(s-t_k)} \mathrm{d}s$$
$$+\sum_{\alpha=1}^{\varsigma} \lambda_{a_{k+1}} \int_{t_{k+1}}^{t_{k+2}} g^{k(\omega_k)}\left[s, x(s), u_k^{(\omega_k,O)\omega_\alpha}(s), u_{k+1}^{(\omega_\alpha,Y)\omega_k}(s)\right] \mathrm{e}^{-r(s-t_k)} \mathrm{d}s$$
$$+\mathrm{e}^{-r(t_{k+2}-t_k)} q^{k(\omega_k)}[t_{k+2}, x(t_{k+2})],$$

subject to dynamics (1), where r is the discount rate.

In the last time interval $[t_\upsilon, t_{\upsilon+1}]$ when the generation $\upsilon - 1$ player is of type $\omega_{\upsilon-1}$ and the generation υ player is of type ω_υ, the type $\omega_{\upsilon-1}$ generation $\upsilon - 1$ player seeks to maximize:

$$\int_{t_\upsilon}^{t_{\upsilon+1}} g^{\upsilon-1(\omega_{\upsilon-1})}\left[s, x(s), u_{\upsilon-1}^{(\omega_{\upsilon-1},O)\omega_\upsilon}(s), u_\upsilon^{(\omega_\upsilon,Y)\omega_{\upsilon-1}}(s)\right] \mathrm{e}^{-r(s-t_\upsilon)}\mathrm{d}s$$
$$+\mathrm{e}^{-r(t_{\upsilon+1}-t_\upsilon)} q^{\upsilon-1(\omega_{\upsilon-1})}\left[t_{\upsilon+1}, x(t_{\upsilon+1})\right] \tag{2}$$

and the type ω_υ generation υ player seeks to maximize:

$$\int_{t_\upsilon}^{t_{\upsilon+1}} g^{\upsilon(\omega_\upsilon)}\left[s, x(s), u_{\upsilon-1}^{(\omega_{\upsilon-1},O)\omega_\upsilon}(s), u_\upsilon^{(\omega_\upsilon,Y)\omega_{\upsilon-1}}(s)\right] \mathrm{e}^{-r(s-t_\upsilon)}\mathrm{d}s$$
$$+\mathrm{e}^{-r(t_{\upsilon+1}-t_\upsilon)} q^{\upsilon(\omega_\upsilon)}\left[t_{\upsilon+1}, x(t_{\upsilon+1})\right], \tag{3}$$

subject to dynamics (1).

The game formulated is a finite overlapping generations version of Jørgensen and Yeung's [4] infinite generations game.

2.2 Noncooperative Outcomes

To obtain a characterization of a noncooperative solution to the asynchronous horizons game mentioned above, we first consider the solutions of the games in the last time interval $[t_\upsilon, t_{\upsilon+1}]$, that is the game (2)–(3). One way to characterize and derive a feedback solution to the games in $[t_\upsilon, t_{\upsilon+1}]$ is to invoke the Hamilton–Jacobi–Bellman equations approach and obtain:

Lemma 1. *If the generation $\upsilon - 1$ player is of type $\omega_{\upsilon-1} \in \{\omega_1, \omega_2, \cdots, \omega_\varsigma\}$ and the generation υ player is of type $\omega_\upsilon \in \{\omega_1, \omega_2, \cdots, \omega_\varsigma\}$ in the time interval $[t_\upsilon, t_{\upsilon+1}]$, a set of feedback strategies $\left\{\phi_{\upsilon-1}^{(\omega_{\upsilon-1},O)\omega_\upsilon}(t,x); \phi_\upsilon^{(\omega_\upsilon,Y)\omega_{\upsilon-1}}(t,x)\right\}$ constitutes a Nash equilibrium solution for the game (2)–(3), if there exist continuously differentiable functions $V^{\upsilon-1(\omega_{\upsilon-1},O)\omega_\upsilon}(t,x) : [t_\upsilon, t_{\upsilon+1}] \times R^m \to R$ and $V^{\upsilon(\omega_\upsilon,Y)\omega_{\upsilon-1}}(t,x) : [t_\upsilon, t_{\upsilon+1}] \times R^m \to R$ satisfying the following partial differential equations:*

$$-V_t^{\upsilon-1(\omega_{\upsilon-1},O)\omega_\upsilon} = \max_{u_{\upsilon-1}} \left\{ g^{\upsilon-1(\omega_{\upsilon-1})}\left[t, x, u_{\upsilon-1}, \phi_\upsilon^{(\omega_\upsilon,Y)\omega_{\upsilon-1}}(t,x)\right] \mathrm{e}^{-r(t-t_\upsilon)} \right.$$
$$\left. + V_x^{\upsilon-1(\omega_{\upsilon-1},O)\omega_\upsilon} f\left[t, x, u_{\upsilon-1}, \phi_\upsilon^{(\omega_\upsilon,Y)\omega_{\upsilon-1}}(t,x)\right]\right\},$$
$$V^{\upsilon-1(\omega_{\upsilon-1},O)\omega_\upsilon}(t_{\upsilon+1}, x) = \mathrm{e}^{-r(t_{\upsilon+1}-t_\upsilon)} q^{\upsilon-1(\omega_{\upsilon-1})}(t_{\upsilon+1}, x), \text{ and}$$
$$-V_t^{\upsilon(\omega_\upsilon,Y)\omega_{\upsilon-1}} = \max_{u_\upsilon} \left\{ g^{\upsilon(\omega_\upsilon)}\left[t, x, \phi_{\upsilon-1}^{(\omega_{\upsilon-1},O)\omega_\upsilon}(t,x), u_\upsilon\right] \mathrm{e}^{-r(t-t_\upsilon)} \right.$$
$$\left. + V_x^{\upsilon(\omega_\upsilon,Y)\omega_{\upsilon-1}} f\left[t, x, \phi_{\upsilon-1}^{(\omega_{\upsilon-1},O)\omega_\upsilon}(t,x), u_\upsilon\right]\right\},$$
$$V^{\upsilon(\omega_\upsilon,Y)\omega_{\upsilon-1}}(t_{\upsilon+1}, x) = \mathrm{e}^{-r(t_{\upsilon+1}-t_\upsilon)} q^{\upsilon(\omega_\upsilon)}\left[(t_{\upsilon+1}, x(t_{\upsilon+1})\right].$$

Proof. Follow the proof of Theorem 6.16 in Chap. 6 of Basar and Olsder [1]. □

For ease of exposition and sidestepping the issue of multiple equilibria, the analysis focuses on solvable games in which a particular noncooperative Nash equilibrium is chosen by the players in the entire subgame.

We proceed to examine the game in the second last interval $[t_{\upsilon-1}, t_\upsilon)$. If the generation $\upsilon - 2$ player is of type $\omega_{\upsilon-2} \in \{\omega_1, \omega_2, \cdots, \omega_\varsigma\}$ and the generation $\upsilon - 1$ player is of type $\omega_{\upsilon-1} \in \{\omega_1, \omega_2, \cdots, \omega_\varsigma\}$. Then the type $\omega_{\upsilon-2}$ generation $\upsilon - 2$ player seeks to maximize:

$$\int_{t_{\upsilon-1}}^{t_\upsilon} g^{\upsilon-2(\omega_{\upsilon-2})}\left[s, x(s), u_{\upsilon-2}^{(\omega_{\upsilon-2},O)\omega_{\upsilon-1}}(s), u_{\upsilon-1}^{(\omega_{\upsilon-1},Y)\omega_{\upsilon-2}}(s)\right] \mathrm{e}^{-r(s-t_{\upsilon-1})}\mathrm{d}s$$
$$+\mathrm{e}^{-r(t_\upsilon-t_{\upsilon-1})} q^{\upsilon-2(\omega_{\upsilon-2})}\left[t_\upsilon, x(t_\upsilon)\right].$$

As shown in Jørgensen and Yeung [4], the terminal condition of the type $\omega_{\upsilon-1}$ generation $\upsilon - 1$ player in the game interval $[t_{\upsilon-1}, t_\upsilon)$ can be expressed as:

$$\sum_{\alpha=1}^{\varsigma} \lambda_\alpha V^{\upsilon-1(\omega_{\upsilon-1},O)\omega_\alpha}(t_\upsilon, x).$$

The type $\omega_{\upsilon-1}$ generation $\upsilon - 1$ player then seeks to maximize:

$$\int_{t_{\upsilon-1}}^{t_\upsilon} g^{\upsilon-1(\omega_{\upsilon-1})}\left[s, x(s), u_{\upsilon-2}^{(\omega_{\upsilon-2},O)\omega_{\upsilon-1}}(s), u_{\upsilon-1}^{(\omega_{\upsilon-1},Y)\omega_{\upsilon-2}}(s)\right] \mathrm{e}^{-r(s-t_{\upsilon-1})}\mathrm{d}s$$
$$+\mathrm{e}^{-r(t_\upsilon-t_{\upsilon-1})} \sum_{\alpha=1}^{\varsigma} \lambda_\alpha V^{\upsilon-1(\omega_{\upsilon-1},O)\omega_\alpha}(t_\upsilon, x(t_\upsilon)).$$

Similarly, the terminal condition of the type ω_k generation k player in the game interval $[t_k, t_{k+1})$ can be expressed as:

$$\sum_{\alpha=1}^{\varsigma} \lambda_\alpha V^{k(\omega_k,O)\omega_\alpha}(t_{k+1}, x), \text{for } k \in \{1, 2, \cdots, \upsilon - 3\}.$$

Consider the game in the time interval $[t_k, t_{k+1})$ involving the type ω_k generation k player and the type ω_{k-1} generation $k - 1$ player, for $k \in \{1, 2, \cdots, \upsilon - 3\}$. The type ω_{k-1} generation $k - 1$ player will maximize the payoff

$$\int_{t_k}^{t_{k+1}} g^{k-1(\omega_{k-1})}\left[s, x(s), u_{k-1}^{(\omega_{k-1},O)\varpi_k}(s), u_k^{(\omega_k,Y)\omega_{k-1}}(s)\right] \mathrm{e}^{-r(s-t_k)}\mathrm{d}s$$
$$+\mathrm{e}^{-r(t_{k+1}-t_k)} q^{k-1(\omega_{k-1})}\left[t_{k+1}, x(t_{k+1})\right], \tag{4}$$

and the type ω_k generation k player will maximize the payoff:

$$\int_{t_k}^{t_{k+1}} g^{k(\omega_k)}\left[s, x(s), u_{k-1}^{(\omega_{k-1},O)\omega_k}(s), u_k^{(\omega_k,Y)\omega_{k-1}}(s)\right] \mathrm{e}^{-r(s-t_k)} \mathrm{d}s$$
$$+ \mathrm{e}^{-r(t_{k+1}-t_k)} \sum_{\alpha=1}^{\varsigma} \lambda_\alpha V^{k(\omega_k,O)\omega_\alpha}(t_{k+1}, x) \tag{5}$$

subject to (1).

A feedback solution to the game (4)–(5) can be characterized as:

Lemma 2. *A set of feedback strategies* $\{\phi_{k-1}^{(\omega_{k-1},O)\omega_k}(t,x); \phi_k^{(\omega_k,Y)\omega_{k-1}}(t,x)\}$ *constitutes a Nash equilibrium solution for the game (4)–(5), if there exist continuously differentiable functions* $V^{k-1(\omega_{k-1},O)\omega_k}(t,x) : [t_k, t_{k+1}] \times R^m \to R$ *and* $V^{k(\omega_k,Y)\omega_{k-1}}(t,x) : [t_k, t_{k+1}] \times R^m \to R$ *satisfying the following partial differential equations:*

$$-V_t^{k-1(\omega_{k-1},O)\omega_k} = \max_{u_{k-1}} \left\{ g^{k-1(\omega_{k-1})}\left[t, x, u_{k-1}, \phi_k^{(\omega_k,Y)\omega_{k-1}}(t,x)\right] \mathrm{e}^{-r(t-t_k)} \right.$$
$$\left. + V_x^{k-1(\omega_{k-1},O)\omega_k} f\left[t, x, u_{k-1}, \phi_k^{(\omega_k,Y)\omega_{k-1}}(t,x)\right] \right\},$$
$$V^{k-1(\omega_{k-1},O)\omega_k}(t_{k+1}, x) = \mathrm{e}^{-r(t_{k+1}-t_k)} q^{k-1(\omega_{k-1})}(t_{k+1}, x), \textit{and}$$
$$-V_t^{k(\omega_k,Y)\omega_{k-1}} = \max_{u_k} \left\{ g^{(k,\omega_k)}\left[t, x, \phi_{k-1}^{(\omega_{k-1},O)\omega_k}, u_k\right] \mathrm{e}^{-r(t-t_k)} \right.$$
$$\left. + V_x^{k(\omega_k,Y)\omega_{k-1}} f\left[t, x, \phi_{k-1}^{(\omega_{k-1},O)\omega_k}, u_k\right] \right\},$$
$$V^{k(\omega_k,Y)\omega_{k-1}}(t_{k+1}, x) = \mathrm{e}^{-r(t_{k+1}-t_k)} \sum_{\alpha=1}^{\varsigma} \lambda_\alpha V^{k(\omega_k,O)\omega_\alpha}(t_{k+1}, x).$$

Proof. Again follow the proof of Theorem 6.16 in Chap. 6 of Basar and Olsder [1]. □

3 Dynamic Cooperation among Coexisting Players

Now consider the case when coexisting players want to cooperate and agree to act and allocate the cooperative payoff according to a set of agreed-upon optimality principles. The agreement on how to act cooperatively and allocate cooperative payoff constitutes the solution optimality principle of a cooperative scheme. In particular, the solution optimality principle for the cooperative game includes (a) an agreement on a set of cooperative strategies/controls, and (b) an imputation of their payoffs.

Consider the game in the time interval $[t_k, t_{k+1})$ involving the type ω_k generation k player and the type ω_{k-1} generation $k-1$ player. Let $\varpi_h^{(\omega_{k-1},\omega_k)}$ denote the probability that the type ω_k generation k player and the type ω_{k-1} generation $k-1$ player would agree to the solution imputation $\left[\xi^{k-1(\omega_{k-1},O)\omega_k[h]}(t,x), \xi^{k(\omega_k,Y)\omega_{k-1}[h]}(t,x)\right]$ over the time interval $[t_k, t_{k+1})$, where $\sum_{h=1}^{\varsigma(\omega_{k-1},\omega_k)} \varpi_h^{(\omega_{k-1},\omega_k)} = 1$

At time t_1, the agreed-upon imputation for the type ω_1 generation 0 player and the type ω_2 generation 1 player 1 are known.

The solution imputation may be governed by many specific principles. For instance, the players may agree to maximize the sum of their expected payoffs and equally divide the excess of the cooperative payoff over the noncooperative payoff. As another example, the solution imputation may be an allocation principle in which the players allocate the total joint payoff according to the relative sizes of the players' noncooperative payoffs. Finally, it is also possible that the players refuse to cooperate. In that case, the imputation vector becomes $\left[V^{k-1(\omega_{k-1},O)\omega_k}(t,x), V^{k(\omega_k,Y)\omega_{k-1}}(t,x)\right]$.

Both group optimality and individual rationality are required in a cooperative plan. Group optimality requires the players to seek a set of cooperative strategies/controls that yields a Pareto optimal solution. The allocation principle has to satisfy individual rationality in the sense that neither player would be no worse off than before under cooperation.

3.1 Group Optimality

Since payoffs are transferable, group optimality requires the players coexisting in the same time interval to maximize their expected joint payoff. Consider the last time interval $[t_\upsilon, t_{\upsilon+1}]$, in which the generation $\upsilon - 1$ player is of type $\omega_{\upsilon-1} \in \{\omega_1, \omega_2, \cdots, \omega_\varsigma\}$ and the generation υ player is of type $\omega_\upsilon \in \{\omega_1, \omega_2, \cdots, \omega_\varsigma\}$. The players maximize their joint profit

$$\int_{t_\upsilon}^{t_{\upsilon+1}} \Big(g^{\upsilon-1(\omega_{\upsilon-1})} \left[s, x(s), u_{\upsilon-1}^{(\omega_{\upsilon-1},O)\omega_\upsilon}(s), u_\upsilon^{(\omega_\upsilon,Y)\omega_{\upsilon-1}}(s) \right]$$
$$+ g^{\upsilon(\omega_\upsilon)} \left[s, x(s), u_{\upsilon-1}^{(\omega_{\upsilon-1},O)\omega_\upsilon}(s), u_\upsilon^{(\omega_\upsilon,Y)\omega_{\upsilon-1}}(s) \right] \Big) e^{-r(s-t_\upsilon)} ds$$
$$+ e^{-r(t_{\upsilon+1}-t_\upsilon)} \Big(q^{\upsilon-1(\omega_{\upsilon-1})} [t_{\upsilon+1}, x(t_{\upsilon+1})] + q^{\upsilon(\omega_\upsilon)} [t_{\upsilon+1}, x(t_{\upsilon+1})] \Big), \quad (6)$$

subject to (1).

Invoking Bellman's [2] techniques of dynamic programming an optimal solution of the problem (6)–(1) can be characterized as:

Lemma 3. *A set of Controls* $\left\{ \psi_{\upsilon-1}^{(\omega_{\upsilon-1},O)\omega_\upsilon}(t,x); \psi_\upsilon^{(\omega_\upsilon,Y)\omega_{\upsilon-1}}(t,x) \right\}$ *constitutes an optimal solution for the control problem (6)–(1), if there exist continuously differentiable functions* $W^{[t_\upsilon,t_{\upsilon+1}](\omega_{\upsilon-1},\omega_\upsilon)}(t,x) : [t_\upsilon, t_{\upsilon+1}] \times R^m \to R$ *satisfying the following partial differential equations:*

$$
\begin{aligned}
-W_t^{[t_\upsilon, t_{\upsilon+1}](\omega_{\upsilon-1},\omega_\upsilon)}(t,x) = \max_{u_{\upsilon-1},u_\upsilon} \Big\{ & g^{\upsilon-1(\omega_{\upsilon-1})}[t,x,u_{\upsilon-1},u_\upsilon]\mathrm{e}^{-r(t-t_\upsilon)} \\
& + g^{v(\omega_\upsilon)}[t,x,u_{\upsilon-1},u_\upsilon]\mathrm{e}^{-r(t-t_\upsilon)} \\
& + W_x^{[t_\upsilon, t_{\upsilon+1}](\omega_{\upsilon-1},\omega_\upsilon)}(t,x) f[t,x,u_{\upsilon-1},u_\upsilon] \Big\},
\end{aligned}
$$

$$
\begin{aligned}
& W^{[t_\upsilon, t_{\upsilon+1}](\omega_{\upsilon-1},\omega_\upsilon)}(t_{\upsilon+1},x) \\
& \quad = \mathrm{e}^{-r(t_{\upsilon+1}-t_\upsilon)} \left[q^{\upsilon-1(\omega_{\upsilon-1})}(t_{\upsilon+1},x) + q^{\upsilon(\omega_\upsilon)}(t_{\upsilon+1},x) \right].
\end{aligned}
$$

We proceed to examine joint payoff maximization problem in the time interval $[t_{\upsilon-1}, t_\upsilon)$ involving the type $\omega_{\upsilon-1}$ generation $\upsilon-1$ player and type $\omega_{\upsilon-2}$ generation $\upsilon-2$ player. A critical problem is to determent the expected terminal valuation to the $\omega_{\upsilon-1}$ generation $\upsilon-1$ player at time t_υ in the optimization problem within the time interval $[t_{\upsilon-1}, t_\upsilon)$. By time t_υ, the $\omega_{\upsilon-1}$ generation $\upsilon-1$ player may coexist with the $\omega_\upsilon \in \{\omega_1, \omega_2, \cdots, \omega_\varsigma\}$ generation υ player with probabilities $\{\lambda_1, \lambda_2, \cdots, \lambda_\varsigma\}$. Consider the case in the time interval $[t_\upsilon, t_{\upsilon+1})$ in which the type $\omega_{\upsilon-1}$ generation $\upsilon-1$ player and the type ω_υ generation υ player co-exist. The probability that the type $\omega_{\upsilon-1}$ generation $\upsilon-1$ player and the type ω_υ generation υ player would agree to the solution imputation

$$
\begin{aligned}
& [\xi^{\upsilon-1(\omega_{\upsilon-1},O)\omega_\upsilon[h]}(t,x), \xi^{\upsilon(\omega_\upsilon,Y)\omega_{\upsilon-1}[h]}(t,x)] \text{ is } \varpi_h^{(\omega_{\upsilon-1},\omega_\upsilon)} \\
& \quad \text{where } \sum_h \varpi_h^{(\omega_{\upsilon-1},\omega_\upsilon)} = 1.
\end{aligned}
$$

In the optimization problem within the time interval $[t_{\upsilon-1}, t_\upsilon)$, the expected terminal reward to the $\omega_{\upsilon-1}$ generation $\upsilon-1$ player at time t_υ can be expressed as:

$$
\sum_{\alpha=1}^{\varsigma} \sum_{h=1}^{\varsigma(\omega_{\upsilon-1},\omega_\alpha)} \varpi_h^{(\omega_{\upsilon-1},\omega_\alpha)} \xi^{\upsilon-1(\omega_{\upsilon-1},O)\omega_\alpha[h]}(t_\upsilon,x). \tag{7}
$$

Similarly for the optimization problem within the time interval $[t_k, t_{k+1})$, the expected terminal reward to the ω_k generation k player at time t_{k+1} can be expressed as:

$$
\sum_{\alpha=1}^{\varsigma} \sum_{h=1}^{\varsigma(\omega_k,\omega_\alpha)} \varpi_h^{(\omega_k,\omega_\alpha)} \xi^{k(\omega_k,O)\omega_\alpha[h]}(t_{k+1},x), \text{ for } k \in \{1,2,\cdots,\upsilon-3\}. \tag{8}
$$

The joint maximization problem in the time interval $[t_k, t_{k+1})$, for $k \in \{1,2,\cdots,$ $\upsilon-3\}$, involving the type ω_k generation k player and type ω_{k-1} generation $k-1$ player can be expressed as:

$$\max_{u_{k-1},u_k}\left\{\int_{t_k}^{t_{k+1}}\Big(g^{k-1(\omega_{k-1})}[s,x(s),u_{k-1}^{(\omega_{k-1},O)\omega_k}(s),u_k^{(\omega_k,Y)\omega_{k-1}}(s)]\right.$$
$$+g^{k(\omega_k)}[s,x(s),u_{k-1}^{(\omega_{k-1},O)\omega_k}(s),u_k^{(\omega_k,Y)\omega_{k-1}}(s)]\Big)\mathrm{e}^{-r(s-t_k)}\mathrm{d}s$$
$$+\mathrm{e}^{-r(t_{k+1}-t_k)}\Big(q^{k-1(\omega_{k-1})}[t_{k+1},x(t_{k+1})]$$
$$\left.+\sum_{\alpha=1}^{\varsigma}\sum_{h=1}^{\varsigma(\omega_k,\omega_\alpha)}\varpi_h^{(\omega_k,\omega_\alpha)}\xi^{k(\omega_k,O)\omega_\alpha[h]}(t_{k+1},x(t_{k+1}))\Big)\right\}, \tag{9}$$

subject to (1).

The conditions characterizing an optimal solution of the problem (9)–(1) are given as follows.

Theorem 1. *A set of Controls* $\left\{\psi_{k-1}^{(\omega_{k-1},O)\omega_k}(t,x);\psi_k^{(\omega_k,Y)\omega_{k-1}}(t,x)\right\}$ *constitutes an optimal solution for the control problem (9)–(1), if there exist continuously differentiable functions* $W^{[t_k,t_{k+1}](\omega_{k-1},\omega_k)}(t,x):[t_k,t_{k+1})\times R^m\to R$ *satisfying the following partial differential equations:*

$$-W_t^{[t_k,t_{k+1}](\omega_{k-1},\omega_k)}(t,x)=\max_{u_{k-1},u_k}\Big\{g^{k-1(\omega_{k-1})}[t,x,u_{k-1},u_k]\mathrm{e}^{-r(t-t_k)}$$
$$+g^{k(\omega_k)}[t,x,u_{k-1},u_k]\mathrm{e}^{-r(t-t_k)}+W_x^{[t_k,t_{k+1}](\omega_{k-1},\omega_k)}(t,x)f[t,x,u_{k-1},u_k]\Big\},$$
$$W^{[t_k,t_{k+1}](\omega_{k-1},\omega_k)}(t_{k+1},x)=\mathrm{e}^{-r(t_{k+1}-t_k)}\Big(q^{k-1(\omega_{k-1})}(t_{k+1},x)$$
$$+\sum_{\alpha=1}^{\varsigma}\sum_{h=1}^{\varsigma(\omega_k,\omega_\alpha)}\varpi_h^{(\omega_k,\omega_\alpha)}\xi^{k(\omega_k,O)\omega_\alpha[h]}(t_{k+1},x)\Big). \tag{10}$$

Proof. Invoking Bellman's [2] technique of dynamic programming we obtain the conditions characterizing an optimal solution of the problem (9)–(1) as in (10). □

Substituting the set of cooperative strategies into (1) yields the dynamics of the cooperative state trajectory in the time interval $[t_k,t_{k+1})$

$$\dot{x}(s)=f\left[s,x(s),\psi_{k-1}^{(\omega_{k-1},O)\omega_k}(s,x(s)),\psi_k^{(\omega_k,Y)\omega_{k-1}}(s,x(s))\right], \tag{11}$$

for $s\in[t_k,t_{k+1}),k\in\{1,2,\cdots,\upsilon\}$ and $x(t_k)=x_{t_k}\in X$.

Let $\{x^{(\omega_{k-1},\omega_k)*}(t)\}_{t=t_k}^{t_{k+1}}$ denote the cooperative solution path governed by (11). For simplicity in exposition, we denote $x^{(\omega_{k-1},\omega_k)*}(t)$ by $x_t^{(\omega_{k-1},\omega_k)*}$.

To fulfill group optimality, the imputation vectors have to satisfy:

$$\xi^{k-1(\omega_{k-1},O)\omega_k[h]}(t,x)+\xi^{k(\omega_k,Y)\omega_{k-1}[h]}(t,x)=W^{[t_k,t_{k+1}](\omega_{k-1},\omega_k)}(t,x)$$

for $t\in[t_k,t_{k+1})$, $\omega_k\in\{\omega_1,\omega_2,\cdots,\omega_\varsigma\}$, $\omega_{k-1}\in\{\omega_1,\omega_2,\cdots,\omega_\varsigma\}$ and $k\in\{0,1,2,\cdots,\upsilon\}$.

3.2 Individual Rationality

In a dynamic framework, individual rationality requires that the imputation received by a player has to be no less than his/her noncooperative payoff throughout the time interval in concern. Hence for individual rationality to hold along the cooperative trajectory $\{x^{(\omega_{k-1},\omega_k)^*}(t)\}_{t=t_k}^{t_{k+1}}$,

$$\xi^{k-1(\omega_{k-1},O)\omega_k[h]}(t,x_t^*) \geq V^{k-1(\omega_{k-1},O)\omega_k}(t,x_t^*) \text{ and}$$
$$\xi^{k(\omega_k,Y)\omega_{k-1}[h]}(t,x_t^*) \geq V^{k(\omega_k,Y)\omega_{k-1}}(t,x_t^*),$$

for $t \in [t_k,t_{k+1})$, $\omega_k \in \{\omega_1,\omega_2,\cdots,\omega_\varsigma\}$, $\omega_{k-1} \in \{\omega_1,\omega_2,\cdots,\omega_\varsigma\}$ and $k \in \{0,1,2,\cdots,\upsilon\}$, where x_t^* is the short form for $x_t^{(\omega_{k-1},\omega_k)^*}$.

Using the results derived, an imputation vector equally dividing the excess of the cooperative payoff over the noncooperative payoff can be expressed as:

$$\begin{aligned}
\xi^{k-1(\omega_{k-1},O)\omega_k[h]}(t,x_t^*) &= V^{k-1(\omega_{k-1},O)\omega_k[h]}(t,x_t^*) + 0.5\\
&\quad\times[W^{[t_k,t_{k+1}](\omega_{k-1},\omega_k)}(t,x_t^*)\\
&\qquad -V^{k-1(\omega_{k-1},O)\omega_k[h]}(t,x_t^*) - V^{k(\omega_k,Y)\omega_{k-1}}(t,x_t^*)], \text{ and}\\
\xi^{k(\omega_k,Y)\omega_{k-1}[h]}(t,x_t^*) &= V^{k(\omega_k,Y)\omega_{k-1}}(t,x_t^*) + 0.5\\
&\quad\times[W^{[t_k,t_{k+1}](\omega_{k-1},\omega_k)}(t,x_t^*)]\\
&\qquad -V^{k-1(\omega_{k-1},O)\omega_k[h]}(t,x_t^*) - V^{k(\omega_k,Y)\omega_{k-1}}(t,x_t^*).
\end{aligned} \tag{12}$$

One can readily see that the imputations in (12) satisfy individual rationality and group optimality.

4 Dynamically Consistent Solutions and Payoff Distribution

A stringent requirement for solutions of cooperative differential games to be dynamically stable is the property of subgame consistency. Under subgame consistency, 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 particular, when the game proceeds, at each instant of time the players are guided by the same optimality principles, and hence do not have any ground for deviation from the previously adopted optimal behavior throughout the game.

According to the solution optimality principle, the players agree to share their cooperative payoff according to the imputations

$$\left[\xi^{k-1(\omega_{k-1},O)\omega_k[h]}(t,x_t^*),\, \xi^{k(\omega_k,Y)\omega_{k-1}[h]}(t,x_t^*)\right] \tag{13}$$

over the time interval $[t_k,t_{k+1})$.

To achieve dynamic consistency, a payment scheme has to be derived so that imputation (13) will be maintained throughout the time interval $[t_k, t_{k+1})$. Following Yeung and Petrosyan [8,10], we formulate a payoff distribution procedure (PDP) over time so that the agreed imputations (13) can be realized. Let $B_{k-1}^{(\omega_{k-1},O)\omega_k[h]}(s)$ and $B_k^{(\omega_k,Y)\omega_{k-1}[h]}(s)$ denote the instantaneous payments at time $s \in [t_k, t_{k+1})$ allocated to the type ω_{k-1} generation $k-1$ (old) player and type ω_k generation k (young) player. In particular, the imputation vector can be expressed as:

$$
\begin{aligned}
&\xi^{k-1(\omega_{k-1},O)\omega_k[h]}(t, x_t^*) \\
&\quad = \int_{t_k}^{t_{k+1}} B_{k-1}^{(\omega_{k-1},O)\omega_k[h]}(s) e^{-r(s-t_k)} ds + e^{-r(t_{k+1}-t_k)} q^{k-1(\omega_{k-1})} \\
&\qquad [t_{k+1}, x^*(t_{k+1})], \\
&\xi^{k(\omega_k,Y)\omega_{k-1}[h]}(t, x_t^*)] = \int_{t_k}^{t_{k+1}} B_k^{(\omega_k,Y)\omega_{k-1}[h]}(s) e^{-r(s-t_k)} ds \\
&\quad + \sum_{\alpha=1}^{\varsigma} \sum_{h=1}^{\varsigma(\omega_k,\omega_\alpha)} \varpi_h^{(\omega_k,\omega_\alpha)} \xi^{k(\omega_k,O)\omega_\alpha[h]}(t_{k+1}, x^*(t_{k+1})),
\end{aligned}
$$

for $k \in \{1, 2, \cdots, \upsilon - 1\}$, and

$$
\begin{aligned}
&\xi^{\upsilon-1(\omega_{\upsilon-1},O)\omega_\upsilon[h]}(t, x_t^*) \\
&\quad = \int_{t_\upsilon}^{t_{\upsilon+1}} B_{\upsilon-1}^{(\omega_{\upsilon-1},O)\omega_\upsilon[h]}(s) e^{-r(s-t_\upsilon)} ds + e^{-r(t_{\upsilon+1}-t_\upsilon)} q^{\upsilon-1(\omega_{\upsilon-1})} \\
&\quad \times [t_{\upsilon+1}, x^*(t_{\upsilon+1})], \\
&\xi^{\upsilon(\omega_\upsilon,Y)\omega_{\upsilon-1}[h]}(t, x_t^*) \\
&\quad = \int_{t_\upsilon}^{t_{\upsilon+1}} B_\upsilon^{(\omega_\upsilon,Y)\omega_{\upsilon-1}[h]}(s) e^{-r(s-t_\upsilon)} ds + e^{-r(t_{\upsilon+1}-t_\upsilon)} q^{\upsilon(\omega_\upsilon)} [t_{\upsilon+1}, x^*(t_{\upsilon+1})].
\end{aligned}
$$

Using the analysis in Yeung and Petrosyan [10] and Petrosyan and Yeung [7], we obtain:

Theorem 2. *If the imputation vector* $[\xi^{k-1(\omega_{k-1},O)\omega_k[h]}(t, x_t^*), \xi^{k(\omega_k,O)\omega_{k-1}[h]}(t, x_t^*)]$ *are functions that are continuously differentiable in* t *and* x_t^**, a PDP with an instantaneous payment at time* $t \in [t_k, t_{k+1})$*:*

$$
\begin{aligned}
&B_{k-1}^{(\omega_{k-1},O)\omega_k[h]}(t) = -\xi_t^{k-1(\omega_{k-1},O)\omega_k[h]}(t, x_t^*) \\
&\quad - \xi_x^{k-1(\omega_{k-1},O)\omega_k[h]}(t, x_t^*) f \left[t, x_t^*, \psi_{k-1}^{(\omega_{k-1},O)\omega_k}(t, x_t^*), \psi_k^{(\omega_k,Y)\omega_{k-1}}(t, x_t^*) \right]
\end{aligned}
$$

allocated to the type ω_{k-1} generation $k-1$ player; and an instantaneous payment at time $t \in [t_k, t_{k+1})$:

$$B_k^{(\omega_k,Y)\omega_{k-1}[h]}(t) = -\xi_t^{k(\omega_k,Y)\omega_{k-1}[h]}(t, x_t^*)$$
$$- \xi_x^{k(\omega_k,Y)\omega_{k-1}[h]}(t, x_t^*) f\left[t, x_t^*, \psi_{k-1}^{(\omega_{k-1},O)\omega_k}(t, x_t^*), \psi_k^{(\omega_k,Y)\omega_{k-1}}(t, x_t^*)\right]$$

allocated to the type ω_k generation k player yields a mechanism leading to the realization of the imputation vector

$$\left[\xi^{k-1(\omega_{k-1},O)\omega_k[h]}(t, x_t^*), \xi^{k(\omega_k,Y)\omega_{k-1}[h]}(t, x_t^*)\right], \text{ for } k \in \{1, 2, \cdots, \upsilon\}.$$

Proof. Follow the proof leading to Theorem 4.4.1 in Yeung and Petrosyan [10] with the imputation vector in present value (rather than in current value). □

5 An Illustration in Resource Extraction

Consider the game in which there are 4 overlapping generations of players with generation 0 and generation 1 players in $[t_1, t_2)$, generation 1 and generation 2 players in $[t_2, t_3)$, generation 2 and generation 3 players in $[t_3, t_4]$. Players are of either type 1 or type 2. The instantaneous payoffs and terminal rewards of the type 1 generation k player and the type 2 generation k player are, respectively:

$$\left[(u_k)^{1/2} - \frac{c_1}{x^{1/2}} u_k\right] \text{ and } q_1 x^{1/2} \text{ and } \left[(u_k)^{1/2} - \frac{c_2}{x^{1/2}} u_k\right] \text{ and } q_2 x^{1/2}$$

At initial time t_1, it is known that the generation 0 player is of type 1 and the generation 1 player is of type 2. It is also known that the generation 2 and generation 3 players may be of type 1 with probability $\lambda_1 = 0.4$ and of type 2 with probability $\lambda_1 = 0.6$.
The state dynamics of the game is characterized by:

$$\dot{x}(s) = a x(s)^{1/2} - b x(s) - u_{k-1}(s) - u_k(s), \text{for } s \in [t_k, t_{k+1}) \text{and } k \in \{1, 2, 3\};$$
$$x(t_1) = x_0 \in X \subset R. \tag{14}$$

The game is an asynchronous horizons version of the synchronous-horizon resource extraction game in Yeung and Petrosyan [10]. The state variable $x(s)$is the biomass of a renewable resource. $u_k(s)$ is the harvest rate of the generation k extraction firm. The death rate of the resource is b. The rate of growth is $a/x^{1/2}$ which reflects the decline in the growth rate as the biomass increases. The type $i \in \{1, 2\}$ generation k extraction firm's extraction cost is $c_i u_k(s) x(s)^{-1/2}$.

This asynchronous horizon game can be expressed as follows. In the time interval $[t_k, t_{k+1})$, for $k \in \{1,2\}$, consider the case with a type $i \in \{1,2\}$ generation $k-1$ firm and a type $j \in \{1,2\}$ generation k firm, the game becomes

$$\max_{u_{k-1}} \left\{ \int_{t_k}^{t_{k+1}} \left[[u_{k-1}^{(i,O)j}(s)]^{1/2} - \frac{c_i}{x(s)^{1/2}} u_1^{(i,O)j}(s) \right] \exp[-r(s-t_k)] \mathrm{d}s \right.$$
$$\left. + \exp[-r(t_{k+1}-t_k)] q_i x(t_{k+1})^{\frac{1}{2}} \right\},$$
$$\max_{u_k} \left\{ \int_{t_k}^{t_{k+1}} \left[\left[u_k^{(j,Y)i}(s) \right]^{1/2} - \frac{c_j}{x(s)^{1/2}} u_2^{(j,Y)i}(s) \right] \exp[-r(s-t_k)] \mathrm{d}s \right.$$
$$+ \sum_{\alpha=1}^{2} \lambda_\alpha \int_{t_3}^{t_4} [[u_k^{(j,O)\alpha}(s)]^{1/2} - \frac{c_j}{x(s)^{1/2}} u_k^{(j,O)\alpha}(s)] \exp[-r(s-t_k)] \mathrm{d}s$$
$$\left. + \exp[-r(t_{k+2}-t_k)] q_j x(t_{k+2})^{\frac{1}{2}} \right\} \tag{15}$$

subject to (14).

In the time interval $[t_3, t_4]$, consider the case with a type $i \in \{1,2\}$ generation 2 firm and a type $j \in \{1,2\}$ generation 3 firm, the game becomes

$$\max_{u_2} \left\{ \int_{t_3}^{t_4} \left[\left[u_2^{(i,O)j}(s) \right]^{1/2} - \frac{c_i}{x(s)^{1/2}} u_2^{(i,O)j}(s) \right] \exp[-r(s-t_3)] \, \mathrm{d}s \right.$$
$$\left. + \exp[-r(t_4-t_3)] \, q_i x(t_4)^{\frac{1}{2}} \right\},$$
$$\max_{u_3} \left\{ \int_{t_3}^{t_4} \left[\left[u_3^{(j,O)i}(s) \right]^{1/2} - \frac{c_j}{x(s)^{1/2}} u_2^{(j,O)i}(s) \right] \exp[-r(s-t_3)] \, \mathrm{d}s \right.$$
$$\left. + \exp[-r(t_4-t_3)] \, q_j x(t_4)^{\frac{1}{2}} \right\}, \tag{16}$$

subject to (14).

5.1 Noncooperative Outcomes

In this section, we characterize the noncooperative outcome of the asynchronous horizons game (14)–(16).

Proposition 1. *The value functions for the type $i \in \{1,2\}$ generation $k-1$ firm and the type $j \in \{1,2\}$ generation k firm coexisting in the game interval $[t_k, t_{k+1})$ can be obtained as:*

$$V^{k-1(i,O)j}(t,x) = \exp[-r(t-t_k)] \left[A_{k-1}^{(i,O)j}(t) x^{1/2} + C_{k-1}^{(i,O)j}(t) \right], \text{ and}$$
$$V^{k(j,Y)i}(t,x) = \exp[-r(t-t_k)] \left[A_k^{(j,Y)i}(t) x^{1/2} + C_k^{(j,Y)i}(t) \right], \tag{17}$$

for $k \in \{1,2,3\}$ *and* $i, j \in \{1,2\}$, *where* $A_{k-1}^{(i,O)j}(t)$, $C_{k-1}^{(i,O)j}(t)$, $A_k^{(j,Y)i}(t)$ *and* $C_k^{(j,Y)i}(t)$ *satisfy:*

$$\dot{A}_{k-1}^{(i,O)j}(t) = \left[r + \frac{b}{2}\right] A_{k-1}^{(i,O)j}(t) - \frac{1}{2\left[c_i + A_{k-1}^{(i,O)j}(t)/2\right]} + \frac{c_i}{4\left[c_i + A_{k-1}^{(i,O)j}(t)/2\right]^2}$$
$$+ \frac{A_{k-1}^{(i,O)j}(t)}{8\left[c_i + A_{k-1}^{(i,O)j}(t)/2\right]^2} + \frac{A_{k-1}^{(i,O)j}(t)}{8\left[c_j + A_k^{(j,Y)i}(t)/2\right]^2}$$

$$\dot{C}_{k-1}^{(i,O)j}(t) = rC_{k-1}^{(i,O)j}(t) - \frac{a}{2}A_{k-1}^{(i,O)j}(t),$$

$$A_{k-1}^{(i,O)j}(t_{k+1}) = q_i \ \ \text{and} \ \ C_{k-1}^{(i,O)j}(t_{k+1}) = 0, \ \text{for}\ k \in \{1,2,3\}; \tag{18}$$

$$\dot{A}_k^{(j,Y)i}(t) = \left[r + \frac{b}{2}\right] A_k^{(j,Y)i}(t) - \frac{1}{2\left[c_j + A_k^{(j,Y)i}(t)/2\right]} + \frac{c_j}{4\left[c_j + A_k^{(j,Y)i}(t)/2\right]^2}$$
$$+ \frac{A_k^{(j,Y)i}(t)}{8\left[c_j + A_k^{(j,Y)i}(t)/2\right]^2} + \frac{A_k^{(j,Y)i}(t)}{8\left[c_i + A_{k-1}^{(i,O)j}(t)/2\right]^2}$$

$$\dot{C}_k^{(j,Y)i}(t) = rC_k^{(j,Y)i}(t) - \frac{a}{2}A_k^{(j,Y)i}(t), \ \text{for}\ k \in \{1,2,3\};$$

$$A_k^{(j,Y)i}(t_{k+1}) = e^{-r(t_{k+1}-t_k)} \sum_{l=1}^{2} \lambda_l A_k^{(j,O)l}(t_{k+1}) \ \text{and}$$

$$C_k^{(j,Y)i}(t_{k+1}) = e^{-r(t_{k+1}-t_k)} \sum_{l=1}^{2} \lambda_l C_k^{(j,O)l}(t_{k+1}),$$

$$\text{for} k \in \{1,2\}, \ \text{and}\ A_3^{(j,Y)i}(t_4) = q_j \ \text{and}\ C_3^{(j,Y)i}(t_4) = 0 \tag{19}$$

Proof. Using Lemmas 1 and 2 and the analysis in Proposition 4.1.1 in Yeung and Petrosyan [10], one can obtain the value functions in (17). □

The solution time paths $A_{k-1}^{(i,O)j}(t)$, $C_{k-1}^{(i,O)j}(t)$, $A_k^{(j,Y)i}(t)$ and $C_k^{(j,Y)i}(t)$ for the system of first-order differential equations in (18)–(19) can be computed numerically for given values of the model parameters $r, q_1, q_2, c_1, c_2, a, b, \lambda_1$, and λ_2.

Following Yeung and Petrosyan [10], the game equilibrium strategies can be expressed as: $\phi_{k-1}^{(i,O)j}(t,x) = \frac{x}{4[c_i + A_{k-1}^{(i,O)j}(t)/2]^2}$ and $\phi_k^{(j,Y)i}(t,x) = \frac{x}{4[c_j + A_k^{(j,Y)i}(t)/2]^2}$.

5.2 Dynamic Cooperation

Now consider the case when coexisting firms want to cooperate and agree to act and allocate the cooperative payoff according to a set of agreed-upon optimality principles. Let there be three acceptable imputations.

Imputation I: the firms would share the excess gain from cooperation equally with weights $w_{k-1}^1 = w_k^1 = 0.5$.

Imputation II: the generation $k-1$ firm acquires $w_{k-1}^2 = 0.6$ of the excess gain from cooperation and the generation k firm acquires $w_k^2 = 0.4$ of the gain.

Imputation III: the generation $k-1$ firm acquires $w_{k-1}^3 = 0.4$ of the excess gain from cooperation and the generation k firm acquires $w_k^3 = 0.6$ of the gain.

In time interval $[t_k, t_{k+1})$, if both the generation $k-1$ firm and the generation k firm are of type 1, the probabilities that the firms would agree to Imputations I, II, and III are, respectively, $\varpi_1^{(1,1)} = 0.8$, $\varpi_2^{(1,1)} = 0.1$ and $\varpi_3^{(1,1)} = 0.1$.

If both the generation $k-1$ firm and the generation k firm are of type 2, the probabilities that the firms would agree to Imputations I, II, and III are, respectively, $\varpi_1^{(2,2)} = 0.7$, $\varpi_2^{(2,2)} = 0.15$ and $\varpi_3^{(2,2)} = 0.15$.

If the generation $k-1$ firm is of type 1 and the generation k firm are of type 2, the probabilities that the firms would agree to Imputations I, II, and III are, respectively, $\varpi_1^{(1,2)} = 0.15$, $\varpi_2^{(1,2)} = 0.75$ and $\varpi_3^{(1,2)} = 0.1$.

If the generation $k-1$ firm is of type 2 and the generation k firm are of type 1, the probabilities that the firms would agree to Imputations I, II, and III are, respectively, $\varpi_1^{(2,1)} = 0.15$, $\varpi_2^{(2,1)} = 0.1$ and $\varpi_3^{(2,1)} = 0.75$.

At initial time t_1, the type 1 generation 0 firm and the type 2 generation 1 firm are assumed to have agreed to Imputation II.

Since payoffs are transferable, group optimality requires the firms coexisting in the same time interval to maximize their joint payoff. Consider the last time interval $[t_3, t_4]$, in which the generation 2 firm is of type $i \in \{1, 2\}$ and the generation 3 firm is of type $j \in \{1, 2\}$. The firms maximize their joint profit

$$\left\{ \int_{t_3}^{t_4} \left[\left[u_2^{(i,O)j}(s) \right]^{1/2} - \frac{c_i}{x(s)^{1/2}} u_2^{(i,O)j}(s) \right] \exp\left[-r(s - t_3)\right] \mathrm{d}s \right.$$
$$+ \int_{t_3}^{t_4} \left[\left[u_3^{(j,O)i}(s) \right]^{1/2} - \frac{c_j}{x(s)^{1/2}} u_2^{(j,O)i}(s) \right] \exp\left[-r(s - t_3)\right] \mathrm{d}s$$
$$\left. + \exp\left[-r(t_4 - t_3)\right] q_i x(t_4)^{1/2} + \exp\left[-r(t_4 - t_3)\right] q_j x(t_4)^{1/2} \right\},$$

subject to (14).

Proposition 2. *The maximized joint payoff with type $i \in \{1, 2\}$ generation 2 firm and the type $j \in \{1, 2\}$ generation 3 firm coexisting in the game interval $[t_3, t_4)$ can be obtained as:*

$$W^{[t_3,t_4](i,j)}(t, x) = \exp\left[-r(t - t_3)\right] \left[A^{[t_3,t_4](i,j)}(t) x^{1/2} + C^{[t_3,t_4](i,j)}(t) \right], \quad (20)$$

where $A^{[t_3,t_4](i,j)}(t)$ and $C^{[t_3,t_4](i,j)}(t)$ satisfy:

$$\dot{A}^{[t_3,t_4](i,j)}(t) = \left[r + \frac{b}{2}\right] A^{[t_3,t_4](i,j)}(t) - \frac{1}{2\left[c_i + A^{[t_3,t_4](i,j)}(t)/2\right]}$$
$$-\frac{1}{2\left[c_j + A^{[t_3,t_4](i,j)}(t)/2\right]} + \frac{c_i}{4\left[c_i + A^{[t_3,t_4](i,j)}(t)/2\right]^2}$$
$$+\frac{c_j}{4\left[c_j + A^{[t_3,t_4](i,j)}(t)/2\right]^2}$$
$$+\frac{A^{[t_3,t_4](i,j)}(t)}{8\left[c_i + A^{[t_3,t_4](i,j)}(t)/2\right]^2} + \frac{A^{[t_3,t_4](i,j)}(t)}{8\left[c_j + A^{[t_3,t_4](i,j)}(t)/2\right]^2},$$
$$\dot{C}^{[t_3,t_4](i,j)}(t) = rC^{[t_3,t_4](i,j)}(t) - \frac{a}{2}A^{[t_3,t_4](i,j)}(t),$$
$$A^{[t_3,t_4](i,j)}(t_4) = q_i + q_j \quad \text{and} \quad C^{[t_3,t_4](i,j)}(t_4) = 0. \tag{21}$$

Proof. Using Lemma 3 and the analysis in example 4.2.1 in Yeung and Petrosyan [10], one can obtain (20)–(21). □

The solution time paths $A^{[t_3,t_4](i,j)}(t)$ and $C^{[t_3,t_4](i,j)}(t)$ for the system of first-order differential equations in (20)–(21) can be computed numerically for given values of the model parameters r, q_1, q_2, c_1, c_2, a and b.

In the game interval $[t_3, t_4)$ if type $i \in \{1, 2\}$ generation 2 firm and the type $j \in \{1, 2\}$ generation 3 firm coexisting, the imputations of the firms under cooperation can be expressed as:

$$\xi^{2(i,O)j[h]}(t,x)$$
$$= V^{2(i,O)j}(t,x) + w_2^h\left[W^{[t_3,t_4](i,j)}(t,x) - V^{2(i,O)j}(t,x) - V^{3(j,Y)i}(t,x)\right],$$
$$\xi^{3(j,Y)i[h]}(t,x)$$
$$= V^{3(j,Y)i}(t,x) + w_3^h\left[W^{[t_3,t_4](i,j)}(t,x) - V^{2(i,O)j}(t,x) - V^{3(j,Y)i}(t,x)\right],$$
$$\text{for } h \in \{1,2,3\}. \tag{22}$$

Now we proceed to the second last interval $[t_k, t_{k+1})$ for $k = 2$. Consider the case the case in which the generation k firm is of type $j \in \{1, 2\}$ and the generation $k-1$ firm is known to be of type $i = 2$. Following the analysis in (7) and (8), the expected terminal reward to the type j generation k firm at time t_{k+1} can be expressed as:

$$\sum_{l=1}^{2} \lambda_l \sum_{h=1}^{3} \varpi_h^{(j,l)} \xi^{k(j,O)l[h]}(t_{k+1}, x), \text{ for } k = 2. \tag{23}$$

A review of Proposition 1, Proposition 2 and (22) shows the term in (23) can be written as:

$$A_k^{\xi(j,O)} x^{1/2} + C_k^{\xi(j,O)},$$

where $A_k^{\xi(j,O)}$ and $C_k^{\xi(j,O)}$ are constant terms.

The joint maximization problem in the time interval $[t_k, t_{k+1})$, for $k \in \{1,2\}$, involving the type j generation k player and type i generation $k-1$ player can be expressed as:

$$\begin{aligned}\max_{u_{k-1},u_k} \Bigg\{ & \int_{t_k}^{t_{k+1}} \left[\left[u_{k-1}^{(i,O)j}(s)\right]^{1/2} - \frac{c_i}{x(s)^{1/2}} u_{k-1}^{(i,O)j}(s)\right] \exp[-r(s-t_k)] \mathrm{d}s \\ & + \int_{t_3}^{t_4} \left[\left[u_k^{(j,O)i}(s)\right]^{1/2} - \frac{c_j}{x(s)^{1/2}} u_k^{(j,O)i}(s)\right] \exp[-r(s-t_k)] \mathrm{d}s \\ & + \exp[-r(t_{k+1}-t_k)] \left[q_i x(t_{k+1})^{1/2} + A_k^{\xi(j,O)} x(t_{k+1})^{1/2} + C_k^{\xi(j,O)}\right] \Bigg\},\end{aligned}$$

subject to (14).

Proposition 3. *The maximized joint payoff with type $i \in \{1,2\}$ generation $k-1$ firm and the type $j \in \{1,2\}$ generation k firm coexisting in the game interval $[t_k, t_{k+1})$, for $k \in \{1,2\}$ can be obtained as:*

$$W^{[t_k,t_{k+1}](i,j)}(t,x) = \exp[-r(t-t_k)] \left[A^{[t_k,t_{k+1}](i,j)}(t)\, x^{1/2} + C^{[t_k,t_{k+1}](i,j)}(t)\right], \tag{24}$$

where $A^{[t_k,t_{k+1}](i,j)}(t)$ and $C^{[t_k,t_{k+1}](i,j)}(t)$ satisfy:

$$\begin{aligned}\dot{A}^{[t_k,t_{k+1}](i,j)}(t) &= \left[r + \frac{b}{2}\right] A^{[t_k,t_{k+1}](i,j)}(t) \\ & - \frac{1}{2\left[c_i + A^{[t_k,t_{k+1}](i,j)}(t)/2\right]} - \frac{1}{2\left[c_j + A^{[t_k,t_{k+1}](i,j)}(t)/2\right]} \\ & + \frac{c_i}{4\left[c_i + A^{[t_k,t_{k+1}](i,j)}(t)/2\right]^2} + \frac{c_j}{4\left[c_j + A^{[t_k,t_{k+1}](i,j)}(t)/2\right]^2} \\ & + \frac{A^{[t_k,t_{k+1}](i,j)}(t)}{8\left[c_i + A^{[t_k,t_{k+1}](i,j)}(t)/2\right]^2} + \frac{A^{[t_k,t_{k+1}](i,j)}(t)}{8\left[c_j + A^{[t_k,t_{k+1}](i,j)}(t)/2\right]^2}, \\ \dot{C}^{[t_k,t_{k+1}](i,j)}(t) &= r C^{[t_k,t_{k+1}](i,j)}(t) - \frac{a}{2} A^{[t_k,t_{k+1}](i,j)}(t), \\ A^{[t_k,t_{k+1}](i,j)}(t_{k+1}) &= q_i + A_k^{\xi(j,O)} \text{ and } C^{[t_k,t_{k+1}](i,j)}(t_{k+1}) = C_k^{\xi(j,O)}.\end{aligned} \tag{25}$$

Proof. Using Theorem 1 and the analysis in example 4.2.1 in Yeung and Petrosyan [10], one can obtain the results in (24) and (25). □

The solution time paths $A^{[t_k,t_{k+1}](i,j)}(t)$ and $C^{[t_k,t_{k+1}](i,j)}(t)$ for the system of first-order differential equations in (24)–(25) can be computed numerically for given values of the model parameters r,q_1, q_2, c_1, c_2, a, b, λ_1, λ_2, and $\varpi_h^{(j,l)}$ for $h \in \{1,2,3\}$ and $j,l \in \{1,2\}$.

Following Yeung and Petrosyan [10], the optimal cooperative controls can then be obtained as:

$$\psi_{k-1}^{(i,O)j}(t,x) = \frac{x}{4\left[c_i + A^{[t_k,t_{k+1}](i,j)}(t)/2\right]^2}, \quad \text{and}$$
$$\psi_k^{(j,Y)i}(t,x) = \frac{x}{4\left[c_j + A^{[t_k,t_{k+1}](i,j)}(t)/2\right]^2}.$$

Substituting these control strategies into (14) yields the dynamics of the state trajectory under cooperation. The optimal cooperative state trajectory in the time interval $[t_k, t_{k+1})$ can be obtained as:

$$x^{(i,j)*}(t) = \left[\Omega_{(i,j)}(t_k,t)\right]^2 \left[\left(x_{t_k}\right)^{1/2} + \int_{t_k}^{t} \Omega_{(i,j)}^{-1}(t_k,s)\frac{a}{2}\mathrm{d}s\right]^2,$$

where $\Omega_{(i,j)}(t_k,t) = \exp\left[\int_k^l H_{(i,j)}(\upsilon)\mathrm{d}\upsilon\right]$ and

$$H_{(i,j)}(s) = -\left[\frac{b}{2} + \frac{1}{8\left[c_i + A^{[t_k,t_{k+1}](i,j)}(s)/2\right]^2} + \frac{1}{8\left[c_j + A^{[t_k,t_{k+1}](i,j)}(s)/2\right]^2}\right].$$

The term x_t^* is used to denote $x^{(i,j)*}(t)$ whenever there is no ambiguity.

5.3 *Dynamically Consistent Payoff Distribution*

According to the solution optimality principle, the players agree to share their cooperative payoff according to the solution imputations:

$$\xi^{k-1(i,O)j[h]}(t,x) = V^{k-1(i,O)j}(t,x) + w_{k-1}^h$$
$$\left[W^{[t_k,t_{k+1}](i,j)}(t,x) - V^{k-1(i,O)j}(t,x) - V^{k(j,Y)i}(t,x)\right],$$
$$\xi^{k(j,Y)i[h]}(t,x) = V^{k(j,Y)i}(t,x) + w_k^h$$
$$\left[W^{[t_k,t_{k+1}](i,j)}(t,x) - V^{k-1(i,O)j}(t,x) - V^{k(j,Y)i}(t,x)\right],$$
$$\text{for } h \in \{1,2,3\}, \ i,j \in \{1,2\} \text{ and } k \in \{1,2,3\}.$$

These imputations are continuous differentiable in x and t. If an imputation vector $\left[\xi^{k-1(i,O)j[h]}(t,x), \xi^{k(j,Y)i[h]}(t,x)\right]$ is chosen, a crucial process is to derive a payoff distribution procedure (PDP) so that this imputation could be realized for $t \in [t_k, t_{k+1})$ along the cooperative trajectory $\{x_t^*\}_{t=t_k}^{t_{k+1}}$.

Following Theorem 2, a PDP leading to the realization of the imputation vector $\left[\xi^{k-1(i,O)j[h]}(t,x), \xi^{k(j,Y)i[h]}(t,x)\right]$ can be obtained as:

Corollary 1. *A PDP with an instantaneous payment at time $t \in [t_k, t_{k+1})$:*

$$B_{k-1}^{(i,O)j[h]}(t) = -\xi_t^{k-1(i,O)j[h]}\left(t, x_t^*\right) - \xi_x^{k-1(i,O)j[h]}\left(t, x_t^*\right)\left[a(x_t^*)^{1/2} - bx_t^* - \frac{x_t^*}{4\left[c_i + A^{[t_k,t_{k+1}](i,j)}(t)/2\right]^2} - \frac{x_t^*}{4\left[c_j + A^{[t_k,t_{k+1}](i,j)}(t)/2\right]^2}\right], \tag{26}$$

allocated to the type i generation $k-1$ player; and an instantaneous payment at time $t \in [t_k, t_{k+1})$:

$$B_{k}^{(j,Y)i[h]}(t) = -\xi_t^{k(j,Y)i[h]}\left(t, x_t^*\right) - \xi_x^{k(j,Y)i[h]}\left(t, x_t^*\right)\left[a(x_t^*)^{1/2} - bx_t^* - \frac{x_t^*}{4\left[c_i + A^{[t_k,t_{k+1}](i,j)}(t)/2\right]^2} - \frac{x_t^*}{4\left[c_j + A^{[t_k,t_{k+1}](i,j)}(t)/2\right]^2}\right] \tag{27}$$

allocated to the type j generation k player, yields a mechanism leading to the realization of the imputation vector

$$\left[\xi^{k-1(i,O)j[h]}(t,x), \xi^{k(j,Y)i[h]}(t,x)\right], \textit{ for } k \in \{1,2,3\},$$
$$h \in \{1,2,3\} \textit{ and } i, j \in \{1,2\}.$$

Since the imputations $\xi^{k-1(i,O)j[h]}(t,x)$ and $\xi^{k(j,Y)i[h]}(t,x)$ are in terms of explicit differentiable functions, the relevant derivatives in Corollary 1 can be derived using the results in Propositions 1, 2, and 3. Hence, the PDP $B_{k-1}^{(i,O)j[h]}(t)$ and $B_k^{(j,Y)i[h]}(t)$ in (26) and (27) can be obtained explicitly.

6 Concluding Remarks and Extensions

This paper considers cooperative differential games in which players enter the game at different times and have diverse horizons. Moreover, the types of future players are not known with certainty. Dynamically consistent cooperative solutions and analytically tractable payoff distribution mechanisms leading to the realization of these solutions are derived. The asynchronous horizons game presented can be extended in a couple of directions. First, more complicated stochastic processes can be adopted in the analysis. For instance, the random variable governing the types of future players can be a series of nonidentical random variables $\omega_{a_k}^k \in \left\{\omega_1^k, \omega_2^k, \cdots, \omega_{\zeta_k}^k\right\}$ with probabilities $\lambda_{a_k}^k \in \left\{\lambda_1^k, \lambda_2^k, \cdots, \lambda_{\zeta_k}^k\right\}$, for $k \in \{2, 3, \cdots, \upsilon\}$.

Second, the overlapping generations of players can be extended to more complex structures. The game horizon of the players can include more than two time intervals and be different across players. The number of players in each time interval can also be more than two and be different across intervals. The analysis can be formulated as a general class of differential games with asynchronous horizons structure. In particular, the type ω_{a_k} generation k player's game horizon is $\left[t_k, t_{k+\eta_k}\right)$, where $\eta_k \geq 1$. The term $u_k^{(\omega_k, S1)}(s)$ is used to denote the vector of controls of the type ω_{a_k} generation k player in his/her first game interval $[t_k, t_{k+1})$; and $u_k^{(\omega_k, S2)}(s)$ is that in his/her second game interval $[t_{k+1}, t_{k+2})$ and so on. This results in a general class of differential games with asynchronous horizons structure. Theorems 1 and 2 can be readily extended to this general structure with more than two players in each time interval.

Finally, this is the first time that dynamically consistent cooperative solutions are analyzed and derived in differential games with asynchronous players' horizons, further research along this line is expected.

Acknowledgements Financial supports by the HKSAR UGC Grant (CERG-202807), HKBU Strategic Development Fund (03-17-224) and the EU TOCSIN Project Grants are gratefully acknowledged.

References

1. Basar, T., Olsder, G.J.: Dynamic Noncooperative Game Theory. SIAM Edition, SIAM's Classics in Applied mathematics, Society for Industrial and Applied Mathematics, Philadelphia (1999)
2. Bellman, R.: Dynamic Programming. Princeton University Press, Princeton, NJ (1957)
3. Haurie, A.: A Note on Nonzero-Sum Differential Games with Bargaining Solutions. Journal of Optimization Theory and Applications **18**, 31–39 (1976)
4. Jørgensen, S., Yeung, D.W.K.: Overlapping Generations Stochastic Differential Games. Automatica **41**, 69–74 (2005)
5. Petrosyan, L.A.: Agreeable Solutions in Differential Games. International Journal of Mathematics, Game Theory and Algebra **7**, 165–177 (1997)
6. Petrosyan, L.A., Danilov, N.N.: Cooperative Differential Games and Their Applications. Izd. Tomskogo University, Tomsk (1982)
7. Petrosyan, L.A., Yeung, D.W.K.: Subgame-consistent Cooperative Solutions in Randomly-furcating Stochastic Differential Games. International Journal of Mathematical and Computer Modelling (Special Issue on Lyapunov's Methods in Stability and Control) **45**, 1294–1307 (2007)

8. Yeung, D.W.K., Petrosyan, L.A.: Subgame Consistent Cooperative Solutions in Stochastic Differential Games. Journal of Optimization Theory and Applications **120**, 651–666 (2004)
9. Yeung, D.W.K., Petrosyan, L.A.: Subgame Consistent Solution of A Cooperative Stochastic Differential Game with Nontransferable Payoffs. Journal of Optimization Theory and Applications **124**, 701–724 (2005)
10. Yeung, D.W.K., Petrosyan, L.A.: Cooperative Stochastic Differential Games. Springer, New York (2006)

Part V
Applications and Numerical Approaches

1. *Alberto Bressan and Wen Shen*, Measure-Valued Solutions to a Harvesting Game with Several Players
2. *Vivek S. Borkar and Joy Kuri*, Optimal Distributed Uplink Channel Allocation: A Constrained MDP Formulation
3. *Yan Dolinsky and Yuri Kifer*, Binomial Approximations for Barrier Options of Israeli Style
4. *Laurent Drouet, Alain Haurie, Jean-Philippe Vial and Marc Vielle*, A Game of International Climate Policy Solved by a Homogeneous Oracle-Based Method for Variational Inequalities
5. *Elżbieta Z. Ferenstein and Adam Pasternak-Winiarski*, Optimal Stopping of a Risk Process with Disruption and Interest Rates
6. *Mikhail Krastanov and Rossen Rozenov*, Examples in Dynamic Optimal Taxation
7. *Heinz-Uwe Kuënle*, On a Discounted Inventory Game
8. *Yulia Pavlova*, Time-Consistent Emission Reduction in a Dynamic Leader-Follower Game
9. *Anna Rettieva*, Fish Wars with Changing Area for a Fishery

Measure-Valued Solutions to a Harvesting Game with Several Players

Alberto Bressan and Wen Shen

Abstract We consider Nash equilibrium solutions to a harvesting game in one-space dimension. At the equilibrium configuration, the population density is described by a second-order O.D.E. accounting for diffusion, reproduction, and harvesting. The optimization problem corresponds to a cost functional having sublinear growth, and the solutions in general can be found only within a space of measures. In this chapter, we derive necessary conditions for optimality, and provide an example where the optimal harvesting rate is indeed measure valued. We then consider the case of many players, each with the same payoff. As the number of players approaches infinity, we show that the population density approaches a well-defined limit, characterized as the solution of a variational inequality. In the last section, we consider the problem of optimally designing a marine park, where no harvesting is allowed, so that the total catch is maximized.

1 Introduction

We consider the noncooperative harvesting game introduced in [8]. Let $\phi(x)$ denote the density of a fish population, or some other marine resource, at the location x. As "players" we consider N fishing companies, whose strategies are described by measures $\mu_1, \ldots, \mu_N$. Here, μ_i describes the intensity of harvesting effort by the i-th player. In one space dimension, a steady-state configuration is characterized as the solution to the two-point boundary value problem

$$\phi'' + g(x, \phi) = \sum_{i=1}^{N} \phi \, \mu_i, \tag{1}$$

A. Bressan (✉) and W. Shen
Department of Mathematics, Penn State University, University Park, PA 16802, USA
e-mail: bressan@math.psu.edu; shen_w@math.psu.edu

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_20,

with boundary conditions

$$\phi'(0) \;=\; \phi'(R) \;=\; 0, \tag{2}$$

where primes denote derivatives w.r.t. the space variable x. The first term accounts for diffusion, the nonlinear function g describes population growth, while $\phi\mu_i$ is the amount of fish harvested by the i-th fishing company. In general, $\mu_1, \ldots, \mu_N$ are positive Radon measures supported on the closed interval $[0, R]$. Notice that the conditions (2) imply that no flux occurs across the boundary.

The goal of the i-th player is to maximize his/her net payoff

$$J_i \;=\; \int_0^R (\phi - c_i(x))\, \mathrm{d}\mu_i\,,$$

where $c_i(\cdot)$ is a strictly positive function, accounting for the harvesting cost.

Under suitable assumptions, the existence of solutions to the noncooperative game was proved in [8], within the class of nonnegative Radon measures. In this chapter, we derive a set of necessary conditions satisfied by these solutions. Since the domain is one-dimensional, we can use a variable transformation that transforms the optimization problem into a standard optimal control problem, to which the Pontryagin maximum principle can then be applied. Our results are formulated, more generally, in the presence of a drift coefficient, and with a variable diffusion coefficient.

A couple of examples are worked out in more details. In particular, we show that if the cost function is discontinuous, then the optimal solution can be a measure containing Dirac masses. This is indeed the case when a marine park is present, i.e., there is an open subset $\P \subset [0, R]$ such that $c_i(x) = +\infty$ for all $x \in \P$.

We then study the density of the fish population in the case of a large number of fishermen, each with the same harvesting cost $c_i(x) \;=\; c(x)$. Calling ϕ_N the population density corresponding to a Nash equilibrium solution with N fishermen, as $N \to \infty$ we prove the convergence $\phi_N \to \phi_\infty$, where ϕ_∞ provides the largest subsolution to the boundary value problem $\phi'' + g(x, \phi) = 0$, $\phi'(0) = \phi'(R) = 0$ satisfying the pointwise constraint $\phi(x) \leq c(x)$. Equivalently, ϕ_∞ can also be characterized as the unique solution to a variational inequality. Indeed, consider the family of functions

$$\begin{aligned}\mathscr{K}_c \doteq \Big\{\phi : [0, R] \mapsto I\!R, \quad &\phi \text{ is Lipschitz continuous with } \phi' \in BV, \\ &\phi'(0+) \geq 0, \quad \phi'(R-) \leq 0, \quad \phi(x) \leq c(x) \forall x \in [0, R]\Big\},\end{aligned} \tag{3}$$

where the lower semicontinuous function $c(\cdot)$ plays the role of an obstacle. Then the limiting density ϕ_∞ of the fish population satisfies $\phi_\infty \in \mathscr{K}_c$ and

$$\int_0^R \phi'_\infty \left(\psi' - \phi'_\infty\right) \mathrm{d}x + \int_0^R g(x, \phi_\infty) \cdot (\phi_\infty - \psi)\, \mathrm{d}x \;\geq\; 0 \qquad \forall \psi \in \mathscr{K}_c. \tag{4}$$

Roughly speaking, our results show that when the number of players becomes large, in a noncooperative game each single individual has no incentive to care for the environment: each player keeps harvesting until the fish population is so low that the profit is completely offset by the harvesting cost, at each point of the domain.

In the last section of the chapter, we consider the problem of the optimal designing of a marine park. Namely, we seek an open subset $\P \subset [0, R]$ such that, imposing the new harvesting cost

$$c^{(\P)}(x) = \begin{cases} c(x) & \text{if } x \in [0, R] \setminus \P, \\ +\infty & \text{if } x \in \P, \end{cases}$$

the corresponding solution $\phi^{(\P)}$ to the variational inequality (3)–(4) maximizes the total harvest, measured by

$$H(\P) \doteq \int_0^R g\left(x, \phi^{(\P)}(x)\right) \mathrm{d}x. \tag{5}$$

We prove that, if the initial cost function $c(\cdot)$ is continuous, then there exists an open set $\P \subset [0, R]$ for which the quantity in (5) is maximized. An example is explicitly worked out.

For more general results on the optimization of variational inequalities, we refer to [1, 2, 4–7].

2 Necessary Conditions for Measure-Valued Optimal Solutions

Given nonnegative cost function $c : [0, R] \mapsto I\!R_+ \cup \{+\infty\}$, we consider the optimal control problem

$$\text{maximize: } J(\sigma) \doteq \int_0^R \big(\phi(x) - c(x)\big) \, \mathrm{d}\sigma. \tag{6}$$

The maximum is sought over all pairs (σ, ϕ), where $\sigma \in \mathscr{M}_+\big([0, R]\big)$ is a nonnegative Radon measure on the closed interval $[0, R]$ and $\phi : [0, R] \mapsto I\!R_+$ provides a distributional solution to the second order boundary value problem

$$\big(\alpha(x)\phi'\big)' - \big(\beta(x)\phi\big)' + g(x, \phi) - \phi\nu = \phi\,\sigma, \tag{7}$$

with boundary conditions

$$\alpha(0)\phi'(0) - \beta(0)\phi(0) \;=\; \alpha(R)\phi'(R) - \beta(R)\phi(R) \;=\; 0. \tag{8}$$

Here and in the sequel, a prime denotes differentiation w.r.t. x. In connection with a noncooperative game, (6)–(8) describe the optimization problem faced by one of the players. The measure ν accounts for the combined fishing effort of all the other players. Notice that (7) describes a more general situation than (1), because we allow here a nonconstant diffusion coefficient, and the presence of a drift term. The boundary condition (8) formally implies that the flux through the boundary vanishes.

Denoting by f_ϕ the partial derivative of a function $f = f(x,\phi)$ w.r.t. ϕ, our basic assumptions will be

(A1) The functions α, β are of class $\mathscr{C}^1$, with $\alpha(x) > 0$ for all $x \in [0, R]$, while $\nu \in \mathscr{M}_+\big([0, R]\big)$ is a given, nonnegative Radon measure. The cost function c is lower semicontinuous and strictly positive, namely $c(x) \geq c_0 > 0$ for all $x \in [0, R]$.

(A2) The source term g can be written in the form $g(x,\phi) = f(x,\phi)\,\phi$, where the function $f = f(x,\phi)$ is continuous w.r.t. both variables and twice continuously differentiable w.r.t. ϕ. Moreover, for some continuous function $h = h(x)$ one has

$$f(x,0) > 0, \qquad f_\phi(x,\phi) < 0, \qquad f\big(x,h(x)\big) = 0 \quad \text{for all } \ x \in [0, R], \ \ \phi \geq 0. \tag{9}$$

One can think of $h(x)$ as the maximum population density supported by the environment at the location x. Since we need to consider not only classical solution of (7)–(8) but, more generally, measure-valued solutions, a precise definition is needed.

Definition 1. By a solution of the boundary value problem (7)–(8), we mean a Lipschitz continuous map $x \mapsto \phi(x)$ such that

(i) The map $x \mapsto \phi'(x)$ has bounded variation and satisfies

$$\begin{cases} \alpha(0)\phi'(0+) - \Big[\beta(0) + \nu\big(\{0\}\big) + \sigma\big(\{0\}\big)\Big]\phi(0) & = 0, \\ \alpha(R)\phi'(R-) - \Big[\beta(R) - \nu\big(\{R\}\big) - \sigma\big(\{R\}\big)\Big]\phi(R) & = 0. \end{cases}$$

(ii) For every test function $\eta \in \mathscr{C}^1_c\big(]0, R[\big)$, in the space of continuously differentiable functions whose support is a compact subset of the open interval $]0, R[$, one has

$$\int_0^R \Big\{ -\alpha\phi'\eta' + \beta\phi\eta' + g(x,\phi)\,\eta \Big\}\,\mathrm{d}x - \int_0^R \phi\,\eta\,(\mathrm{d}\nu + \mathrm{d}\sigma) = 0.$$

If the measures ν, σ are absolutely continuous w.r.t. Lebesgue measure, and if the function ϕ has an absolutely continuous derivative, then $\nu\big(\{0\}\big) = \sigma\big(\{0\}\big) = \nu\big(\{R\}\big) = \sigma\big(\{R\}\big) = 0$ and the above definition coincides with the classical one. To see a consequence of Definition 1, call d $\subset [0, R]$ the set of points where ϕ is differentiable. Since ϕ is Lipschitz continuous, meas(d) $= R$. For every $x_1, x_2 \in$ d one has

$$\begin{aligned} &\alpha(x_2)\phi'(x_2) - \alpha(x_1)\phi'(x_1) \\ &\quad = \beta(x_2)\phi(x_2) - \beta(x_1)\phi(x_1) - \int_{x_1}^{x_2} g\big(x,\phi(x)\big)\,\mathrm{d}x + \int_{x_1}^{x_2} \phi(x)\,(\mathrm{d}\nu + \mathrm{d}\sigma). \end{aligned}$$

Moreover, at any point $x \in]0, R[$, letting $x_1 \to x-$, $x_2 \to x+$, one checks that the left and right limits of the derivative $\phi'(x\pm)$ satisfy

$$\phi'(x+) - \phi'(x-) = \frac{\phi(x)}{\alpha(x)} \cdot (\nu + \sigma)\big(\{x\}\big).$$

The following construction reduces the measure-valued optimization problem (6)–(8) to a standard optimal control problem, with control functions in $\mathbf{L}^\infty$. Fix any point $z \in [0, R]$ and let δ_z be the Dirac measure concentrating a unit mass at the point z. This will allow us to compare the optimal strategy σ with some other strategy containing a point mass at z. Introduce the variable

$$s(x) = x + \nu\big([0, x]\big) + \sigma\big([0, x]\big) + \delta_z\big([0, x]\big),$$

so that, as $x \in [0, R]$,

$$s(x) \in [0, S], \qquad S = R + \nu\big([0, R]\big) + \sigma\big([0, R]\big) + 1.$$

The function $x \mapsto s(x)$ admits a Lipschitz continuous inverse: $x = x(s)$, with

$$\theta^x(s) \doteq \frac{\mathrm{d}}{\mathrm{d}s} x(s) \in [0, 1]. \tag{10}$$

Let $\tilde{\mu} \doteq \mathscr{L} + \nu + \sigma + \delta_z$ be the positive Radon measure on $[0, R]$ obtained as the sum of the Lebesgue measure $\mathscr{L}$ plus the three measures ν, σ, and δ_z. To define the densities $\theta^\nu = \frac{\mathrm{d}\nu}{\mathrm{d}\tilde{\mu}}$, $\theta^\sigma = \frac{\mathrm{d}\sigma}{\mathrm{d}\tilde{\mu}}$, and $\theta^z = \frac{\mathrm{d}\delta_z}{\mathrm{d}\tilde{\mu}}$ as functions of the variable s, we proceed as follows. For every $s \in [0, S]$, let

$$\begin{cases} s^+(s) \doteq \max\big\{\xi \in [0, S]\,;\;\; x(\xi) = x(s)\big\}, \\ s^-(s) \doteq \min\big\{\xi \in [0, S]\,;\;\; x(\xi) = x(s)\big\}. \end{cases} \tag{11}$$

Define the functions

$$\begin{cases} s \mapsto y^\nu(s) \doteq \nu\Big(\big[0,\, x(s)\big[\Big) + \nu\Big(\big\{x(s)\big\}\Big) \cdot \dfrac{s - s^-(s)}{s^+(s) - s^-(s)}, \\ s \mapsto y^\sigma(s) \doteq \sigma\Big(\big[0,\, x(s)\big[\Big) + \sigma\Big(\big\{x(s)\big\}\Big) \cdot \dfrac{s - s^-(s)}{s^+(s) - s^-(s)}, \\ s \mapsto y^z(s) \doteq \delta_z\Big(\big[0,\, x(s)\big[\Big) + \delta_z\Big(\big\{x(s)\big\}\Big) \cdot \dfrac{s - s^-(s)}{s^+(s) - s^-(s)}. \end{cases} \tag{12}$$

Observe that these functions are positive and nondecreasing. Moreover, (10) and (12) yield

$$x(s) + y^\nu(s) + y^\sigma(s) + y^z(s) = s \qquad \forall\, s \in [0, S]. \tag{13}$$

We can thus define

$$\theta^\nu(s) \doteq \frac{\mathrm{d}}{\mathrm{d}s} y^\nu(s), \qquad \theta^\sigma(s) \doteq \frac{\mathrm{d}}{\mathrm{d}s} y^\sigma(s), \qquad \theta^z(s) \doteq \frac{\mathrm{d}}{\mathrm{d}s} y^z(s).$$

Because of (13), the above definitions imply $\theta^x(s) + \theta^\nu(s) + \theta^\sigma(s) + \theta^z(s) = 1$ for a.e. $s \in [0, S]$.

For convenience, we shall now write $c(s) \doteq c\big(x(s)\big)$, and similarly for $\alpha(s)$ and $\beta(s)$. On the interval $[0, S]$, we also consider the functions

$$s \mapsto \phi(s) \doteq \phi\big(x(s)\big), \qquad\qquad s \mapsto \psi(s) \doteq \alpha\big(x(s)\big)\,\phi'\big(x(s)\big),$$

where ϕ' denotes the derivative of ϕ w.r.t. x. Notice that the map ϕ is well defined and continuous. However, if $\mu\big(\{x\}\big) > 0$, then ϕ' is discontinuous at x, hence it is not well defined as a function of the parameter s. To take care of points where μ has a point mass, and ϕ' thus has a jump, recalling (11) we define

$$\psi(s) \doteq \alpha\big(x(s)\big) \cdot \left(\frac{s - s^-}{s^+ - s^-} \cdot \phi'\big(x(s)+\big) \;+\; \frac{s^+ - s}{s^+ - s^-} \cdot \phi'\big(x(s)-\big) \right).$$

By (7), the maps ϕ and $\psi = \phi'$ provide a solution to the system of O.D.E's

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}s}\phi(s) = \theta^x(s)\,\dfrac{\psi(s)}{\alpha(s)}, \\ \dfrac{\mathrm{d}}{\mathrm{d}s}\psi(s) = \big(\theta^\nu(s) + \theta^\sigma(s)\big) \cdot \phi(s) \\ \qquad\qquad +\theta^x(s)\left[\beta'(s)\phi(s) + \dfrac{\beta(s)}{\alpha(s)}\psi(s) - g\big(x(s), \phi(s)\big)\right], \end{cases}$$

with boundary data

$$\psi(0) = \beta(0)\phi(0), \qquad\qquad \psi(S) = \beta(S)\phi(S).$$

From the optimality of the measure σ, it now follows that the control functions $u_1(s) \equiv 1$, $u_2(s) \equiv 0$ are optimal for the problem

$$\text{maximize:}\;\; J(u_1, u_2) \doteq \int_0^S \big(y_1(s) - c(s)\big)\left[\theta^\sigma(s)u_1(s) + \theta^z(s)u_2(s)\right]\mathrm{d}s, \tag{14}$$

for the control system

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}s}y_1 = \dfrac{\theta^x(s)}{\alpha(s)}\,y_2, \\ \dfrac{\mathrm{d}}{\mathrm{d}s}y_2 = \Big(\theta^\nu(s) + \theta^\sigma(s)u_1 + \theta^z u_2\Big)\,y_1 \\ \qquad\qquad +\theta^x(s)\left[\beta'(s)y_1 + \dfrac{\beta(s)}{\alpha(s)}y_2 - g\big(x(s), y_1\big)\right], \end{cases} \tag{15}$$

with boundary conditions

$$y_2(0) = \beta(0)y_1(0), \qquad y_2(S) = \beta(S)y_1(S). \tag{16}$$

The control $u = (u_1, u_2)$ in (14)–(16) ranges over all couples of nonnegative functions $u_1, u_2 : [0, S] \mapsto IR_+$. Notice that in this optimization problem all maps $\alpha, \beta, c, \theta, x$ are given functions of $s \in [0, S]$ and do not depend on the particular choice of the controls u_1, u_2.

Since $(u_1, u_2) \equiv (1, 0)$ is optimal, by Pontryagin's maximum principle there exists an adjoint vector $p = (p_1, p_2)$ such that the following equations hold.

$$\begin{cases} \dfrac{d}{ds} y_1 = \dfrac{\theta^x}{\alpha}\, y_2, \\ \dfrac{d}{ds} y_2 = (\theta^\nu + \theta^\sigma)\cdot y_1 + \theta^x \cdot \left[\beta' y_1 + \dfrac{\beta}{\alpha} y_2 - g(x, y_1)\right], \end{cases} \tag{17}$$

$$\begin{cases} \dfrac{d}{ds} p_1 = -\Big[(\theta^\nu + \theta^\sigma) + \theta^x\big(\beta' - g_\phi(x, y_1)\big)\Big] p_2 - \theta^\sigma, \\ \dfrac{d}{ds} p_2 = -\dfrac{\theta^x}{\alpha}\,(p_1 + \beta\, p_2), \end{cases} \tag{18}$$

together with the boundary conditions

$$\begin{cases} y_2(0) - \beta(0)y_1(0) \ = 0, \\ y_2(S) - \beta(S)y_1(S) = 0, \end{cases} \qquad \begin{cases} p_1(0) + \beta(0)p_2(0) \ = 0, \\ p_1(S) + \beta(S)p_2(S) = 0, \end{cases} \tag{19}$$

and moreover, for almost every $s \in [0, S]$ the following maximality condition holds:

$$\begin{aligned} &\theta^\sigma\big[(p_2 + 1)y_1 - c\big] \\ &\quad = \max_{\omega_1, \omega_2 \geq 0} \Big\{\theta^\sigma\big[(p_2 + 1)y_1 - c\big]\omega_1 + \theta^z\big[(p_2 + 1)y_1 - c\big]\omega_2\Big\}. \end{aligned} \tag{20}$$

Notice that (20) is equivalent to the two conditions

$$\theta^\sigma \cdot \big[(p_2 + 1)y_1 - c\big] = 0, \qquad \theta^z \cdot \big[(p_2 + 1)y_1 - c\big] \leq 0, \tag{21}$$

for a.e. $s \in [0, S]$.

It is convenient to rewrite the above conditions in terms of the original space variable x. Recall that $y_1 = \phi$, $y_2 = \alpha\phi'$, and set $q = p_2$. Observing that

$$\frac{d(\alpha\phi')}{dx} = \frac{d(\alpha\phi')}{ds}\cdot\frac{ds}{dx} = \frac{1}{\theta^x}\,\frac{dy_2}{ds}, \qquad q' = \frac{dq}{dx} = \frac{1}{\theta^x}\,\frac{dp_2}{ds} = -\frac{p_1 + \beta p_2}{\alpha},$$

from (17)–(19), we obtain the second-order equations

$$\begin{cases} (\alpha\phi')' + (\beta\phi)' + g(x, \phi) = \phi(\nu + \sigma), \\ (\alpha q')' + \beta q' + g_\phi(x, \phi)q = q(\nu + \sigma) + \sigma, \end{cases} \tag{22}$$

with boundary conditions

$$\begin{cases} \alpha(0)\phi'(0) - \beta(0)\phi(0) &= 0, \\ \alpha(R)\phi'(R) - \beta(R)\phi(R) = 0, \end{cases} \qquad \begin{cases} q'(0) = 0. \\ q'(R) = 0. \end{cases} \tag{23}$$

By first identity in (21) there exists a set $\mathcal{N} \subset [0, R]$ with $\sigma(\mathcal{N}) = 0$ such that $(q(x)+1)\phi(x) - c(x) = 0$ at every point $x \notin \mathcal{N}$. Moreover, at the particular point z, the second inequality in (21) implies $(q(z)+1)\phi(z) - c(z) \le 0$. We now observe that the previous construction can be performed with an arbitrary choice of the point $z \in [0, R]$. Our analysis can thus be summarized as follows.

Theorem 1. *Assume that the couple (σ, ϕ) provides an optimal solution to the optimization problem (6)–(8), where σ ranges within the class of all nonnegative Radon measures on the interval $[0, R]$, and ϕ is a corresponding solution of (7)–(8). Then there exists an adjoint function $q : [0, R] \mapsto I\!R$ such that the boundary value problem (22)–(23) is satisfied, in the sense of Definition 1. Moreover, one has the optimality conditions*

$$(q(x) + 1)\phi(x) - c(x) = 0 \qquad \sigma\text{-a.e. on } [0, R], \tag{24}$$

$$(q(x) + 1)\phi(x) - c(x) \le 0 \qquad \forall x \in [0, R]. \tag{25}$$

In the special case $\alpha(x) \equiv 1$, $\beta(x) \equiv 0$, the (22)–(25) reduce to

$$\begin{cases} \phi'' + g(x, \phi) &= \phi\,(\nu + \sigma), \\ q'' + g_\phi(x, \phi)q &= q(\nu + \sigma) + \sigma, \end{cases} \tag{26}$$

with boundary conditions

$$\phi'(0) = \phi'(R) = 0, \qquad q'(0) = q'(R) = 0, \tag{27}$$

The optimality conditions are still given by (24)–(25).

We remark that for a multidimensional optimization problem related to a linear elliptic PDE, necessary conditions of a similar type were derived in [9].

3 Examples

Relying on the necessary conditions established in the previous section, we now examine more in detail the solution to the optimization problem

$$\text{maximize:} \;\; J(\sigma) \doteq \int_0^R \big(\phi - c\big)\, d\sigma.$$

subject to

$$\phi'' + g(x, \phi) = \phi\,\sigma, \qquad \phi'(0) = \phi'(R) = 0. \tag{28}$$

Example 1. Assume $g(\phi) = (2-\phi)\phi$, $c(x) \equiv 1$. This corresponds to a space-homogeneous optimal control problem. As already observed in [12], in this case the optimal strategy σ is the measure having constant density $1/2$ w.r.t. Lebesgue measure. The corresponding optimal solution is $\phi(x) = 3/2$. The optimality conditions (24)–(27) are satisfied taking as adjoint function $q(x) = -1/3$.

Example 2. We now show that if the cost function c is discontinuous, the optimal strategy can be a measure σ containing point masses. On the interval $[0,2]$, consider the functions

$$g(x,\phi) = (2-\phi)\phi, \qquad c(x) = \begin{cases} 1 & \text{if} \quad x \le 1\,, \\ 3 & \text{if} \quad x > 1\,. \end{cases} \tag{29}$$

Since the cost function $c(\cdot)$ is lower semicontinuous, the existence of an optimal solution (σ,ϕ) to (6)–(8) is provided by Theorem 1 in [8]. We observe that σ must be a nonzero measure whose support satisfies

$$\mathrm{Supp}(\sigma) \subset \Gamma \doteq \Big\{x \in [0,2]\,;\ \ \phi(x) > c(x)\Big\} \subseteq [0,1]. \tag{30}$$

Otherwise, the alternative measure $\tilde\sigma$, defined as $\tilde\sigma(A) \doteq \sigma(A\cap\Gamma)$ for every Borel set A, would achieve a strictly better payoff.

As shown in Lemma 1 in [8], if the fish density ϕ vanishes at some point $x_0 \in [0,2]$, then ϕ is identically zero. In this case, we claim that

$$1 \le \phi(x) \le 2 \qquad\qquad \forall x \in [0,2].$$

Indeed, for any positive measure σ, every nonnegative solution of (28) satisfies $\phi \le 2$. On the other hand, if the open set $S \doteq \big\{x\ \in]0,2[\,;\ \ \phi(x) < 1\big\}$ is nonempty, consider a maximal open interval $]a,b[\ \subseteq\ S$. Observing that

$$0 < \phi(x) < 1, \qquad \phi'' \ =\ -(2-\phi)\phi\ <\ 0 \qquad \forall x \in]a,b[,$$

$$\begin{cases} \phi'(a) = 0 & \text{if}\ \ a = 0, \\ \phi(a) = 1 & \text{if}\ \ a > 0, \end{cases} \qquad\qquad \begin{cases} \phi'(b) = 0 & \text{if}\ \ b = 2, \\ \phi(b) = 1 & \text{if}\ \ b < 2, \end{cases}$$

an application of the maximum principle for parabolic equations yields $\phi(x) \ge 1$ for all $x \in]a,b[$.

By Theorem 1, there exists an adjoint function q such that

$$\begin{cases} \phi'' = -(2-\phi)\phi + \phi\sigma, \\ q'' = (2\phi-2)q + (1+q)\sigma, \end{cases} \qquad\qquad \begin{cases} \phi'(0) = \ \phi'(2)\ =\ 0, \\ q'(0) = \ q'(2)\ =\ 0, \end{cases} \tag{31}$$

$$\begin{cases} (1+q)\phi\ =\ 1 & \sigma\text{-a.e.} \\ (1+q)\phi\ \le\ c(x) & \text{for a.e. } x \in\ [0,2]. \end{cases} \tag{32}$$

Observing that the functions ϕ, q are Lipschitz continuous, (32) can be rewritten as

$$\begin{cases} (1+q)\phi = 1 & x \in \mathrm{Supp}(\sigma) \\ (1+q)\phi \le 1 & x \in [0,1], \\ (1+q)\phi \le 3 & x \in]1,2]. \end{cases} \tag{33}$$

Recalling (29), (30), and observing that $g(\phi) \le 0$ for $\phi \ge 2$, we conclude that the function ϕ must satisfy the strict inequalities

$$1 < \phi(x) < 2 \qquad \forall x \in [0,2]. \tag{34}$$

Since q is continuous and $\phi \in [1,2]$, the first equality in (33) implies

$$q = \frac{1}{\phi} - 1 \in \left] -\frac{1}{2},\, 0 \right[\qquad \forall x \in \mathrm{Supp}(\sigma).$$

Next, we prove the following claim:
(C) The measure σ is absolutely continuous w.r.t. Lebesgue measure on the half-open interval $[0,1[$, but contains a positive mass at the point $x = 1$. Moreover, $0 \in \mathrm{Supp}(\sigma)$.

Indeed, by (34) and (30) it is clear that $\mathrm{Supp}(\sigma) \subseteq [0,1]$. To prove that $0 \in \mathrm{Supp}(\sigma)$, assume that, on the contrary, $a \doteq \min\{x\,;\ \ x \in \mathrm{Supp}(\sigma)\} > 0$. Then, from (31) it follows that $\phi''(x) < 0$, $q''(x) < 0$ for all $x \in]0,a[$, and hence also $q'(x) < 0$ and $p'(x) < 0$ for all $x \in]0,a[$. Calling $\Theta(x) = \big(1+q(x)\big)\phi(x)$ the switching function, its derivative satisfies

$$\Theta' = (1+q)\phi' + q'\phi < 0 \qquad \forall x \in]0,a[.$$

In turn, this implies $\Theta(x) > \Theta(a) = 1$, $\forall x \in]0,a[$, providing a contradiction with the second inequality in (33).

On the other hand, if $b \doteq \max\{x; x \in \mathrm{Supp}(\sigma)\} < 1$, a contradiction is obtained by an entirely similar argument. For $x \in]b,2[$, we have $\phi''(x) < 0$, $q''(x) < 0$, and hence $q'(x) > 0$ and $p'(x) > 0$. This implies

$$\Theta'(x) > 0 \qquad \forall x \in]b,2[.$$

In turn, this implies $1 = \Theta(b) < \Theta(x)$, $\forall x \in]b,2[$, providing a contradiction with the second inequality in (33).

Next, assume $\sigma\big(\{1\}\big) = 0$. Since $1 \in \mathrm{Supp}(\sigma)$ and the functions ϕ, q are continuous, there must exist an increasing sequence of points $x_n \in \mathrm{Supp}(\sigma)$, with $x_n \to 1$. By a nonsmooth version of the intermediate value theorem, there exist a sequence of points $y_n \to 1$ such that $0 \in \partial\Theta(y_n)$ for every $n \ge 1$. Here, $\partial\Theta$ denotes the Clarke generalized gradient of Θ. Since $\Theta'(1+) > 0$, this shows that $\sigma\big([y_n, 1]\big) \ge c_0$ for some constant $c_0 > 0$ and all $n \ge 1$. Hence $\sigma\big(\{1\}\big) \ge c_0$, proving that σ must contain a Dirac mass at $x = 1$.

Finally, if the restriction of σ to $[0,1[$ is not absolutely continuous w.r.t. Lebesge measure, we could find a sequence of intervals $[a_n, b_n] \subset [0,1[$ with $a_n, b_n \in \mathrm{Supp}(\sigma)$ and $\sigma\big([a_n, b_n]\big) \geq n(b_n - a_n) > 0$, $\forall n \geq 1$. Observing that

$$\begin{aligned}\Theta'' &= (1+q)\phi'' + 2q'\phi' + q''\phi \\ &= (1+q)\big[(2-\phi)\phi + \phi\sigma\big] + 2q'\phi' + \big[(2\phi-2)q + (1+q)\sigma\big]\phi\end{aligned}$$

we conclude that $\Theta'(b_n+) > \Theta'(a_n-)$ for all n sufficiently large. This provides a contradiction with the assumptions $\Theta(a_n) = \Theta(b_n) = 1$ and $\Theta(x) \leq 1$ for all x in a neighborhood of the interval $[a_n, b_n]$. This completes the proof of our claim **(C)**.

Relying on the necessary conditions (31)–(32), to construct an optimal solution we proceed as follows. Assuming that $\mathrm{Supp}(\sigma) = [0,1]$, for $x \in [0,1[$ the optimality conditions yield

$$\begin{cases} \phi'' = (\phi - 2 + u)\phi, \\ q'' = (2\phi - 2 + u)q + u, \end{cases} \tag{35}$$

and

$$\begin{cases} \phi(1+q) = 1, \\ \phi'(1+q) + \phi q' = 0, \\ \phi''(1+q) + 2\phi' q' + \phi q'' = 0. \end{cases} \tag{36}$$

In turn these imply

$$u = \frac{\phi''}{\phi} - \phi + 2, \qquad\qquad q = \frac{1}{\phi} - 1. \tag{37}$$

From the third equation in (36), using (37), the second equation in (36) and then the second equation in (35), one gets $\frac{\phi''}{\phi} + 2\phi'\left(-\frac{\phi'}{\phi}(1+q)\right) + \phi(2\phi - 2 + u)q + \phi u = 0$ and finally

$$\frac{\phi''}{\phi} + 2\phi'\left(-\frac{\phi'}{\phi}\frac{1}{\phi}\right) + \phi\left(2\phi - 2 + \frac{\phi''}{\phi} - \phi + 2\right)\left(\frac{1}{\phi} - 1\right) + \phi\left(\frac{\phi''}{\phi} - \phi + 2\right) = 0,$$

which leads to

$$\phi'' = \frac{(\phi')^2}{\phi} + \left(\phi - \frac{3}{2}\right)\phi^2 .$$

Combine this with the first equation in (35), we get

$$u = \left(\frac{\phi'}{\phi}\right)^2 + \phi^2 - \frac{5}{2}\phi + 2 \geq \left(\frac{\phi'}{\phi}\right)^2 + \frac{7}{16} > 0.$$

To construct the optimal solution, we seek a continuous function $\phi : [0,2] \mapsto [1,2]$ such that

$$\phi'' = \frac{(\phi')^2}{\phi} + \left(\phi - \frac{3}{2}\right)\phi^2 \qquad x \in]0,1[, \tag{38}$$

$$\phi'' = (\phi - 2)\phi \qquad \text{if} \quad x \in]1, 2[, \tag{39}$$

and satisfies the boundary conditions

$$\phi'(0) = 0, \qquad \phi'(2) = 0. \tag{40}$$

Notice that ϕ is Lipschitz continuous but ϕ' is expected to have a discontinuity at $x = 1$.

In addition, we seek a solution q to

$$q'' = (2\phi - 2)q \qquad x \in]1, 2[\tag{41}$$

with boundary conditions

$$q(1) = \frac{1}{\phi(1)} - 1, \qquad q'(2) = 0, \quad \frac{q'(1+)}{1 + q(1)} = \frac{\phi'(1+) - 2\phi'(1-)}{\phi(1)}. \tag{42}$$

Notice that the first identity in (40) is derived from (36), while the last one follows from the jump conditions

$$\begin{cases} \phi'(1+) - \phi'(1-) = \phi(1)\,\sigma\big(\{1\}\big), \\ q'(1+) - q'(1-) = \big(1 + q(1)\big)\,\sigma\big(\{1\}\big), \end{cases}$$

observing that (36) implies $\frac{q'(1-)}{1+q(1)} = -\frac{\phi'(1-)}{\phi(1)}$. The optimal solution $\phi(\cdot)$ is now determined by solving the three second-order O.D.E's (38), (39), and (41), together with six boundary conditions, namely (40), (42), and the trivial continuity relation $\phi(1+) = \phi(1-)$. A numerical solution to this problem is given in Fig. 1. A posteriori, we check that the assumption $\text{Supp}(\sigma) = [0, 1]$ is satisfied by the numerically

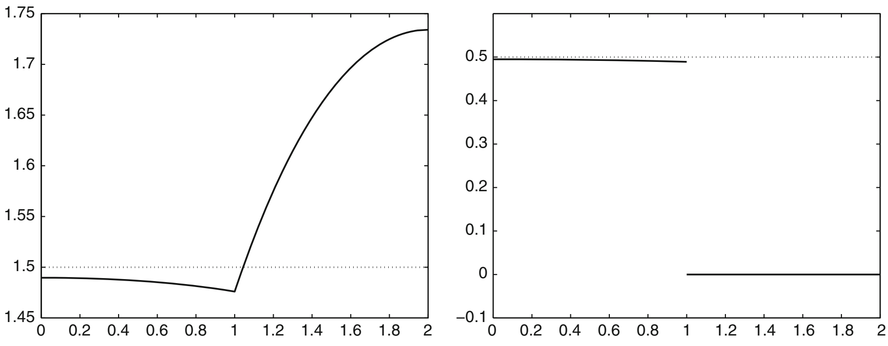


Fig. 1 Numerically computed solutions: the fish population density ϕ (*left*) and the fishing intensity density u (*right*) for Example 1 (*dotted line*) and Example 2 (*solid line*). In the second case, the fishing intensity contains a point mass of 0.4075 at $x = 1$

computed solution. We note that the derivative ϕ' of the fish population density has an upward jump at $x = 1$, corresponding to a point mass in the measure σ, which is equal to 0.4075 in this case. Moreover, on the subinterval $[0, 1[$ the optimal harvesting effort σ has density very close to 0.5. This is the optimal density in the spatially independent setting considered in Example 1.

4 Necessary Conditions for the Differential Game

Consider a differential game for N equal players, where each one wishes to maximize his/her payoff

$$\text{maximize: } J(\sigma) \doteq \int_0^R \big(\phi(x) - c(x)\big)\, d\sigma. \tag{43}$$

The maximum is sought over all pairs (σ, ϕ), where $\sigma \in \mathscr{M}_+\big([0, R]\big)$ is a nonnegative Radon measure on the closed interval $[0, R]$ and $\phi : [0, R] \mapsto I\!R_+$ provides a distributional solution to the second-order boundary value problem

$$\phi'' + g(x, \phi) - \phi\nu = \phi\,\sigma, \tag{44}$$

with boundary conditions

$$\phi'(0) \;=\; \phi'(R) \;=\; 0. \tag{45}$$

If the couple (σ, ϕ) provides an optimal solution of the above problem, in connection with the measure $\nu = (N-1)\sigma$, then we say that (σ, ϕ) a *symmetric Nash equilibrium solution* to the N-players, noncooperative differential game.

Applying (26)–(27), with $\nu = (N-1)\sigma$, we obtain the existence of an adjoint function q such that

$$\begin{cases} \phi'' + g(x, \phi) = N\phi\sigma, \\ q'' + g_\phi(x, \phi)q = (Nq + 1)\sigma, \end{cases} \tag{46}$$

with boundary conditions

$$\phi'(0) = \phi'(R) = 0, \qquad\qquad q'(0) = q'(R) = 0. \tag{47}$$

Moreover, the conditions in (24)–(25) remain the same

$$\begin{aligned} &(q(x) + 1)\phi(x) - c(x) = 0 \qquad && \sigma\text{-a.e. on } [0, R], \\ &(q(x) + 1)\phi(x) - c(x) \le 0 \qquad && \forall x \in [0, R]. \end{aligned} \tag{48}$$

Example 3. For the problem (43)–(45), assume that $c(x) \equiv \gamma > 0$, while $g(x, \phi) = (2 - \phi)\phi$. Then a spatially independent symmetric Nash equilibrium solution is found by solving the algebraic system

$$\begin{cases} (2-\phi) = & Nu, \\ (2-2\phi)q = & (Nq+1)u, \\ (q+1)\phi = & \gamma. \end{cases}$$

This yields

$$\begin{aligned} q = \frac{\gamma}{\phi} - 1, \qquad\qquad & u = \frac{2-\phi}{N}, \\ (2-2\phi)\left(\frac{\gamma}{\phi} - 1\right) &= \left\{N\left(\frac{\gamma}{\phi}-1\right)+1\right\}\frac{2-\phi}{N}, \\ N(2-2\phi)(\gamma-\phi) &= (N\gamma - N\phi + \phi)(2-\phi), \\ (N+1)\phi^2 - (N\gamma+2)\phi &= 0, \\ \phi &= \frac{N\gamma+2}{N+1}. \end{aligned}$$

On an interval of length R, the total catch is $Nu\phi\, R = g(\phi)\, R \;=\; \left(2 - \frac{N\gamma+2}{N+1}\right) \frac{N\gamma+2}{N+1}\, R$, while the total payoff is

$$Nu(\phi-\gamma)R = \left(2 - \frac{N\gamma+2}{N+1}\right)\left(\frac{N\gamma+2}{N+1} - \gamma\right) R.$$

Notice that as $N \to \infty$, the fish density satisfies $\phi_N \to \gamma$. The total catch approaches $(2-\gamma)\gamma\, R$, while the total payoff approaches zero.

Example 4. We now modify Example 3, assuming that a marine park is created on the open domain $]\xi, R]$. The new cost function thus takes the form

$$c(x) = \begin{cases} \gamma & \text{if} \quad x \le \xi\,, \\ +\infty & \text{if} \quad x > \xi\,. \end{cases}$$

Given an integer $N \ge 1$, a symmetric Nash equilibrium solution with N players can be computed by the same techniques used in Example 2.

Assuming that $\mathrm{Supp}(\sigma) = [0, \xi]$, for $x \in [0, \xi[$ the optimality conditions yield

$$\begin{cases} \phi'' = (\phi - 2 + Nu)\phi, \\ q'' = (2\phi - 2 + Nu)q + u, \end{cases} \tag{49}$$

$$\begin{cases} \phi(1+q) = \gamma, \\ \phi'(1+q) + \phi q' = 0, \\ \phi''(1+q) + 2\phi' q' + \phi q'' = 0. \end{cases} \tag{50}$$

In turn these imply

$$Nu = \frac{\phi''}{\phi} - \phi + 2, \qquad q = \frac{\gamma}{\phi} - 1. \tag{51}$$

From the third equation in (50), using (51), the second equation in (50) and then the second equation in (49), one gets

$$\frac{\gamma\phi''}{\phi} + 2\phi'\left(-\frac{\phi'}{\phi}(1+q)\right) + \phi(2\phi - 2 + Nu)q + \phi u = 0,$$

and finally

$$\frac{\gamma\phi''}{\phi} + 2\phi'\left(-\frac{\phi'}{\phi}\frac{\gamma}{\phi}\right) + \phi\left(2\phi - 2 + \frac{\phi''}{\phi} - \phi + 2\right)\left(\frac{\gamma}{\phi} - 1\right)$$
$$+ \frac{\phi}{N}\left(\frac{\phi''}{\phi} - \phi + 2\right) = 0,$$

which gives

$$\left(\frac{2\gamma}{\phi} - \frac{N-1}{N}\right)\phi'' - 2\gamma\left(\frac{\phi'}{\phi}\right)^2 + \left(\gamma + \frac{2}{N}\right)\phi - \frac{N+1}{N}\phi^2 = 0.$$

To construct the optimal solution, we seek a continuous function $\phi : [0, R] \mapsto [\gamma, 2]$ such that

$$\phi'' = \left(\frac{2\gamma}{\phi} - \frac{N-1}{N}\right)^{-1}\left[2\gamma\left(\frac{\phi'}{\phi}\right)^2 - \left(\gamma + \frac{2}{N}\right)\phi + \frac{N+1}{N}\phi^2\right] \quad x \in]0, \xi[, \tag{52}$$

$$\phi'' = (\phi - 2)\phi \qquad x \in]\xi, R[, \tag{53}$$

and satisfies the boundary conditions

$$\phi'(0) = 0, \qquad \phi'(R) = 0. \tag{54}$$

Notice that ϕ is Lipschitz continuous but ϕ' is expected to have a discontinuity at $x = \xi$.

In addition, we seek a solution q to

$$q'' = (2\phi - 2)q \qquad x \in]\xi, R[\tag{55}$$

with boundary conditions

$$
\begin{aligned}
q(\xi) &= \frac{\gamma}{\phi(\xi)} - 1, \quad q'(R) = 0, \\
q'(\xi+) &= \frac{1+Nq}{N\phi}\,\phi'(\xi+) - \frac{1+N+2Nq}{N\phi}\,\phi'(\xi-).
\end{aligned}
\tag{56}
$$

Notice that the first identity in (56) is derived from (50), while the last one follows from the jump conditions

$$
\begin{cases}
\phi'(\xi+) - \phi'(\xi-) = N\phi(\xi)\,\sigma\big(\{\xi\}\big), \\
q'(\xi+) - q'(\xi-) = \big(1 + Nq(\xi)\big)\,\sigma\big(\{\xi\}\big),
\end{cases}
$$

observing that the second equation in (50) implies $\frac{q'(\xi-)}{1+q(\xi)} = -\frac{\phi'(\xi-)}{\phi(\xi)}$. The optimal solution $\phi(\cdot)$ can now be determined by solving the three second-order O.D.E's (52), (53), and (55), together with six boundary conditions, namely (54), (56), and the trivial continuity relation $\phi(\xi+) = \phi(\xi-)$.

Figure 2 shows the plots of total catch as functions of ξ, where ¶ $=]\xi, 2]$ is the location of the marine park where no fishing is allowed. Here, the domain is $[0, R] = [0, 2]$ and the number of fishermen is $N = 40$. Two different fishing costs are considered: $\gamma = 0.5$ and $\gamma = 0.3$. For the smaller fishing cost $\gamma = 0.3$, we see that the marine park yielding the largest catch is ¶ $\approx]0.45, 2]$. On the other hand, for the larger cost $\gamma = 0.5$, the optimal marine reserve is smaller, namely ¶ $\approx]0.75, 2]$.

Figure 3 shows the total catch and total payoff as functions of the number of fishermen. Here, the marine park is ¶ $=]1, 2]$, while the fishing cost outside the park is $\gamma = 0.3$. We see that the total catch decreases to a nonzero limit as N becomes large, with the maximum value reached with 2 fishermen. But for total payoff, it is a decreasing function and goes to 0 as N increases, with the maximum value reached with 1 fisherman.

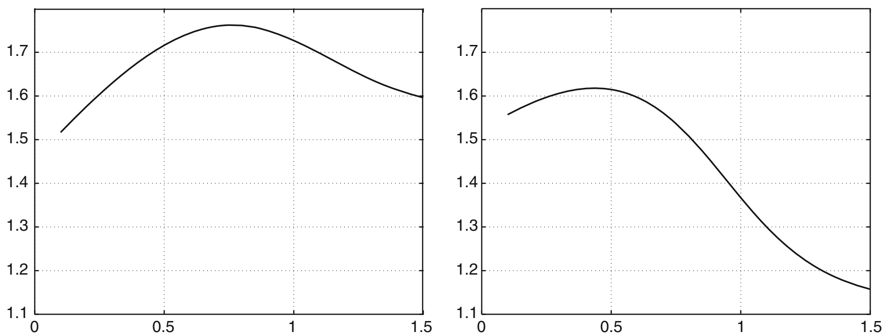


Fig. 2 Plots of total catch depending on the location of marine reserve ξ, with $N = 40$ fishermen. On the left the cost of fishing (outside the park) is $\gamma = 0.5$, on the right $\gamma = 0.3$

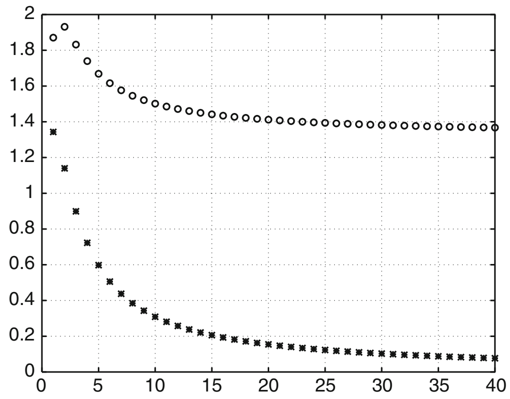


Fig. 3 The total catch (marked with 'o') and total payoff (marked with '*') depending on the number of fishermen. Here ¶ $=]1,2]$ and $\gamma = 0.3$

5 The Fish Population for a Large Number of Fishermen

Let the assumptions (A1)–(A2) hold. Assume that for each $N \geq 1$, the noncooperative game (43)–(45) admits a symmetric Nash equilibrium solution, say (σ_N, ϕ_N).

As the number N of fishermen grows without bound, a natural problem is to study the limiting density of the fish population. In this section, we will prove that the limit

$$\phi_\infty(x) \doteq \lim_{N\to\infty} \phi_N(x)$$

indeed exists, and can be characterized as the largest subsolution to

$$\phi'' + g(x,\phi) = 0, \qquad \phi'(0) = \phi'(R) = 0, \tag{57}$$

which satisfies the additional constraint

$$\phi(x) \leq c(x) \qquad \forall x \in [0,R]. \tag{58}$$

Equivalently, the limit ϕ_∞ can also be characterized as the unique strictly positive solution to a variational inequality. To state these results more precisely, we begin with some definitions.

Definition 2. By a subsolution of the boundary value problem (57), we mean a Lipschitz continuous map $\phi : [0,T] \mapsto I\!R$ such that

(i) The map $x \mapsto \phi'(x)$ has bounded variation and satisfies

$$\phi'(0+) \geq 0, \qquad \phi'(R-) \leq 0.$$

(ii) For every nonnegative test function $\eta \in \mathscr{C}_c^1\big(]0,R[\big)$, one has

$$\int_0^R \Big\{-\phi'\eta' + g(x,\phi)\,\eta\Big\}\,\mathrm{d}x \;\geq\; 0.$$

We remark that the largest subsolution of (57) satisfying the constraint (58) can be characterized as the solution to a variational inequality. Indeed, calling BV the space of functions with bounded variation, consider the family of functions

$$\mathcal{K}_c \doteq \Big\{\phi : [0, R] \mapsto I\!R, \qquad \phi \text{ is Lipschitz continuous with } \phi' \in BV, \\ \phi'(0+) \geq 0, \quad \phi'(R-) \leq 0, \quad \phi(x) \leq c(x) \ \ \forall x \in [0, R]\Big\}.$$

Here, the lower semicontinuous function $c(\cdot)$ plays the role of an obstacle. Then the limiting density ϕ_∞ of the fish population satisfies $\phi_\infty \in \mathcal{K}_c$ and

$$\int_0^R \phi'_\infty(\psi' - \phi'_\infty)\,\mathrm{d}x + \int_0^R g(x, \phi_\infty)\cdot(\phi_\infty - \psi)\,\mathrm{d}x \ \geq \ 0 \quad \forall \psi \in \mathcal{K}_c.$$

In particular, this implies

$$\begin{aligned} &\phi''_\infty + g(x, \phi_\infty) \geq 0 && \text{for all } \ x \in]0, R[, \\ &\phi''_\infty + g(x, \phi_\infty) = 0 && \text{on the open set where } \ \phi_\infty(x) < c(x). \end{aligned}$$

Theorem 2. *Let the function g and c satisfy the assumptions in (A1)–(A2). For each $N \geq 1$, let (ϕ_N, σ_N) be a symmetric Nash equilibrium solution to the N-players noncooperative differential game (43)–(45), with $\phi_N > 0$. Then, as $N \to \infty$, one has the uniform convergence $\phi_N \to \phi_\infty$ in $\mathcal{C}^0\big([0, R]\big)$, where ϕ_∞ provides the largest positive subsolution to (57) which satisfies the additional constraint (58).*

Proof. We divide the argument in several steps.

(i) We first observe that the sequence of nonnegative Radon measures $N\sigma_N$ remains uniformly bounded. Indeed, since the support of each measure σ_N is contained in the region where $\phi_N(x) \geq c(x) \geq c_0 > 0$, we have

$$\int_{[0,R]} c_0 N\,\mathrm{d}\sigma_N \leq \int_{[0,R]} \phi_N\, N\,\mathrm{d}\sigma_N = \int_0^R g\big(x, \phi_N(x)\big)\,\mathrm{d}x \\ \leq \int_0^R \Big(\max_\xi g(x, \xi)\Big)\,\mathrm{d}x.$$

In turn, the boundedness of the measures $N\sigma_N$ implies that the positive functions ϕ_N are uniformly Lipschitz continuous. By possibly taking a subsequence, we thus obtain the existence of a Lipschitz map ϕ_∞ such that $\phi_N(x) \to \phi_\infty(x)$ as $N \to \infty$, uniformly on $[0, R]$.

(ii) Recalling the assumptions in (A1), introduce the constants

$$h_{\min} \doteq \min_{x\in[0,R]} h(x) \ > \ 0, \qquad\qquad h_{\max} \doteq \max_{x\in[0,R]} h(x).$$

A comparison argument now shows that the positive functions ϕ_N satisfy the uniform bounds

$$0 \ < \ \min\{c_0, h_{\min}\} \ \leq \ \phi_N(x) \ \leq \ h_{\max}. \tag{59}$$

In turn, this implies

$$0 \ < \ \min\{c_0, h_{\min}\} \ \leq \ \phi_\infty(x) \ \leq \ h_{\max}.$$

Next, define the largest subsolution ϕ^* by setting

$$\begin{aligned}\phi^*(x) \doteq \ \sup \Big\{\phi(x); \ \ \phi \ \text{ is a subsolution of (57)}, \ 0 \leq \phi(y) \\ \leq c(y) \ \ \forall\, y \in [0, R]\Big\}.\end{aligned} \tag{60}$$

Since the cost function $c(\cdot)$ is lower semicontinuous, it is clear that ϕ^* is a Lipschitz continuous subsolution of (57) and satisfies the constraint (58). In the remainder of the proof, we will establish the equality $\phi_\infty = \phi^*$. In particular, this will show that the limit ϕ_∞ is independent of the choice of the subsequence. Hence, the entire sequence $(\phi_N)_{N\geq 1}$ converges to the same limit.

(iii) The inequality $\phi_\infty \geq \phi^*$ will be proved by showing that, for every $N \geq 1$,

$$\phi_N(x) \geq \phi^*(x) \qquad\qquad \text{for all } \ x \in [0, R]. \tag{61}$$

To prove (61), we first observe that the measure σ_N is supported on the closed set where $\phi_N(x) \geq c(x)$. Otherwise, any player could choose the alternative strategy $\tilde{\sigma}_N \doteq \sigma_N \cdot \chi_{\{\phi_N \geq c\}}$ and achieve a strictly better payoff.

If now $\phi_N(y) < \phi^*(y)$ at some point $y \in [0, R]$, define $\lambda \doteq \max_{x\in[0,R]} \frac{\phi^*(x)}{\phi_N(x)} > 1$. We then have $\lambda\phi_N(\bar{x}) = \phi^*(\bar{x})$ at some point $\bar{x}$, while $\lambda\phi_N(x) \geq \phi^*(x)$ for all $x \in [0, R]$. Introducing the function

$$\varphi(x) \ \doteq \ \lambda\phi_N(x) - \phi^*(x) \ \geq \ 0, \tag{62}$$

a contradiction is obtained as follows.

By continuity, $\phi_N(x) < \phi^*(x) \leq c(x)$ for all x in an open neighborhood $\mathcal{N}_{\bar{x}}$ of the point $\bar{x}$. Hence, recalling the assumption (A2) on the source term,

$$(\phi^*)'' + f(x, \phi^*)\phi^* \geq 0, \qquad \phi_N'' + f(x, \phi_N)\phi_N = 0 \qquad x \in \mathcal{N}_{\bar{x}}.$$

In turn, this yields

$$\begin{aligned}(\varphi)'' + f\big(x, \phi_N(x)\big)\,\varphi &= -(\phi^*)'' - f(x, \phi_N)\phi^* \\ &\leq \Big[f\big(x, \phi^*(x)\big) - f\big(x, \phi_N(x)\big)\Big]\phi^*(x) \ < \ 0.\end{aligned} \tag{63}$$

Indeed, $\phi^*(x) > \phi_N(x) > 0$ and by the assumption (9) the map $\phi \mapsto f(x, \phi)$ is strictly decreasing.
Since φ is continuous and $\varphi(\bar{x}) = 0$, (63) implies that $\varphi'' < 0$ in a neighborhood of $\bar{x}$. Three cases must be considered.
If $0 < \bar{x} < R$, we immediately obtain a contradiction with the inequality in (62).
If $\bar{x} = 0$, since $\sigma_N(\{\bar{x}\}) = 0$, we have $(\phi^*)'(0+) \geq 0$, $\phi_N'(0+) = \phi_N(0)\sigma_N(\{0\}) = 0$. In a neighborhood of the origin, the inequalities

$$\varphi(0) = 0, \qquad \varphi'(0+) = \lambda\phi_N'(0+) - (\phi^*)'(0+) \leq 0, \qquad \varphi''(x) < 0$$

clearly yield a contradiction with (62).
In a similar way, if $\bar{x} = R$, we deduce

$$\varphi(R) = 0, \qquad \varphi'(R-) = \lambda\phi_N'(R-) - (\phi^*)'(R-) \geq 0, \qquad \varphi''(x) < 0,$$

reaching again a contradiction with (62).
Since (61) holds for every $N \geq 1$, this establishes the inequality $\phi_\infty \geq \phi^*$.

(iv) In the next two steps, we work toward the converse inequality $\phi_\infty \leq \phi^*$.
For each N, call (ϕ_N, q_N) the solution to the corresponding boundary value problem (46)–(47). Aim of this step is to prove that

$$\lim_{N\to\infty} \|q_N\|_{\mathscr{C}^0} = 0. \tag{64}$$

Indeed, q_N provides a solution to the linear, nonhomogeneous boundary value problem

$$q_N'' + \Big(g_\phi(x, \phi_N) - N\sigma_N\Big) q_N = \sigma_N, \qquad q_N'(0) = q_N'(R) = 0.$$

Hence, we have a representation $q_N(x) = \int_{[0,R]} K_N(x, y)\, d\sigma_N(y)$, where K_N is the Green kernel for the linear operator

$$\Lambda\psi \doteq \psi'' + \Big(g_\phi(x, \phi_N) - N\sigma_N\Big)\psi,$$

whose domain consists of functions satisfying $\psi'(0) = \psi'(R) = 0$.
By step **1**, as $N \to \infty$, the total mass of the measure σ_N approaches zero: $\sigma_N([0, R]) \to 0$. To establish the limit (64) it thus suffices to prove that all the kernels $K_N(\cdot, \cdot)$ are uniformly bounded. Toward this goal, fix $y \in [0, R]$ and call $\psi(x) \doteq K_N(x, y)$. Then ψ can be characterized as the solution to

$$\psi'' + \Big(g_\phi(x, \phi_N) - N\sigma_N\Big)\psi = \delta_y, \qquad \psi'(0) = \psi'(R) = 0,$$

where δ_y denotes the Dirac measure concentrating a unit mass at the point y.
We now recall that $g_\phi(x, \phi_N) = f(x, \phi_N) + f_\phi(x, \phi_N)\phi_N$. Moreover, the

functions ϕ_N satisfy the uniform bounds (59) and, for some constant c_f, the assumption (9) yields

$$f_\phi(x,\phi_N) \ \leq\ -c_f\ <\ 0 \qquad\qquad \forall x \in [0,R].$$

Introduce the functions

$$z_N \doteq \frac{\phi_N'}{\phi_N}, \qquad\qquad z = \frac{\psi'}{\psi}.$$

Observe that z_N has bounded total variation, uniformly w.r.t. N, and provides a measurable solution to the boundary value problem

$$z_N' + z_N^2 + f(x,\phi_N) = N\sigma_N, \qquad\qquad z_N(0) = z_N(R) = 0. \tag{65}$$

On the other hand, the function $z = \psi'/\psi$ satisfies $z(0) = z(R) = 0$, together with

$$z' + z^2 + f(x,\phi_N) \ =\ N\sigma_N - f_\phi(x,\phi_N)\phi_N \ >\ N\sigma_N + c_f, \tag{66}$$

separately on the subintervals $[0,y[$ and $]y,R]$. Comparing (66) with (65), we conclude that

$$z(y-) - z(y+) \ \geq\ c_0\ >\ 0, \tag{67}$$

for some positive constant c, independent of N, y.
The function $\psi(\cdot) = K_N(\cdot,y)$ can now obtained as

$$\psi(x) \ =\ \begin{cases} A\cdot\exp\left(\displaystyle\int_0^x z(s)\,\mathrm{d}s\right) & \text{if } x\in[0,y],\\[2mm] B\cdot\exp\left(-\displaystyle\int_x^R z(s)\,\mathrm{d}s\right) & \text{if } x\in[y,R],\end{cases}$$

choosing the constants A, B so that $\psi(y+) = \psi(y-)$, $\psi'(y+) - \psi'(y-) = 1$. This leads to the linear algebraic system

$$\begin{cases} B\cdot\exp\left(-\displaystyle\int_x^R z(s)\,\mathrm{d}s\right) - A\cdot\exp\left(\displaystyle\int_0^x z(s)\,\mathrm{d}s\right) \ =\ 0,\\[2mm] B\cdot\exp\left(-\displaystyle\int_x^R z(s)\,\mathrm{d}s\right) z(y+) - A\cdot\exp\left(\displaystyle\int_0^x z(s)\,\mathrm{d}s\right) z(y-) \ =\ 1.\end{cases} \tag{68}$$

From the uniform bounds on z, and the lower bound (67), we conclude that the constants A, B in (68) remain uniformly bounded, for all N, y. This establishes the uniform bound on K_N, proving our claim. In turn, this implies (64).

(v) From the optimality conditions (48), we deduce $(q_N + 1)\phi_N - c \ \leq 0$ for all $x \in [0,R]$. Letting $N\to\infty$ we conclude

$$\phi_\infty(x) \ \leq\ \limsup_{N\to\infty}\ \phi_N(x) \ \leq\ \limsup_{N\to\infty}\Big(c(x) - q_N(x)\phi_N(x)\Big) \ =\ c(x)$$

for all $x \in [0, R]$. We have thus shown that ϕ_∞ is a subsolution of (57) which satisfies $0 < \phi_\infty(x) \leq c(x)$ for all $x \in [0, R]$. By the definition of ϕ^* at (60), this trivially implies that $\phi_\infty \leq \phi^*$, completing the proof.

□

6 Optimizing a Variational Inequality

Motivated by the above result, we now consider the problem of optimally designing a marine park, where $c = +\infty$, in such a way that the total catch is maximized. Calling $\P \subset [0, R]$ the open set where the park is located, it will be convenient to work with the complement $\Sigma \doteq [0, R] \setminus \P$. Given a closed set $\Sigma \subseteq [0, R]$, we thus consider the cost function

$$c(x, \Sigma) = \begin{cases} c(x) & \text{if} \quad x \in \Sigma\,, \\ +\infty & \text{if} \quad x \notin \Sigma\,, \end{cases}$$

and the domain

$$\mathscr{K}_\Sigma \doteq \Big\{\phi : [0, R] \mapsto I\!R, \qquad \phi \text{ is Lipschitz continuous with } \phi' \in BV,$$
$$\phi'(0+) \geq 0, \qquad \phi'(R-) \leq 0, \qquad \phi(x) \leq c(x, \Sigma) \;\; \forall x \in [0, R]\Big\}.$$

We now seek an optimal pair (ϕ, Σ), such that the integral $\int_0^R g(x, \phi)\,dx$ is maximized. Here, Σ ranges over all closed subsets of $[0, R]$, while $\phi \in \mathscr{K}_\Sigma$ provides a solution to the corresponding variational inequality

$$\int_0^R \phi'_\infty(\psi' - \phi'_\infty)\,dx + \int_0^R g(x, \phi_\infty) \cdot (\phi_\infty - \psi)\,dx \geq 0 \qquad \forall \psi \in \mathscr{K}_\Sigma. \quad (69)$$

Theorem 3. *In addition to the assumptions (A1)–(A2), let the cost function $c : [0, R] \mapsto I\!R_+$ be continuous. Then the one-dimensional optimization problem for the variational inequality has an optimal solution (ϕ, Σ).*

Proof. Let $\Sigma_n \subset [0, R]$ be a maximizing sequence of compact sets, and let ϕ_n be the corresponding solutions to the variational inequality, for $n \geq 1$. Since all the ϕ_n are uniformly Lipschitz continuous, by taking a subsequence we can assume that $\phi_n \to \phi$ uniformly on $[0, R]$. Moreover, we can assume that $\Sigma_n \to \Sigma$ in the Hausdorff metric [3], for some compact set $\Sigma \subseteq [0, R]$.

The uniform convergence $\phi_n \to \phi$ implies

$$\int_0^R g(x, \phi_n(x))\,dt \to \int_0^R g(x, \phi(x))\,dx \qquad \text{as } n \to \infty.$$

To conclude the proof, it suffices to show that ϕ provides the solution to the variational inequality (69), corresponding to the compact set $\Sigma \subseteq [0, R]$. Equivalently, we need to show that

(i) ϕ is a subsolution of (57)
(ii) $\phi(x) \leq c(x, \Sigma)$ for all $x \in [0, R]$
(iii) On the open set where $\phi(x) < c(x, \Sigma)$, the function ϕ satisfies

$$\phi''(x) + g(x, \phi(x)) = 0. \tag{70}$$

The property (i) follows from the uniform convergence $\phi_n \to \phi$, because each ϕ_n is a subsolution of (57).

To prove (ii), assume first $x \in \Sigma$. By the Hausdorff convergence $\Sigma_n \to \Sigma$, we can select points $x_n \in \Sigma_n$ such that $x_n \to x$ as $n \to \infty$. By the uniform Lipschitz continuity of the functions ϕ_n, and by the continuity of the cost function $c(\cdot)$, this yields

$$\phi(x) = \lim_{n\to\infty} \phi_n(x_n) \leq \lim_{n\to\infty} c(x_n) = c(x).$$

On the other hand, if $x \notin \Sigma$, one trivially has $\phi(x) < c(x, \Sigma) = +\infty$. This establishes (ii).

To prove (iii), assume $\phi(y) < c(y, \Sigma)$. We consider two cases. If $y \in \Sigma$, then $\phi(y) < c(y)$. By the uniform continuity and uniform convergence properties, we deduce $\phi_n(x) < c(x)$ for all n sufficiently large and every x in an open neighborhood $\mathcal{N}_y$ of y. Restricted to $\mathcal{N}_y$, all functions ϕ_n are solutions to the same equation (70). By the uniform convergence $\phi_n \to \phi$ we conclude that, on the open set $\mathcal{N}_y$, ϕ satisfies (70) as well.

On the other hand, if $y \notin \Sigma$, then we can find an open neighborhood $\mathcal{N}_y$ such that $\mathcal{N}_y \cap \Sigma_n = \emptyset$ for all n sufficiently large. In this case, each function ϕ_n satisfies (70) on $\mathcal{N}_y$, for n large enough. By the uniform convergence $\phi_n \to \phi$, we again conclude that, on the open set $\mathcal{N}_y$, ϕ satisfies (70) as well. This completes the proof. □

Example 5. On the domain $[0, R] \doteq [0, 1]$, assume $g(x, \phi) = (2-\phi)\phi$ and consider the cost function

$$c(x) \doteq \left(1 - \frac{x}{2}\right)\gamma \qquad x \in [0, 1]. \tag{71}$$

We claim that, if the constant $\gamma > 0$ is sufficiently small, then the choice $\Sigma = \{0\}$ is the unique optimal one. Indeed, it is clear that $\Sigma = \emptyset$ yields zero total catch, and cannot be optimal. If now $y \in \Sigma$, the corresponding solution ϕ^Σ of the variational inequality (69) will satisfy $\phi^\Sigma(x) \leq \phi^y(x)$, $y \in [0, 1]$, where ϕ^y denotes the solution to

$$\phi'' + g(x, \phi) = 0 \ \forall x \in]0, y[\cup]y, 1[,$$
$$\phi(y) = c(y) \ \ \phi'(0) = \phi'(1) = 0.$$

Notice that, by choosing the constant γ sufficiently small, we can achieve $0 < \phi^y(x) < 1$ for all $x, y \in [0, 1]$.

If Σ contains not only y but also additional points, then $0 < \phi^{\Sigma}(x) < \phi^{y}(x) < 1$ for all $x \neq y$. Hence

$$\int_0^1 \left(2 - \phi^{\Sigma}(x)\right) \phi^{\Sigma}(x)\,\mathrm{d}x \;<\; \int_0^1 (2 - \phi^{y}(x))\, \phi^{y}(x)\,\mathrm{d}x$$

showing that Σ cannot be optimal. By the above remarks, the optimal choice is restricted to singletons: $\Sigma = \{y\}$ for some $y \in [0,1]$. The particular form of the cost function $c(\cdot)$ implies that $\Sigma = \{0\}$ is the unique optimal strategy. This means that if γ in (71) is sufficiently small, in order to maximize the total catch the marine park should be $\P =]0,1]$, and fishing should be allowed only at the point $x = 0$.

Notice that the same conclusion remains valid for every strictly decreasing cost function $c : [0,1] \mapsto I\!R_+$, provided that $c(0)$ is sufficiently small.

Remark 1. In Theorem 3, the continuity assumption on the cost function $c(\cdot)$ is essential. Otherwise, a counterexample could be constructed as follows. With reference to Example 5 above, let us replace $c(\cdot)$ in (71) with the lower semicontinuous cost function

$$\tilde{c}(x) \doteq \begin{cases} \left(1 - \dfrac{x}{2}\right)\gamma & \text{if} \quad 0 < x \leq 1, \\[2ex] \dfrac{\gamma}{3} & \text{if} \quad x = 0. \end{cases} \tag{72}$$

Then, choosing

$$\Sigma_n \doteq \left\{\frac{1}{n}\right\}$$

we obtain a maximizing sequence where the total catch converges to the same maximum achieved in Example 5. However, when the cost is given by (72), the set $\Sigma = \{0\}$ is not optimal. In this case, the variational problem does not have any optimal solution.

Remark 2. In this chapter, we analyzed only problems in one-space dimension. However, we expect that Theorems 1 and 2 can be extended to multidimensional problems. On the other hand, Theorem 3 cannot have a direct counterpart valid in dimension $n \geq 2$. Indeed, the results in [10, 11] indicate that the multidimensional optimization problem is not well posed. To have the existence of an optimal solution, an additional cost term is needed. For example, one could add a penalization term proportional to the total length of the boundary $\partial\P$ of the marine park.

Acknowledgements The work of the first author was partially supported by NSF, with grant DMS-0807420. The work of the second author was partially supported by NSF grant DMS-0908047. This paper was written as part of the international research program on Nonlinear Partial Differential Equations at the Centre for Advanced Study at the Norwegian Academy of Science and Letters in Oslo during the academic year 2008–2009.

References

1. Adams, D.R., Lenhart, S.: An obstacle control problem with a source term. Appl. Math. Optim. **47**, 79–95 (2003)
2. Adams, D.R., Lenhart, S., Yong, J.: Optimal control of the obstacle for an elliptic variational inequality. Appl. Math. Optim. **38**, 121–140 (1998)
3. Aubin, J.P., Cellina, A.: Differential Inclusions. Springer, Berlin (1984)
4. Bergounioux, M.: Optimal control of semilinear Elliptic Obstacle Problem. J. Nonlinear Convex Anal. **3** (2002)
5. Bergounioux, M., Lenhart, S.: Optimal control of the obstacle in semilinear variational inequalities. Positivity **8**, 229–242 (2004)
6. Bergounioux, M., Mignot, F.: Optimal control of obstacle problems: existence of Lagrange multipliers. ESAIM Control Optim. Calc. Var. **5**, 45–70 (2000)
7. Bergounioux, M., Zidani, H.: Pontryagin maximum principle for optimal control of variational inequalities. SIAM J. Control Optim. **37**, 1273–1290 (1999)
8. Bressan, A., Shen, W.: Measure valued solutions for a differential game related to fish harvesting. SIAM J. Control Optim. **47**, 3118–3137 (2009)
9. Buttazzo, G., Dal Maso, G.: Shape optimization for Dirichlet problems: relaxed formulation and optimality conditions. Appl. Math. Optim. **23**, 17–49 (1991)
10. Cioranescu, D., Murat, F.: A strange term coming from nowhere. In: Topics in the mathematical modelling of composite materials. Progr. Nonlinear Differential Equations Appl. **31**, 45–93. Birkhäuser, Boston (1997)
11. Dal Maso, G., Murat, F., Orsina, L., Prignet, A.: Renormalized solutions of elliptic equations with general measure data. Ann. Sc. Norm. Super Pisa Cl. Sci. **28**, 741–808 (1999)
12. Neubert, M.G.: Marine reserves and optimal harvesting. Ecol. Lett. **6**, 843–849(2003)

Optimal Distributed Uplink Channel Allocation: A Constrained MDP Formulation

Vivek S. Borkar and Joy Kuri

Abstract Several users share a common channel for transmission, which has an average rate constraint. The packets not yet transmitted are queued. The problem of optimal channel allocation to minimize the average sum queue occupancy subject to this constraint splits into individual Markov decision processes (MDPs) coupled through the Lagrange multiplier for the rate constraint. This multiplier can be computed by an on-line stochastic gradient ascent in a centralized or distributed manner. This gives a stochastic dynamic version of Kelly's decomposition. A learning scheme is also presented.

1 Introduction

Figure 1 shows the system that we consider in this paper.

1.1 The Model

Our model can be described as follows. Further elaboration of the model appears after this description.

- A wireless channel is shared by K users.
- Each user wishes to transmit packets to a common "gateway" (like an Access Point or Base Station) that serves as a portal to the external world.

V.S. Borkar (✉)
School of Technology and Computer Science, Tata Institute of Fundamental Research,
Homi Bhabha Road, Mumbai 400005, India
e-mail: borkar@tifr.res.in

J. Kuri
Centre for Electronics Design and Technology, Indian Institute of Science,
Bangalore 560012, India
e-mail: kuri@cedt.iisc.ernet.in

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_21,

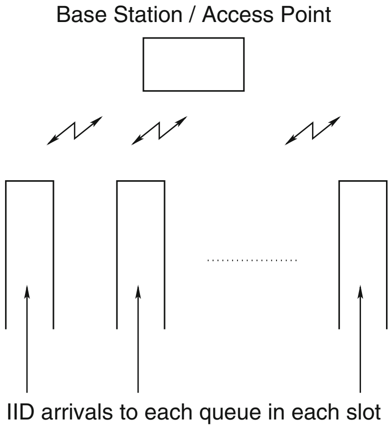


Fig. 1 K nodes share an uplink channel to a Base Station/Access Point. Time is slotted. Each queue sees independent and identically distributed arrivals in each slot

- Time is slotted, and the system is assumed to be synchronized; i.e., all users and the gateway identify slot boundaries correctly, and transmissions from users occur "simultaneously" (see remark below).
- Each user receives a packet stream which is independent and identically distributed (iid) across slots; the streams are also mutually independent across users, but the arrival statistics need not be identical for all users.
- Each user has a large but finite queue to store backlogged packets.
- Each user transmits a variable number of packets per slot, according to some policy, which is a map from the state of its queue to the number of packets to be transmitted.
- The average aggregate transmission rate (from all K users) is subject to a soft constraint imposed by the channel; we refer to this as the "communication constraint."
- The system objective is to minimize the long-run-average aggregate backlog in the users' queues, subject to the communication constraint.

Remark 1. Wireless channel sharing can be based on several multiple access principles: CDMA, FDMA, SDMA and TDMA are examples. In a slotted CDMA or FDMA or SDMA system, users are separated by codes/frequencies/directed antennas, and "simultaneous" transmissions are evident. Even a TDMA-based system fits into our model if we imagine a slot consisting of minislots that are used, one after the other, by the users. With this, from the perspective of a slot, the transmissions from the K users are effectively simultaneous. Thus, our model encompasses wireless channel sharing based on many multiple-access principles.

Because each user receives a random arrival stream, queues can build up occasionally. When this happens, a user would desire to transmit more packets in the slot by spending more power. This is precisely what can be done by exploiting the technique of *power control*. Hence, our assumption that a user can transmit a variable number of packets in a time slot is a very natural one.

The communication constraint is a consequence of the Shannon Theorem. In practice, a user will always have a constraint on the average power that can be spent. Then, Shannon's Theorem says that for reliable communication, the average aggregate transmission rate from the group of K users must be bounded above. We let R denote this upper bound. We note that imposing an upper bound on each *individual* user is unnecessarily restrictive. If this is done, then we run the risk of inefficient operation – a user would not be able to utilize transmission opportunities created by a temporary lack of data in some other users' buffers.

The system objective is natural, because by Little's Theorem, it corresponds to minimizing average packet transfer delays while respecting the communication constraint. This will lead to higher levels of overall user satisfaction.

1.2 Distributed Solutions

Evidently, the problem is for each user to decide how much to transmit such that the objective is achieved while respecting the communication constraint. In a centrally controlled system, it is conceivable that a central coordinator (like the gateway) would work out the amounts to be transmitted by each user, and convey the information to each. This would enable *optimal* solutions to be computed. However, the difficulty with this approach is that the coordinator needs to know the current queue lengths of the stations. In practice, the available queue length information is invariably old. Moreover, this approach can lead to an excessive computational burden on the coordinator. Besides, it does not use the computational capabilities of the stations, treating them as dumb devices.

For these reasons, it has been recognized that *distributed solutions* are preferable, which makes this a game problem in principle. The computational load is distributed across stations, and there is no central bottleneck. However, the disadvantage is that each station has limited information, *viz.,* its own queue length. The challenge, therefore, is to devise strategies that are based on limited information, and yet achieve what the centralized system would be able to.

About a decade ago, Kelly observed that a distributed solution as good as a centralized one is possible [18]. Since then, many researchers have explored this idea in a multitude of contexts. A brief discussion of the literature follows.

1.3 Literature Review

Research into control of wireless networks has a long history. The literature can be classified into two broad categories. The first category of papers focuses on random exogenous arrivals that are *not controllable*, and seeks scheduling and routing algorithms that ensure system stability for the largest set in the arrival rate space (called the "capacity region") [4, 10, 29, 30, 33]. In their seminal paper [33], Tassiulas and Ephremides considered a multihop packet radio network and showed that the "backpressure" algorithm achieved the capacity region. Several authors have studied the same problem on the downlink in cellular networks, and provided scheduling policies achieving the capacity region [4,29,30]. Without insisting on achieving the capacity region, [10] points out that good performance according to a range of criteria can be achieved while ensuring system stability. One common feature of the algorithms in all these papers is that they are *centralized*; this is natural for scheduling problems on the downlink of a cellular system, as the queue length information is available at the Base Station.

The second category of papers considers sources that are *controllable* and have infinite backlogs of data to send. Each source is equipped with a utility function, and the system objective is to maximize aggregate utility. This model appeared first in the seminal papers of Kelly [18, 19] for a wired network, and was followed by several others, including [22, 23] and [20]. The basic idea is that the network provides congestion signals in the form of "prices," and sources modify their data rates in response. The original problem is shown to decompose into several subproblems, *viz.,* flow control and routing, and the focus is on *distributed* algorithms to achieve the objective.

This programme was later carried out for wireless networks, where the additional aspect of wireless link *scheduling* appeared [11–15,21,31,32,35]. It has been shown that the problem of obtaining a distributed algorithm for optimal scheduling is a hard one.

The problem that we consider in this paper straddles the two groups above. We have uncontrollable arrivals, and look for distributed algorithms to achieve optimal system operation. There are two scenarios in which the setup shown in Fig. 1 can operate. In the *noncooperative* scenario, each user is concerned with himself/herself only, and wants to transmit as much of his/her own traffic as possible. To ensure that the communication constraint is respected, the user must be induced to "behave." This suggests that each user must be *charged* for packet transmissions, so that there is a disincentive for excessive transmissions. As mentioned above, this theme of *pricing* has been researched extensively over the last decade, following the important work of Kelly [18, 19]. In the noncooperative context, the price must be conveyed to each user by the gateway. We note that our model covers users dispersed over a large geographical area also, as in a wide-area cellular network, where the Base Station plays the role of the gateway.

In a cooperative context, the users form a "team" and are willing to work together to achieve the system objective while respecting the communication constraint. We will see that in this case, there is no need for a central coordinator (like the gateway) to supply prices to the users. Each user can work out the price for himse/herself, from the knowledge of the total number of transmissions from the team in a slot. Of course, this means that there must be a mechanism for each user to monitor the total number of transmissions in a slot.

Thus, in either context, users have incentives to limit transmission rates, so that the communication constraint can be respected.

In the related works [27] and [28], per slot (hard) rate constraints are used as in previous literature, and learning algorithms are proposed. Our use of average (soft) rate constraints has the advantage that it permits a Kelly-type decomposition and is also in tune with information theoretic usage.

2 System Model

Because the wireless channel is *shared* among the K stations, each station's data transfer rate is constrained. Let $R_k(t)$ denote the number of bits that station k transfers (according to some channel sharing policy) in time $\{0, 1, \ldots, t\}$, $1 \le k \le K$. We assume that the constraint takes the following form

$$\limsup_{t\to\infty} \frac{1}{t}\mathbb{E}^{\pi}_{\mathbf{x}_0} \sum_{k=1}^{K} R_k(t) \le R \tag{1}$$

for a specified $R > 0$, where π is the policy followed and $\mathbf{x}_0$ represents the initial state of the system.

Let $A_k(t)$ denote the number of arrivals to station k in slot t. These are assumed to be iid with $P\{A_k(t) = w_k\} = \tilde{p}^{(k)}(w_k)$. Let $U_k(t)$ be the number served by k in slot t. Then $R_k(t) = \sum_{s=0}^{t} U_k(s)$. We assume that arrivals in a slot are available for service only in the next slot. Let $X_k(t)$ denote the queue length at k at the beginning of slot t. Then, the dynamics of each station can be represented as

$$\begin{aligned} X_k(t+1) &= X_k(t) - U_k(t) + A_k(t) \\ U_k(t) &\le X_k(t) \end{aligned}$$

In this setting, we are interested in optimal operation of the system. Our objective is to ensure that the long-run average system population

$$\limsup_{T\to\infty} \frac{1}{T+1}\mathbb{E}^{\pi}_{\mathbf{x}_0} \sum_{t=0}^{T}\sum_{k=1}^{K} X_k(t)$$

is minimized.

Remark 2. We assume that the arrival rates into each queue are such that the system is stable under some policy.

$U_k(t)_{t=0,1,2,\ldots}$ represents a sequence of decisions that station k must make. So, we have a sequential decision problem, subject to the constraint in (1). Our assumptions allow us to pose the problem as a *constrained* Markov Decision Process (MDP) [3], described next.

3 The MDP

In this section, we describe the MDP.

3.1 Elements of the MDP

We assume that in each slot, the random variable $A_k(t)$ representing arrivals is nonnegative integer-valued, and so is the service variable $U_k(t)$, $0 \le U_k(t) \le X_k(t)$. These imply that the queue length variable $X_k(t)$ is nonnegative and integer-valued as well.

The elements of the MDP can be listed as follows:

(i) *State space:* $\mathbb{N}_+^K$. K-element vectors of queue lengths constitute the state space.

(ii) *Action space:* If the state is $\mathbf{x}$, then the action space is given by

$$\{\mathbf{u} \in \mathbb{N}_+^K : 0 \le u_k \le x_k, 1 \le k \le K\}$$

(iii) *One-step cost:* The nonnegative real-valued function $\mathbf{c}(\mathbf{x}(t), \mathbf{u}(t))$ is the one-step cost function. We write

$$\mathbf{c}(\mathbf{x}(t), \mathbf{u}(t)) = \sum_{k=1}^{K} c_k(x_k(t), u_k(t))$$

when c_k is the one-step cost function for the kth user. Later, we shall specialize to $c_k(x_k, u_k) = x_k$. Another example is $c_k(x_k, u_k) = x_k + f_k(u_k)$, where $f_k(.)$ is positive and strictly convex increasing, denoting power consumption for transmitting u_k packets.

(iv) *Transition law:* This is denoted as

$$p(\mathbf{X}_{k+1} = \mathbf{x}_{k+1} | \mathbf{X}_k = \mathbf{x}_k, \mathbf{U}_k = \mathbf{u}_k).$$

We write this as

$$\prod_{k=1}^{K} p^{(k)}(X_{k+1} = x_{k+1} | X_k = x_k, U_k = u_k)$$

to reflect decentralized operation.

(v) *Objective:* For a finite T, we define

$$J_T(\pi, \mathbf{x}_0) := \mathbb{E}_{\mathbf{x}_0}^{\pi} \sum_{t=0}^{T} c(\mathbf{X}(t), \mathbf{U}(t))$$

Then, the long-run average cost corresponding to policy π and starting state $\mathbf{x}_0$ is

$$J(\pi, \mathbf{x}_0) := \limsup_{T \to \infty} \frac{J_T(\pi, \mathbf{x}_0)}{T + 1}$$

Our objective is to find a policy π^* such that

$$J(\pi^*, \mathbf{x}_0) = \inf_{\pi \in \Pi} J(\pi, \mathbf{x}_0), \forall \mathbf{x}_0 \in \mathbb{R}_+^K$$

subject to the constraint

$$\limsup_{T\to\infty} \frac{1}{T+1} \mathbb{E}_{\mathbf{x}_0}^{\pi} \sum_{t=0}^{T} \sum_{k=1}^{K} U_k(t) \leq R$$

3.2 Optimal Policy: Existence and Optimality Equation

In [2], the authors prove the existence of an optimal stationary policy for a problem that is more general than the one considered here. The model in [2] is more general in three respects:

- The state space is not merely integer valued, but real valued.
- Arrivals are not iid, but Markovian.
- The wireless channel varies over time due to fading.

Fluid approximations to the queue length are employed. Consequently, the state space in [2] is a 3-tuple, consisting of the (fluid) queue length, the arrivals in the previous slot and the channel state.

In this paper, the state of the MDP is given by the K-vector of (integer) queue lengths. The wireless channel manifests itself by imposing an upper bound on the sum of transmission rates that can be achieved by the K stations together.

It can be seen that the state space and the one-step cost function in this problem are special cases of the corresponding elements in [2]. Consequently, the existence of an optimal stationary policy and the form of the Average Cost Optimality Equation (ACOE) follow by appealing to the results in [2].

Since this is a constrained MDP, the ACOE corresponds to the ACOE for the unconstrained MDP with $(\mathbf{c}(\mathbf{x}, \mathbf{u}) + \boldsymbol{\lambda}^T \mathbf{u})$ as the one-step cost, where $\boldsymbol{\lambda}$ is the vector of Lagrange multipliers [3].

4 The Linear Programming Approach

The approach to study MDPs via linear programming was initiated in [24]. For the long-run average cost criterion, it was developed further in [5] and [6]. Hernández-Lerma and Lasserre [16] and [17] provide extensive accounts of this approach.

For the long-run average cost criterion, one gets a primal linear program defined on a vector space of measures, and its dual defined on a vector space of functions.

4.1 The Linear Programs

In the linear programming viewpoint, the long-run average cost is looked upon as an inner product between a measure $\mu(\mathbf{x},\mathbf{u})$, on the state $\times$ action space, and the cost function $\mathbf{c}(\mathbf{x}(t),\mathbf{u}(t))$, which is a function on the state $\times$ action space, when both are expressed as vectors in view of their discrete domains:

$$\sum_{(\mathbf{x},\mathbf{u})} \mathbf{c}(\mathbf{x},\mathbf{u})\,\mu(\mathbf{x},\mathbf{u}).$$

Here, μ is the so-called ergodic occupation measure defined by

$$\mu(\mathbf{x},\mathbf{u}) := \pi(\mathbf{x})\,p(\mathbf{u}|\mathbf{x}),$$

where π is the stationary distribution of the Markov chain corresponding to a given policy, and $p(\mathbf{u}|\mathbf{x})$ is the conditional probability of choosing control $\mathbf{u}$ when in state $\mathbf{x}$. The long-run average cost problem is viewed as one of selecting the appropriate ergodic occupation measure μ, subject to the rate constraint, such that the above sum is minimized. We note that $\sum_{(\mathbf{x},\mathbf{u})} \mathbf{c}(\mathbf{x},\mathbf{u})\,\mu(\mathbf{x},\mathbf{u})$ is a linear functional on the space of measures on the state $\times$ action space.

The linear program may be written as follows

$$\min \textstyle\sum_{(\mathbf{x},\mathbf{u})} \mathbf{c}(\mathbf{x},\mathbf{u})\,\mu(\mathbf{x},\mathbf{u}) \tag{2}$$

subject to

$$\sum_{(\mathbf{y},\mathbf{u})} \mu(\mathbf{y},\mathbf{u})\,p(\mathbf{x}|(\mathbf{y},\mathbf{u})) = \sum_{\mathbf{u}} \mu(\mathbf{x},\mathbf{u}) \;\; \forall \;\; \mathbf{x} \tag{3}$$

$$\sum_{(\mathbf{x},\mathbf{u})} \mu(\mathbf{x},\mathbf{u}) \left(\sum_{k=1}^{K} u_k \right) \leq R \tag{4}$$

$$\sum_{(\mathbf{x},\mathbf{u})} \mu(\mathbf{x},\mathbf{u}) = 1 \tag{5}$$

$$\mu(\mathbf{x},\mathbf{u}) \geq 0 \;\; \forall \;\; (\mathbf{x},\mathbf{u}) \tag{6}$$

Equation (3) imposes a consistency condition on the measure μ, which, along with (5) and (6), characterizes it as an ergodic occupation measure. Equation (4) expresses the rate constraint. See [5, 6] for details of this formulation.

4.1.1 The Primal LP

As $\mathbf{c}(\mathbf{x},\mathbf{u}) = \sum_{k=1}^{K} c_k(x_k(t), u_k(t))$, the additive structure of the one-step cost can be exploited by considering the *marginals* of the measure μ instead of the joint measure μ. Let $\mu^{(k)}(.,.)$ denote the marginal of $\mu(.,.)$ obtained by retaining the arguments (x_k, u_k) and summing out the rest. Then, the LP given above can be written as

$$\min \sum_{k=1}^{K} c_k(x_k, u_k) \sum_{(x_k,u_k)} \mu^{(k)}(x_k, u_k)$$

subject to

$$\sum_{(y_k,u_k)} \mu^{(k)}(y_k, u_k) p^{(k)}(x_k|y_k, u_k) = \sum_{u_k} \mu^{(k)}(x_k, u_k), \forall x_k,\ 1 \le k \le K \quad (7)$$

$$-\sum_{k=1}^{K} \sum_{(x_k,u_k)} u_k\, \mu^{(k)}(x_k, u_k) \ge -R \quad (8)$$

$$\sum_{(x_k,u_k)} \mu^{(k)}(x_k, u_k) = 1,\ \ 1 \le k \le K \quad (9)$$

$$\mu^{(k)}(x_k, u_k) \ge 0\ \ \forall\ \ (x_k, u_k),\ \ 1 \le k \le K \quad (10)$$

Equation (7) is obtained from (3) by fixing (x_k, u_k) and y_k, and adding over all (x_j, u_j), $j \in \{1, 2, \ldots, K\}$, $j \neq k$. This is done for each k, yielding the K constraints in (7). The same argument is used to arrive at (9) and (10). Similarly, (8) results from (4). To express the primal LP in the so-called "symmetric" form, we multiply throughout by -1 and obtain the $\ge$ sign in (8); this is done so that the corresponding Lagrange multiplier (λ in the next subsection) is nonnegative.

4.1.2 The Dual LP

We will write down the dual of the LP in Sect. 4.1.1. For $1 \le k \le K$, let $-V^{(k)}(x_k)$ represent the Lagrange multipliers (dual variables) corresponding to the constraints in (7); $V^{(k)}(x_k)$ are unrestricted in sign. (Again, the negative sign in $-V^{(k)}(x_k)$ is used so that the dual program can be expressed in the form given below.) Let $\lambda \ge 0$ be the Lagrange multiplier corresponding to constraint (8), and let β_k represent the Lagrange multipliers corresponding to the constraints in (9); β_k are unrestricted in sign.

Then, after some transposition, the dual LP can be expressed as

$$\max \left(\sum_{k=1}^{K} \beta_k - \lambda R \right)$$

subject to

$$\beta_k + V^{(k)}(x_k) - \lambda u_k \le c_k(x_k, u_k) + \sum_{y_k} p^{(k)}\left(y_k|(x_k, u_k)\right) V^{(k)}(y_k), \forall (x_k, u_k),$$

$$1 \le k \le K \lambda \ge 0 \quad (11)$$

From (11), we observe that if λ is given, then the dual LP above *decouples* into *K distinct LPs*. Each LP corresponds to a distinct user, and can be solved independently. This suggests that a *distributed* solution to the MDP is possible if λ is made available to each user.

4.1.3 Discussion

This observation was made first by Kelly [18] in a different context. The problem in [18] was one of the aggregate utility maximizations, and a "static" or "one-shot" optimization problem was formulated. Then, standard results from the theory of convex optimization were used to show that the system optimization problem may be decomposed into smaller optimization problems, one for each user and one for the network, by using price per unit flow as a Lagrange multiplier.

Our approach demonstrates that the same programme can be carried out in a "dynamic" context also. Instead of the one-shot optimization problem in [18], we have a constrained sequential decision problem that extends over the infinite time horizon. Nevertheless, by viewing the constrained MDP as a linear programme in an appropriate space, we are able to recover the same basic insight that given the dual variable λ (which behaves like the price in [18]), the original control problem decouples into several smaller problems that are solvable independently.

4.2 Distributed Solution

Given λ, each user k solves the problem

$$\begin{aligned} &\max \beta_k \\ &\text{subject to} \\ &\beta_k + V^{(k)}(x_k) - \lambda u_k \le c_k(x_k, u_k) + \sum_{y_k} p^{(k)}\left(y_k|(x_k, u_k)\right) V^{(k)}(y_k),\ \forall (x_k, u_k) \end{aligned}$$

Taking λu_k to the right, we get the problem

$$\max \beta_k \tag{12}$$

$$\begin{aligned} &\text{subject to} \\ &\beta_k + V^{(k)}(x_k) \le (c_k(x_k, u_k) + \lambda u_k) + \sum_{y_k} p^{(k)}\left(y_k|(x_k, u_k)\right) V^{(k)}(y_k), \\ &\forall (x_k, u_k) \end{aligned} \tag{13}$$

Let us treat $(c_k(x_k, u_k) + \lambda u_k)$ as a "modified" one-step cost function. Then a solution to this problem corresponds to a solution to the ACOE [16] for a long-run average cost minimization problem, with $(c_k(x_k, u_k) + \lambda u_k)$ as the one-step cost. The ACOE is given by

$$\bar{\beta}_k + \bar{V}^{(k)}(x_k) = \min_{u_k <= x_k} \Big[c_k(x_k, u_k) + \lambda u_k + \sum_{y_k} p^{(k)}(y_k | (x_k, u_k)) \bar{V}^{(k)}(y_k) \Big]. \tag{14}$$

Here, $\bar{\beta}_k$ is uniquely characterized as the corresponding optimal cost and $\bar{V}^{(k)}(\cdot)$ is uniquely characterized up to an additive scalar. We render it unique by imposing $\bar{V}^{(k)}(y_{k0}) = \bar{\beta}_k$ for a prescribed y_{k0}.

4.3 Relative Value Iteration

User k can obtain a solution to his/her problem using *Relative Value Iteration (RVI)* [26]. The iteration is

$$\begin{aligned} & V_{n+1}^{(k)}(x_k) \\ & \quad = \min_{u_k \le x_k} \left(c_k(x_k, u_k) + \lambda u_k + \sum_{y_k} p^{(k)}(y_k | (x_k, u_k)) V_n^{(k)}(y_k) - V_n^{(k)}(y_{k0}) \right) \end{aligned} \tag{15}$$

with $V_0^{(k)}(.) \equiv 0$. As $n \to \infty$, $V_n^{(k)}(x_k)$ converges to $\bar{V}^{(k)}(x_k)$ in (14), and $V_n^{(k)}(y_{k0})$ converges to the solution $\bar{\beta}_k$ of the problem in (12) and (13), which is also the corresponding optimal cost. We set $\lambda = \lambda_n$, computed as described below.

4.4 Iteration for $\boldsymbol{\lambda}$

To obtain the optimal solution to the primal problem (which is our goal), we need to use the *optimal* dual variable λ^*. As this is not known a priori, it is necessary to consider an iterative scheme to produce a sequence of λ values that converge to λ^*.

Such a scheme is given by

$$\lambda_{n+1} = \Gamma \left(\lambda_n + b_n \left(\sum_{k=1}^{K} u_k^*(\lambda_n, X_k(n)) - R \right) \right), \tag{16}$$

where for $1 \le k \le K$, $u_k^*(\lambda_n, X_k(n))$ is the action that achieves the minimum in (15) with $\lambda = \lambda_n$, $\Gamma(.)$ is a projection to $[0, M]$ for some $M >> 1$ that serves to keep the λ iterates bounded and nonnegative, and b_n is a sequence of decreasing stepsizes. The sequence of decreasing step sizes b_n must satisfy

$$\sum_{n=1}^{\infty} b_n = \infty, \quad \sum_{t=1}^{\infty} b_n^2 < \infty$$

When the arrival statistics are known, the RHS in (15) can be computed for x_k, and the minimizing action $u_k^*(\lambda_n, x_k)$ obtained; $U_k(n) := u_k^*(\lambda_n, X_k(n))$ are the values used in (16).

Referring to the primal problem in (2)–(6), let

$$\phi(\mu, \lambda) := \sum_{(\mathbf{x},\mathbf{u})} \mathbf{c}(\mathbf{x}, \mathbf{u})\, \mu(\mathbf{x}, \mathbf{u}) + \lambda \sum_{(\mathbf{x},\mathbf{u})} \mu(\mathbf{x}, \mathbf{u}) \sum_{k=1}^{K} u_k$$

Then, we know that the optimal values μ^*, λ^* constitute a *saddle point* for ϕ, *viz.,*

$$\phi(\mu^*, \lambda) \leq \phi(\mu^*, \lambda^*) \leq \phi(\mu, \lambda^*) \tag{17}$$

Thus, given λ^*, the optimal μ is obtained by solving a minimization problem; similarly, given μ^*, the optimal λ is obtained by solving a maximization problem. For a given λ, the first (minimization) problem is precisely what RVI solves. The V iterations and the λ iterations are, however, coupled. Because $b_n \to 0$ as $n \to \infty$, it follows that in the coupled iterations (15)–(16), for large n, there is hardly any change in λ_n relative to the change in $V_n(.)$. This means that for large n, (15) sees λ_n as nearly a constant, while (16) sees $V_{n+1}^{(k)}(.)$ as having nearly equilibrated; i.e., roughly speaking, the latter is a good approximation to $V_{n+1}^{(k)*}(.)$ for the current λ_n. Thus, for n large enough, we have a close approximation to the optimal policy for the current λ_n.

The foregoing intuitive description can be made precise, as in Lemma 4.1 of [7] (see below).

Next, consider

$$G(\lambda) := \min_{\mu} \phi(\mu, \lambda)$$

The left inequality in (17) suggests that we should maximize $G(\lambda)$ with respect to λ, which, in turn, motivates an ascent algorithm for λ. This can be achieved by invoking the *Envelope Theorem* of mathematical economics [25] (also known as "Danskin's Theorem" in the convex case).

Lemma 1. *The ascent direction with respect to λ_n is given by*

$$\sum_{k=1}^{K} \sum_{(x_k, u_k)} u_k \mu^{*(k),\lambda_n}(x_k, u_k) - R,$$

where $\mu^{*(k),\lambda}$ is the ergodic occupation measure under $u_k^*(\lambda, \cdot)$.

Proof. This follows by taking the partial derivative of ϕ with respect to λ_n, and evaluating the result at the optimal primal variable $\mu^{*(k),\lambda_n}(.,.)$, $1 \le k \le K$. The relation then follows from the Envelope Theorem. □

We note that evaluating $\sum_{k=1}^{K} \sum_{(x_k,u_k)} \left(u_k \mu^{*(k),\lambda_n}(x_k,u_k) - R\right)$ actually corresponds to an *averaging* with respect to the ergodic occupation measure, *viz.,* $\mu^{*(k),\lambda_n}(.,.)$. This suggests that one can employ a Stochastic Approximation (SA) algorithm to update λ_n, taking advantage of the averaging property of SA. This leads directly to the iteration in (16). Again, this can be formalized as in Corollary 4.1 of [7].

The RVI in (15) proceeds on the "natural" timescale, i.e., the timescale on which the iterations, indexed by n, proceed. On the other hand, the iterations in (16) proceed on a timescale defined by the "time increments" $\{b_n\}$. In particular, $b_n \to 0$ as $n \uparrow \infty$, which makes this a slower timescale. The fact that the coupled iterations in (15) and (16) converge can now be established by the "two timescale" analysis as in [8], Chap. 6. We state this as a lemma.

Lemma 2. *The coupled iterations in* (15) *and* (16) *converge.*

Remark 3. Instead of the coupled iterations in (15) and (16), an alternative view is to hold λ_n fixed, run the iteration in (15) in situ till convergence, and then proceed to the next iteration of λ_n. This is possible in the *nonlearning case*, which we are considering at present, because the probabilities $p^{(k)}(y_k|(x_k,u_k))$ are known. Thus, this approach seeks to compute the value function by an "inner loop" first, and then proceeds to the "outer loop" involving the λ_n updates.

Later, in Sect. 6, we will consider the situation where the arrival statistics are not known, and must be *learned*. In that case, the view will be that the learning counterpart of the RVI in (15), which is another Stochastic Approximation, must run on a sufficiently fast timescale, and the "output" generated from it must be used in the iteration for λ that runs on a slower timescale.

5 Structural Properties of an Optimal Policy

Suppose λ is given and consider the problem that user k wishes to solve:

$$\begin{gathered}\max \beta_k \\ \text{subject to} \\ \beta_k + V^{(k)}(x_k) \le (c_k(x_k,u_k) + \lambda u_k) + \sum_{y_k} p^{(k)}(y_k|(x_k,u_k))\, V^{(k)}(y_k)\end{gathered}$$

Equivalently, consider the corresponding ACOE

$$\bar{\beta}_k + \bar{V}^{(k)}(x_k) = \min_{u_k}\left[c_k(x_k,u_k) + \lambda u_k + \sum_{y_k} p^{(k)}(y_k|(x_k,u_k))\bar{V}^{(k)}(y_k)\right]$$

From now, we will consider the special case

$$c_k(x_k, u_k) = x_k$$

to reflect the fact that our interest is in minimizing the average aggregate population in the system.

The results in [2] show that $V^{(k)}(x_k)$ is a convex increasing function in x_k. The optimal action in state k is given by the action that achieves the minimum in the following

$$\min_{0 \le u_k \le x_k} \left(\lambda u_k + \sum_{y_k} p^{(k)} \left(y_k | (x_k, u_k) \right) V^{(k)}(y_k) \right)$$

Using the fact that the arrival process into each queue is iid, the expression above simplifies to

$$\begin{aligned} &\min_{0 \le u_k \le x_k} \left(\lambda u_k + \sum_{w_k} \tilde{p}^{(k)}(w_k) V^{(k)}(x_k - u_k + w_k) \right) \\ &= \min_{0 \le u_k \le x_k} \left(\lambda u_k + G^{(k)}(x_k - u_k) \right), \end{aligned}$$

where we recall that $\tilde{p}^{(k)}(w)$ denotes the probability of w arrivals in a slot at user k, and

$$G^{(k)}(x_k - u_k) := \sum_{w_k} \tilde{p}^{(k)}(w_k) V^{(k)}(x_k - u_k + w_k).$$

As $V^{(k)}(.)$ is convex increasing in its argument, so is $G^{(k)}(.)$. Therefore, the problem above corresponds to finding the minimum of the convex function

$$F^{(k)}(u_k) := \lambda u_k + G^{(k)}(x_k - u_k)$$

over the compact set $0 \le u_k \le x_k$. Hence, the minimum will be *attained* at a point $\in [0, x_k]$.

As $F^{(k)}(u_k)$ is convex, and the feasible set for u_k is described by the linear equations $-u_k \le 0$ and $u_k \le x_k$, the Karush–Kuhn–Tucker conditions are both necessary and sufficient at the optimal point $u_k^{(*)}(x_k)$. The minimum is achieved at a point determined by the slope λ and the derivative $G^{(k)'}(.)$ together.

5.1 *Optimal Point in the Interior of* $[0, x_k]$

Lemma 3. $u_k^*(x_k)$ *will be attained in the interior of* $[0, x_k]$ *if and only if the following is satisfied*

$$G^{(k)'}(x_k - u_k^*(x_k)) = \lambda$$

Proof. Let $\eta_1 \geq 0$ and $\eta_2 \geq 0$ be the optimal Lagrange multipliers corresponding to the constraints $u_k \leq x_k$ and $-u_k \leq 0$, respectively. As neither constraint is active at the optimal point, we must have $\eta_1 = \eta_2 = 0$. Then the KKT gradient condition becomes

$$\lambda - G^{(k)'}(x_k - u_k^*(x_k)) = 0,$$

which yields the result. □

5.2 Optimal Point at x_k

Lemma 4. *The minimum is achieved at* x_k *if and only if there is* $\eta_1 \geq 0$ *such that*

$$G^{(k)'}(0) - \lambda = \eta_1$$
$$\text{i.e., iff} \quad G^{(k)'}(0) \geq \lambda$$

Proof. For this case, we have $\eta_2 = 0$ as the constraint $-u_k \leq 0$ is slack at the optimal point. Then the KKT gradient condition becomes

$$\lambda - G^{(k)'}(x_k - u_k^*(x_k)) + \eta_1 = 0$$

from which the result follows. □

5.3 Optimal Point at 0

Lemma 5. *The minimum is achieved at* 0 *if and only if there is* $\eta_2 \geq 0$ *such that*

$$\lambda - G^{(k)'}(x_k) = \eta_2$$
$$\text{i.e., iff} \quad G^{(k)'}(x_k) \leq \lambda$$

Proof. Here, we have $\eta_1 = 0$ as the constraint $u_k \leq x_k$ is slack at the optimal point. The KKT condition now leads to

$$\lambda - G^{(k)'}(x_k - u_k^*(x_k)) - \eta_2 = 0$$

from which the result follows. □

In Fig. 2, we provide a suggestive sketch of $G^{(k)'}(x_k - u_k^*(x_k))$ versus $(x_k - u_k^*(x_k))$. As $G^{(k)}(.)$ is convex increasing, $G^{(k)'}(.)$ is an increasing function. Depending on the value of λ, the optimal action $u_k^*(x_k)$ can be in one of three regimes.

Intuitively, these results make sense. When λ is very high, packet transmission is expensive, and hence the optimal action is to transmit nothing. Similarly, when λ is

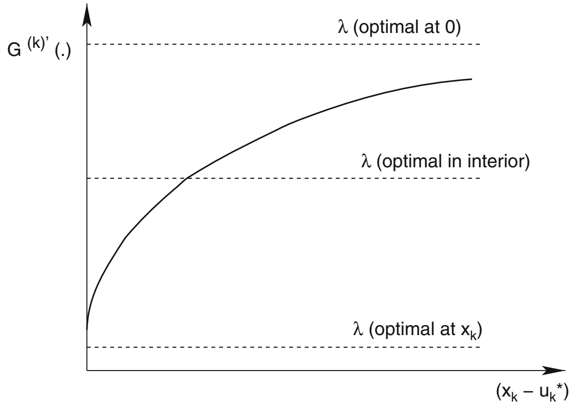


Fig. 2 Sketch of $G^{(k)'}(.)$ versus $(x_k - u_k^*(x_k))$. The three *dashed horizontal lines* represent the values of λ in the three cases

very small, packet transmission does not cost much, and it is meaningful to transmit the entire buffer contents. For intermediate values of λ, our analysis shows that the optimal action is to transmit the entire buffer contents *except for a constant amount* given by $q :=$ the solution to $G^{(k)'}(q) = \lambda$. Analogous analysis is possible for other related criteria.

6 Learning

So far, we have assumed that the statistics of the arrival process is known. This means that $p^{(k)}(y_k|(x_k, u_k))$ is known, and enables the right-hand side of the RVI iteration in (15) to be computed.

In this section, we relax this assumption. Evidently, this is what will occur in practice, because the arrival process statistics will be unknown in most cases. Our interest is in finding the best course of action under these circumstances.

The topic of "optimal adaptive control" deals with such problems, and it has a fairly long history. Following the standard approach, we will view the problem in which the statistics are *learned over time*; the learning occurs concurrently with the control actions chosen over time. Our scheme fits in the online version of Q-learning [34].

We will consider the "post-state" framework; see [8] for details. This framework can be utilized whenever a transition of the controlled Markov Chain can be *separated* into two parts, *viz.,* the effect of the control action and subsequently, the effect of the random event (like arrivals). In our case, a control action u_k results in the queue length changing from x_k to $(x_k - u_k)$; the queue length in the next slot, $(x_k - u_k + w_k)$, is obtained by adding the (random) number of arrivals w_k to it. $(x_k - u_k) := \tilde{x}_k$ is referred to as the "post-state", as it is the state after the control action has been applied, but before the random arrivals have occurred.

The RVI of (15) can be written in terms of the post-state as follows. Let $\tilde{V}_n^{(k)}(\tilde{x}_k)$ denote the value function for user k at iteration n. Then we have

$$\tilde{V}_{n+1}^{(k)}(\tilde{x}_k) = \sum_{w_k} \tilde{p}_{w_k} \min_{u_k \leq \tilde{x}_k + w_k} \Big(c_k(\tilde{x}_k + w_k, u_k) + \lambda_n u_k + \tilde{V}_n^{(k)}(\tilde{x}_k + w_k - u_k) - \tilde{V}_n^{(k)}(0) \Big), \tag{18}$$

where we have used λ_n in place of λ in (15) to indicate the fact that λ also changes over time, albeit slowly. In fact, for the value iteration in (18), λ_n behaves as if it were frozen because updates to λ_n occur on a much slower timescale. This suggests the "two-timescale" Stochastic Approximation framework [8].

The standard approach towards casting this in the Stochastic Approximation framework is to consider an *actual evaluation* of the term in brackets on the right-hand side of (18) at an observed transition, and then making an "incremental" move in this direction. We get

$$\begin{aligned}
\hat{V}_{n+1}^{(k)}(\tilde{x}_k) &= (1 - a_n I X_n^{(k)} = \tilde{x}_k)\hat{V}_n^{(k)}(\tilde{x}_k) \\
&\quad + a_n \Big[I\{X_n^{(k)} = \tilde{x}_k\} \times \min_{u_k \leq \tilde{x}_k + w_k} \Big(c_k(\tilde{x}_k + w_k, u_k) \\
&\quad + \lambda_n u_k + \hat{V}_n^{(k)}(\tilde{x}_k - u_k + w_k) - \hat{V}_n^{(k)}(0) \Big) \Big] \\
&= \hat{V}_n^{(k)}(\tilde{x}_k) + a_n I\{X_n^{(k)} = \tilde{x}_k\} \\
&\quad \times \Big[\min_{u_k \leq \tilde{x}_k + w_k} \Big(c_k(\tilde{x}_k + w_k, u_k) + \lambda_n u_k + \hat{V}_n^{(k)}(\tilde{x}_k - u_k + w_k) \\
&\quad - \hat{V}_n^{(k)}(0) \Big) - \hat{V}_n^{(k)}(\tilde{x}_k) \Big],
\end{aligned} \tag{19}$$

where $I\{A\}$ is the indicator of event A. The indicator $I\{X_n^{(k)} = \tilde{x}_k\}$ is present because the function value at $\tilde{x}_k$ is updated only if the state at the current time is $\tilde{x}_k$, not otherwise. Also, as we remarked before in the discussion following (16), λ_n will be updated on a slow timescale, according to

$$\lambda_{n+1} = \Gamma \left(\lambda_n + b_n \left(\sum_{k=1}^{K} u_k^*(\lambda_n, \tilde{X}_k(n)) - R \right) \right), \tag{20}$$

where b_n is a sequence of stepsizes satisfying $\sum_{n=1}^{\infty} b_n = \infty$, $\sum_{n=1}^{\infty} b_n^2 < \infty$, and *in addition*, $b_n = o(a_n)$. The last condition ensures that (20) moves on a slower timescale compared to (19).

Treating $\hat{V}_n^{(k)}(\tilde{x}_k)$ as a finite-dimensional vector that evolves in time, we can analyze (19), as in [8], by freezing λ_n to λ and considering the limiting o.d.e.

$$\dot{y}(t) = \Lambda(t)(h(y)(t) - y(t)). \tag{21}$$

Here, $y(t)$ represents the vector $\hat{V}_n^{(k)}(\tilde{x}_k)$, $\Lambda(t)$ is a diagonal matrix with non-negative elements summing to 1 on the diagonal and $h(.)$ is a function that acts on the vector $y(.)$ and maps it to $h(y)(.)$ given by

$$h(y)(x) = \sum_{w_k} \tilde{p}_{w_k} \left[\min_{u_k \leq x + w_k} (c_k(x + w_k, u_k) + \lambda_n u_k + y(x - u_k + w_k) - y(0)) \right].$$

Assuming that the elements of $\Lambda(t)$ remain bounded away from zero, it is possible to prove, by a routine adaptation of the arguments in [1], that the o.d.e. in (21) converges. A sufficient condition ensuring that the elements of $\Lambda(t)$ remain positive is the irreducibility of the Markov chain $X_n^{(k)}$ under all stationary policies. We state this as a lemma.

Lemma 6. *If the elements of $\Lambda(t)$ remain bounded away from 0, then $y(t)$ in* (21) *converges to the solution of the Average Cost Optimality Equation in* (14) *with λ frozen at λ_n.*

From this, the optimal stationary policy, for the current λ_n, can be obtained. However, as mentioned before (21), we have argued about the convergence of $y(t)$ assuming that λ_n was frozen, i.e., λ_n was assumed to be quasi-static. Now λ_n itself is updated on a slower timescale. To analyze the behaviour with respect to the evolution of λ, we note that what we want is to maximize the Lagrangean with respect to λ, which corresponds to an ascent in the dual (Lagrangean) variable space.

As we have seen in Sect. 4.4, the ascent direction with respect to λ_n is given by

$$\sum_{k=1}^{K} \sum_{(x_k, u_k)} u_k \mu^{*(k),\lambda_n}(x_k, u_k) - R$$

We note that, as expected, the ascent direction is a function of the queue lengths $\tilde{x}_k$, $1 \leq k \leq K$. If the optimal primal variable were known, we would have followed the ascent direction. As it is not known, we follow the latest "estimate" of the ascent direction, as in (20); the $u_k^*(\lambda_n, \tilde{x}_k)$ there are, of course, functions of $\tilde{x}_k$, $1 \leq k \leq K$. As before, Stochastic Approximation does the averaging.

Lemma 7. *The two-timescale coupled iterations in* (19) *and* (20) *converge.*

Proof. This can be established along the lines of [7]. □

Remark 4. We note that the timescales in (19) and (20) are determined by $a(n)$ and $b(n)$, respectively, with $b(n) = o(a(n))$ ensuring that (20) moves on a slower timescale with respect to (19). In Sects. 4.3 and 4.4, too, we had two timescales, but there (15) moved on the natural timescale, while (16) moved on a comparatively slower timescale. We note that in Sect. 4.4, $b(n) = o(1)$, which is a weaker requirement than $b(n) = o(a(n))$ (recall that $a(n) = o(1)$).

7 Conclusion

We recall that in the noncooperative context, the gateway (Base Station/Access Point) runs the λ-iteration, based on the amounts of traffic that each user has transmitted. Then, λ_n is broadcast to all users. In turn, each user k determines how much to transmit by determining which case in Fig. 2 applies (this is where the value of λ_n is used), and then transmitting the optimal number $u_k^*(\lambda_n, \tilde{X}_k(n))$ of packets.

The original problem for each user, mentioned in Sect. 5, was to decide how many packets to transmit (based on the current queue length) such that the long-run average aggregate population was minimized, subject to an upper bound on the aggregate packet transmission rate. We have seen how the gateway can provide a sequence of prices (the λ_n values) such that, while each user selfishly maximizes her own long-run average net benefit, the goal of optimal system operation is achieved.

In the cooperative context, the K users constitute a team. Let us assume that there is a mechanism for each user to monitor the shared wireless channel and deduce the total number of transmissions in a slot from the team. Such a mechanism is not hard to imagine; for example, in the TDMA system referred to in Sect. 1.1, each user merely needs to monitor every minislot and count the total number of packets transmitted in the minislot. Then, it is possible for each user to run the λ-iteration *herself* and work out the price λ_n, without relying on any central controller. Following this, each user in the team treats the λ_n as the price, and proceeds to maximize his/her average net benefit, as usual.

One can also allow channel measurements with additive zero mean noise. The noise will be averaged out by the stochastic approximation.

In the problem studied, all users were assumed to share a *single* wireless channel. It is straightforward to extend the model to multiple noninterfering wireless channels, with each channel being shared by several users. Every channel will give rise to one constraint on the aggregate rate of the associated user group (as in (1)), and this will lead to a corresponding Lagrange multiplier.

Further, in our problem, the state space was finite because we assumed finite buffers. It seems possible to extend the approach to countably infinite state spaces (corresponding to infinite buffers) by "function approximation" to approximate the value function by using a linear combination of a moderate number of basis functions. See, for example, [9]. In fact, this may be possible for *general* state spaces as well.

References

1. Abounadi, J., Bertsekas, D.P., Borkar, V.S.: Learning Algorithms for Markov Decision Processes with Average Cost. SIAM Journal on Control and Optimization **40**(3), 681–698 (2001)
2. Agarwal, M., Borkar, V.S., Karandikar, A.: Structural Properties of Optimal Tansmission Policies over a Randomly Varying Channel. IEEE Transactions on Automatic Control **53**(6), 1476–1491 (2008)

3. Altman, E.: Constrained Markov Decision Processes. Chapman and Hall/CRC, Boca Raton, Florida, USA (1999)
4. Andrews, M., Kumaran, K., Ramanan, K., Stolyar, A., Whiting, P., Vijayakumar, R.: Providing Quality of Service over a Shared Wireless Link. IEEE Communications Magazine **39**(2), 150–153 (2001)
5. Borkar, V.S.: Control of Markov Chains with Long-run Average Cost Criterion: the Dynamic Programming Equations. SIAM Journal on Control and Optimization **27**(3), 642–657 (1989)
6. Borkar, V.S.: Ergodic Control of Markov Chains with Constraints – the General Case. SIAM Journal on Control and Optimization **32**(1), 176–186 (1994)
7. Borkar, V.S.: An Actor-Critic Algorithm for Constrained Markov Decision Processes. Systems & Control Letters **54**(3), 207–213 (2005)
8. Borkar, V.S.: Stochastic Approximation: A Dynamic Viewpoint. Hindustan Book Agency, New Delhi, India *and* Cambridge University Press, Cambridge, UK (2008)
9. Borkar, V.S.: Reinforcement Learning: A Bridge Between Numerical Methods and Markov Chain Monte Carlo. In: N.S.N. Sastry, T.S.S.R.K. Rao, M. Delampady and B. Rajeev (ed.) Perspectives in Mathematical Sciences I: Probability and Statistics, Statistical Sciences and Interdisciplinary Research, vol. 7, chap. 4, pp. 71–91. World Scientific, Singapore (2009)
10. Buche, R., Kushner, H.J.: Control of Mobile Communication Systems with Time-varying Channels via Stability Methods. IEEE Transactions on Automatic Control **49**(11), 1954–1962 (2004)
11. Bui, L., Eryilmaz, A., Srikant, R., Wu, X.: Joint Asyncronous Congestion Control and Distributed Scheduling for Multihop Wireless Networks. In: Proceedings of INFOCOM (2006)
12. Chen, L., Low, S.H., Chiang, M., Doyle, J.C.: Cross-layer Congestion Control, Routing and Scheduling Design in Ad Hoc Wireless Networks. In: Proceedings of INFOCOM (2006)
13. Eryilmaz, A., Srikant, R.: Joint Congestion Control, Routing and MAC for Fairness and Stability in Wireless Networks. IEEE Journal on Selected Areas in Communications **24**(8), 1514–1524 (2006)
14. Eryilmaz, A., Srikant, R.: Fair Resource Allocation in Wireless Networks using Queue-Length-based Scheduling and Congestion Control. IEEE/ACM Transactions on Networking **15**(6), 1333–1344 (2007)
15. Gupta, A., Lin, X., Srikant, R.: Low Complexity Distributed Scheduling Algorithms for Wireless Networks. In: Proceedings of INFOCOM (2007)
16. Hernández-Lerma, O., Lasserre, J.B.: Discrete-Time Markov Control Processes. Springer, New York (1996)
17. Hernández-Lerma, O., Lasserre, J.B.: Further Topics in Discrete-Time Markov Control Processes. Springer, New York (1999)
18. Kelly, F.P.: Charging and Rate Control for Elastic Traffic. European Transactions on Telecommunications **8**(1), 33–37 (1997)
19. Kelly, F.P., Maulloo, A., Tan, D.: Rate Control for Communication Networks: Shadow Prices, Proportional Fairness and Stability. Journal of the Operational Research Society **49**(3), 237–252 (1998)
20. Kunniyur, S., Srikant, R.: End-to-end Congestion Control Schemes: Utility Functions, Random Losses and ECN marks. IEEE/ACM Transactions on Networking **11**(5), 689–702 (2003)
21. Lin, X., Shroff, N., Srikant, R.: A Tutorial on Cross-Layer Optimization in Wireless Networks. IEEE Journal on Selected Areas in Communications **24**(8), 1452–1463 (2006)
22. Low, S.H.: A Duality Model of TCP and Queue Management Algorithms. IEEE/ACM Transactions on Networking **11**(4), 525–536 (2003)
23. Low, S.H., Lapsley, D.: Optimization Flow Control-I: Basic Algorithm and Convergence. IEEE/ACM Transactions on Networking **7**(6), 861–874 (1999)
24. Manne, A.: Linear Programming and Sequential Decisions. Management Science **6**(3), 259–267 (1960)
25. Milgrom, P., Segal, I.: Envelope Theorems for Arbitrary Choice Sets. Econometrica **70**(2), 583–601 (2002)
26. Puterman, M.L.: Markov Decision Processes. Wiley-Interscience, Hoboken, New Jersey (1994)

27. Salodkar, N., Bhorkar, A., Karandikar, A., Borkar, V.S.: An Online Learning Algorithm for Energy Efficient Delay Constrained Scheduling over a Fading Channel. IEEE Journal on Selected Areas in Communications **26**(4), 732–742 (2008)
28. Salodkar, N., Karandikar, A., Borkar, V.S.: A Stable Online Algorithm for Energy-Efficient Multiuser Scheduling. IEEE Transactions on Mobile Computing **9**(10), 1391–1406 (2010)
29. Shakkottai, S., Stolyar, A.: Scheduling of a Shared Time-Varying Channel: the Exponential Rule Stability. In: Proceedings of INFORMS Applied Probability Conference (2001)
30. Shakkottai, S., Stolyar, A.: Scheduling for Multiple Flows sharing a Time-varying Channel: the Exponential Rule. Translations of the American Mathematical Society **207** (a volume in memory of F. Karpelevich) (2002)
31. Stolyar, A.L.: Maximizing Queueing Network Utility subject to Stability: Greedy Primal-Dual Algorithm. Queueing Systems: Theory and Applications (QUESTA) **50**(4), 401–457 (2005)
32. Tan, L., Zhang, X., Andrew, L.L.H., Zukerman, M.: Price-based Max-Min Fair Rate Allocation in Wireless Multi-hop Networks. IEEE Communications Letters **10**(1), 31–33 (2006)
33. Tassiulas, L., Ephremides, A.: Stability Properties of Constrained Queueing Systems and Scheduling Policies for Maximum Throughput in Multihop Radio Networks. IEEE Transactions on Automatic Control **37**(12), 1936–1948 (1992)
34. Watkins, C.: Learning from Delayed Rewards. Ph.D. thesis, Cambridge University, UK (1989)
35. Wu, X., Srikant, R., Perkins, J.R.: Queue-Length Stability of Maximal Greedy Schedules in Wireless Networks. IEEE Transactions on Mobile Computing **6**(6), 595–605 (2007)

Binomial Approximations for Barrier Options of Israeli Style

Yan Dolinsky and Yuri Kifer

Abstract We show that prices and shortfall risks of game (Israeli) barrier options in a sequence of binomial approximations of the Black–Scholes (BS) market converge to the corresponding quantities for similar game barrier options in the BS market with path dependent payoffs and we estimate the speed of convergence. The results are also new for usual American style options and they are interesting from a computational point of view, since in binomial markets these quantities can be obtained via dynamic programming algorithms. The paper extends [6] and [3] but requires substantial additional arguments in view of peculiarities of barrier options which, in particular, destroy the regularity of payoffs needed in the above papers.

1 Introduction

This paper deals with knock–out double barrier options of the game (Israeli) type sold in a standard securities market consisting of a nonrandom component b_t representing the value of a savings account at time t with an interest rate r and of a random component S_t representing the stock price at time t. As usual, we view $S_t, t > 0$ as a stochastic process on a probability space $(\Omega, \mathscr{F}, P)$ and we assume that it generates a right continuous filtration $\{\mathscr{F}_t\}$. The setup also includes two right continuous with left limits (*cadlag*) stochastic payoff processes $X_t \geq Y_t \geq 0$ adapted to the above filtration. Recall that a game contingent claim (GCC) or game option was defined in [5] as a contract between the seller and the buyer of the option such that both have the right to exercise it any time up to a maturity date (horizon) T which in this paper is assumed to be finite. If the seller exercises at a stopping $\sigma \leq T$ and a buyer at a stopping time $\tau \leq T$, then the former pays to the latter the amount $H(\sigma, \tau) = X_\sigma \mathbb{I}_{\sigma<\tau} + Y_\tau \mathbb{I}_{\tau \leq \sigma}$ where we set $\mathbb{I}_A = 1$ if an event A occurs and $\mathbb{I}_A = 0$ if not.

Y. Dolinsky (✉) and Y. Kifer
Institute of Mathematics, The Hebrew University, Jerusalem 91904, Israel
e-mail: yann1@math.huji.ac.il; kifer@math.huji.ac.il

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_22,

A hedge (for the seller) against a GCC is defined here as a pair (π, σ) which consists of a self-financing strategy π and a stopping time σ, which is the cancellation time for the seller. A hedge is called perfect if, no matter what exercise time the buyer chooses, the seller can cover his liability to the buyer (with probability one). The option price $\mathscr{V}^*$ is defined as the minimal initial capital required for a perfect hedge. Recall (see [5]) that pricing a GCC in a complete market leads to the value of a zero sum optimal stopping (Dynkin's) game (see [7]) with discounted payoffs $\tilde{X}_t = b_0 \frac{X_t}{b_t}$, $\tilde{Y}_t = b_0 \frac{Y_t}{b_t}$ considered under the unique martingale measure $\tilde{P} \sim P$.

We consider a double knock–out barrier option with a two constant barriers L, R such that $0 \leq L < S_0 < R \leq \infty$ which means that the option is worthless to its holder (buyer) at the first time τ_I the stock price S_t exits the open interval $I = (L, R)$. Thus for $t \geq \tau_{(L,R)}$ the payoff is $X_t = Y_t = 0$. For $t < \tau_{(L,R)}$ we consider path dependent payoffs. Such a contract is of potential value to a buyer who believes that the stock price will not exit the interval I up to a maturity date and to a seller who believes otherwise and does not want to have to worry about hedging if the stock price reaches one of the barriers L, R.

The Cox, Ross and Rubinstein (CRR) binomial model which was introduced in [1] is an efficient tool to approximate derivative securities in a Black–Scholes (BS) market. We will show that for a double barrier options in the BS model the option price can be approximated by a sequence of option prices for a barrier option (with the same barriers) in appropriate CRR n–step models with errors bounded by $Cn^{-1/4}(\ln n)^{3/4}$ where C is a constant which does not depend on the value of the barriers. These results provide an algorithm for the computation of this important class of derivative securities since pricing of game options in CRR markets can be done by dynamic programming (see [5]). Binomial approximations of barrier options were studied only for European options (see [8]).

We also deal with partial hedging which becomes relevant if the initial capital of the seller is less than the option price. In this case portfolio shortfall comes into the picture and for this reason we distinguish here between hedges and perfect hedges. In this paper we deal with a certain type of risk called the shortfall risk which was defined for game options in [2]. An investor (seller) whose initial capital x is less than the option price still wants to compute the minimum possible shortfall risk and to find a hedge with the initial capital x which minimizes or "almost" minimizes the shortfall risk.

In [2] we proved that for a game option in the multinomial model with general payoffs there exists a hedge which minimizes the shortfall risk under constraints on the initial capital, and the above hedge together with the corresponding shortfall risk can be computed via a dynamic programming procedure. We will prove that in the BS model the shortfall risk $R(x)$ of a seller with initial capital x for double barrier options is a limit of the shortfall risks $R_n(x)$ for double barrier options in the CRR market with the same barriers and initial capital as in the BS model. For a given initial capital x we will use hedges which minimize the shortfall risk in CRR markets under the above constraint on the initial capital, in order to construct hedges which "almost" minimize the shortfall risk in the BS model under the same constraint on the initial capital. Furthermore we will see that the corresponding portfolios are managed on a finite set of random times.

We also consider another situation where the seller of a game option in the BS model has an initial capital which is a little bit larger than the option price. In this case we use perfect hedges in CRR markets (which was constructed explicitly in [5]) in order to build explicitly hedges with small shortfall risks in the BS model where the corresponding portfolios are managed on a finite set of random times.

The main results of this paper are formulated in the next section where we also discuss the Skorohod type embedding. In Sect. 3 we introduce recursive formulas which enable us to compare various option prices and risks and we derive auxiliary estimates for option prices and risks. In Sect. 4 we complete the proof of the main results of the paper. More details can be found in [3] and [6].

2 Preliminaries and Main Results

First, we describe the setup. Denote by $M[0,t]$ the space of Borel measurable functions on [0,t] with the uniform metric $d_{0t}(\upsilon,\tilde{\upsilon}) = \sup_{0\le s\le t} |\upsilon_s - \tilde{\upsilon}_s|$. For each $t > 0$ let F_t and Δ_t be nonnegative functions on $M[0,t]$ such that for some constant $\mathscr{L} \ge 1$ and for any $t \ge s \ge 0$ and $\upsilon, \tilde{\upsilon} \in M[0,t]$ we have the following, $|F_s(\upsilon) - F_s(\tilde{\upsilon})| + |\Delta_s(\upsilon) - \Delta_s(\tilde{\upsilon})| \le \mathscr{L}(s+1)d_{0s}(\upsilon,\tilde{\upsilon})$ and $|F_t(\upsilon) - F_s(\upsilon)| + |\Delta_t(\upsilon) - \Delta_s(\upsilon)| \le \mathscr{L}(|t-s|(1+\sup_{u\in[0,t]}|\upsilon_u|) + \sup_{u\in[s,t]}|\upsilon_u - \upsilon_s|)$. Next we consider a complete probability space $(\Omega_B, \mathscr{F}^B, P^B)$ together with a standard one-dimensional continuous in time Brownian motion $\{B_t\}_{t=0}^{\infty}$, and the filtration $\mathscr{F}_t^B = \sigma\{B_s | s \le t\}$. A BS financial market consists of a savings account and a stock whose prices b_t and S_t^B at time t, respectively, are given by the formulas $b_t = b_0 e^{rt}$ and $S_t^B = S_0 e^{rt+\kappa B_t^*}$, where $B_t^* = (\frac{\mu}{\kappa} - \frac{\kappa}{2})t + B_t$, $r > 0$ is the interest rate, $\kappa > 0$ is the volatility and μ is a parameter. Denote by $\tilde{S}_t^B = e^{-rt} S_t^B$ the discounted stock price.

For any open interval $I = (L,R)$ such that $0 \le L < S_0 < R \le \infty$ let $\tau_I = \inf\{t \ge 0 | S_t^B \notin I\}$ be the first time the stock price exits from the interval I. Clearly τ_I is a stopping time (not necessary finite since we allow the cases $L=0$ and $R=\infty$). In this paper we assume that either $L > 0$ or $R < \infty$ while the case $L = 0$ and $R = \infty$ of regular options is treated in [6] and [3]. Consider a game option with the payoffs $Y_t^I = F_t(S^B)\mathbb{I}_{t<\tau_I}$ and $X_t^I = G_t(S^B)\mathbb{I}_{t<\tau_I}$ where $G_t = F_t + \Delta_t$, $S^B = S^B(\omega) \in M[0,\infty)$ is a random function taking the value $S_t^B = S_t^B(\omega)$ at $t \in [0,\infty)$. When considering $F_t(S^B)$, $G_t(S^B)$ for $t < \infty$ we take the restriction of S^B to the interval $[0,t]$. Denote by T the horizon of our game option assuming that $T < \infty$. Observe that the contract is "knocked–out" (i.e. becomes worthless to the buyer) at the first time that the stock price exits from the interval I. The discounted payoff function is given by $Q^{B,I}(s,t) = \tilde{X}_s^I \mathbb{I}_{s<t} + \tilde{Y}_t^I \mathbb{I}_{t\le s}$, where $\tilde{Y}_t^I = e^{-rt} Y_t^I$ and $\tilde{X}_t^I = e^{-rt} X_t^I$ are the discounted payoffs. Among examples of barrier options which fit our setup are put or call barrier options, Russian barrier options and Asian barrier options.

Denote by $\tilde{P}^B$ the unique martingale measure for the BS model. Using standard arguments it follows that the restriction of the probability measure $\tilde{P}^B$ to the σ–algebra $\mathcal{F}_t^B$ satisfies $Z_t = \frac{dP^B}{d\tilde{P}^B}|\mathcal{F}_t^B = e^{\frac{\mu}{\kappa}B_t + \frac{1}{2}(\frac{\mu}{\kappa})^2 t}$. Denote by $\mathcal{T}^B$ the set of all stopping times with respect to the Brownian filtration $\mathcal{F}_t^B, t \geq 0$ and let $\mathcal{T}_{0T}^B$ be the set of all stopping times with values in $[0, T]$. From Theorem 3.1 in [5] we obtain the fair price of a game option in the BS model is given by $\mathcal{V}^I = \inf_{\sigma \in \mathcal{T}_{0T}^B} \sup_{\tau \in \mathcal{T}_{0T}^B} \tilde{E}^B Q^{B,I}(\sigma, \tau)$ where $\tilde{E}^B$ is the expectation with respect to $\tilde{P}^B$.

Recall, (see, for instance, [9]) that a self-financing strategy π with a (finite) horizon T and an initial capital x is a process $\pi =|, \{(\beta_t, \gamma_t)\}_{t=0}^T$ of pairs where β_t and γ_t are progressively measurable with respect to the filtration $\mathcal{F}_t^B$, $t \geq 0$ and satisfy $\int_0^T e^{rt}|\beta_t|dt < \infty$ and $\int_0^T (\gamma_t S_t^B)^2 dt < \infty$. The portfolio value V_t^π for a strategy π at time $t \in [0, T]$ satisfies

$$
\begin{aligned}
V_t^\pi &= \beta_t b_t + \gamma_t S_t^B = x + \int_0^t \beta_u db_u + \int_0^t \gamma_u dS_u^B, \\
\tilde{V}_t^\pi &= x + \int_0^t \gamma_u d\tilde{S}_u^B \text{ and } \beta_t = (x + \int_0^t \gamma_u d\tilde{S}_u^B - \gamma_t \tilde{S}_t^B)/b_0
\end{aligned}
\tag{1}
$$

where $\tilde{V}_t^\pi = e^{-rt} V_t^\pi$ is the discounted portfolio value at time t. A self-financing strategy π is called *admissible* if $V_t^\pi \geq 0$ for all $t \in [0, T]$ and the set of such strategies with an initial capital x will be denoted by $\mathcal{A}^B(x)$. Set also $\mathcal{A}^B = \bigcup_{x \geq 0} \mathcal{A}^B(x)$. A pair $(\pi, \sigma) \in \mathcal{A}^B \times \mathcal{T}_{0T}^B$ of an *admissible* self-financing strategy π and of a stopping time σ will be called a hedge. The shortfall risk is defined by (see [2]),

$$
\begin{aligned}
R^I(\pi, \sigma) &= \sup_{\tau \in \mathcal{T}_{0T}^B} E^B[(Q^{B,I}(\sigma, \tau) - \tilde{V}_{\sigma \wedge \tau}^\pi)^+], \\
R^I(\pi) &= \inf_{\sigma \in \mathcal{T}_{0T}^B} R^I(\pi, \sigma) \text{ and } R^I(x) = \inf_{\pi \in \mathcal{A}^B(x)} R^I(\pi)
\end{aligned}
$$

where E^B denotes the expectation with respect to the market probability measure P^B. Thus for a hedge, the shortfall risk is defined as the maximal expectation (with respect to the buyer stopping strategies) of the discounted shortfall.

As in [3] and [6] we consider a sequence of CRR markets on a complete probability space such that for each $n = 1, 2, \ldots$ the bond prices $b_t^{(n)}$ at time t are $b_t^{(n)} = b_0 e^{r[nt/T]T/n} = b_0(1 + r_n)^{[nt/T]}$ $(r_n = e^{rT/n} - 1)$ and stock prices $S_t^{(n)}$ at time t are given by the formulas $S_t^{(n)} = S_0$ for $t \in [0, T/n)$ and $S_t^{(n)} = S_0 \prod_{k=1}^{[nt/T]}(1 + \rho_k^n)$ if $t \geq T/n$ where $\rho_k^n = \exp(\frac{rT}{n} + \kappa(\frac{T}{n})^{1/2}\xi_k) - 1$ and $\xi_1, \xi_2, \ldots$ are i.i.d. random variables taking values 1 and -1 with probabilities $p^{(n)} = (\exp((\kappa - \frac{2\mu}{\kappa})\sqrt{\frac{T}{n}}) + 1)^{-1}$ and $1 - p^{(n)} = (\exp((\frac{2\mu}{\kappa} - \kappa)\sqrt{\frac{T}{n}}) + 1)^{-1}$, respectively. Let $P_n^\xi = \{p^{(n)}, 1 - p^{(n)}\}^\infty$ be the corresponding product probability measure on the space of sequences $\Omega_\xi = \{-1, 1\}^\infty$. Namely, for each n we consider a CRR market with horizon n on the probability space $(\Omega_\xi,\ P_n^\xi)$ with bond prices $b_m = b^{(n)}_{\frac{mT}{n}}$ and stock prices $S_m = S^{(n)}_{\frac{mT}{n}}$. We view $S^{(n)} = S^{(n)}(\omega)$ as a random function on [0,T], so that $S^{(n)}(\omega) \in M[0,T]$ takes the value $S^{(n)}_t = S^{(n)}_t(\omega)$ at $t \in [0,T]$. For $k \leq n$ denote the discounted stock price at the moment kT/n by $\tilde{S}^{(n)}_{\frac{kT}{n}} = (1+r_n)^{-k} S^{(n)}_{\frac{kT}{n}}$. Let $\mathscr{F}^\xi_k = \sigma\{\xi_1,\dots,\xi_k\}$ and $\mathscr{F}^\xi = \sigma\{\xi_1,\xi_2\dots\}$. Denote by $\mathscr{T}^\xi$ the set of all stopping times with respect to the filtration $\mathscr{F}^\xi_k$ and let $\mathscr{T}^\xi_{0n}$ be the set of all stopping times with values in $\{0,1,\dots,n\}$. Given an open interval I introduce a stopping time (with respect to the filtration $\{\mathscr{F}^\xi_k\}_{k=0}^\infty$), $\tau^{(n)}_I = \min\{k \geq 0 | S^{(n)}_{\frac{kT}{n}} \notin I\}$ together with barrier options having the payoffs $Y^{I,n}_k = F_{\frac{kT}{n}}\left(S^{(n)}\right)\mathbb{I}_{k<\tau^{(n)}_I}$ and $X^{I,n}_k = G_{\frac{kT}{n}}\left(S^{(n)}\right)\mathbb{I}_{k<\tau^{(n)}_I}$. The corresponding discounted payoff function is given by $Q^{I,n}(s,k) = \tilde{X}^{I,n}_s \mathbb{I}_{s<k} + \tilde{Y}^{I,n}_k \mathbb{I}_{k\leq s}$, $k,s \leq n$ where $\tilde{X}^{I,n}_k = (1+r_n)^{-k} X^{I,n}_k$ and $\tilde{Y}^{I,n}_k = (1+r_n)^{-k} Y^{I,n}_k$ are the discounted payoffs. Let $\tilde{P}^\xi_n$ be a probability measure on the Ω_ξ such that $\xi_1,\xi_2,\dots$ is a sequence of i.i.d. random variables taking on the values 1 and -1 with probabilities $\tilde{p}^{(n)} = \left(\exp(\kappa\sqrt{\frac{T}{n}})+1\right)^{-1}$ and $1-\tilde{p}^{(n)} = \left(\exp(-\kappa\sqrt{\frac{T}{n}})+1\right)^{-1}$, respectively (with respect to $\tilde{P}^\xi_n$). Observe that for any n the process $\{\tilde{S}^{(n)}_{\frac{mT}{n}}\}_{m=0}^n$ is a martingale with respect to $\tilde{P}^\xi_n$, and so we conclude that $\tilde{P}^\xi_n$ is the unique martingale measure for the above CRR markets. Thus from Theorem 2.1 in [5] it follows that the fair price of the game option in the n–step CRR market is given by $\mathscr{V}^I_n = \min_{\zeta\in\mathscr{T}^\xi_{0n}} \max_{\eta\in\mathscr{T}^\xi_{0n}} \tilde{E}^\xi_n Q^{I,n}(\zeta,\eta)$, where $\tilde{E}^\xi_n$ is the expectation with respect to $\tilde{P}^\xi_n$. The following theorem provides an estimate for the error term in approximations of the fair price of a knock–out game option in the BS model by fair prices of the sequence of knock out game options in the CRR markets defined above.

Theorem 1. *There exists a constant C_1 such that for any open interval I and $n \in \mathbb{N}$,* $|\mathscr{V}^I - \mathscr{V}^I_n| \leq C_1 n^{-\frac{1}{4}}(\ln n)^{\frac{3}{4}}$.

Denote by $\mathscr{A}^{\xi,n}(x)$ the set of all *admissible* self-financing strategies with an initial capital x and set $\mathscr{A}^{\xi,n} = \bigcup_{x\geq 0}\mathscr{A}^{\xi,n}(x)$. A self-financing strategy π with an initial capital x and a horizon n is a sequence $(\pi_1,\dots,\pi_n)$ of pairs $\pi_k = (\beta_k,\gamma_k)$ where β_k,γ_k are $\mathscr{F}^\xi_{k-1}$-measurable random variables representing the number of bond and stock units, respectively, at time k. Thus (see [9]), the discounted portfolio value $\tilde{V}^\pi_k = (1+r_n)^{-k} V^\pi_k$, $k = 0,1,\dots,n$ satisfies

$$\tilde{V}^\pi_k = x + \sum_{i=0}^{k-1} \gamma_{i+1}\left(\tilde{S}^{(n)}_{\frac{(i+1)T}{n}} - \tilde{S}^{(n)}_{\frac{iT}{n}}\right)$$
$$\text{and } \beta_k = \left(x + \sum_{i=0}^{k-1}\gamma_{i+1}\left(\tilde{S}^{(n)}_{\frac{(i+1)T}{n}} - \tilde{S}^{(n)}_{\frac{iT}{n}}\right) - \gamma_k \tilde{S}^{(n)}_{\frac{kT}{n}}\right) \Big/ b_0. \tag{2}$$

We call a self-financing strategy π *admissible* if $V_k^\pi \geq 0$ for any $k \leq n$. A hedge with an initial capital x is an element in the set $\mathscr{A}^{\xi,n}(x) \times \mathscr{T}_{0n}^\xi$. The definitions for the shortfall risks in the CRR markets are similar to the definitions in the BS model. Thus for the n–step CRR market the shortfall risks are given by

$$R_n^I(\pi,\sigma) = \max_{\tau \in \mathscr{T}_{0n}^\xi} E_n^\xi \left(Q^{I,n}(\sigma,\tau) - \tilde{V}_{\sigma\wedge\tau}^\pi \right)^+,$$
$$R_n^I(\pi) = \min_{\sigma \in \mathscr{T}_{0n}^\xi} R_n^I(\pi,\sigma) \text{ and } R_n^I(x) = \inf_{\pi \in \mathscr{A}^{\xi,n}(x)} R_n^I(\pi) \tag{3}$$

where E_n^ξ is the expectation with respect to P_n^ξ.

Theorem 2. *For any open interval I, $lim_{n\to\infty} R_n^I(x) = R^I(x)$. Furthermore, there exists a constant C_2 (which does not depend on the interval I) such that for any $n \in \mathbb{N}$*

$$R^I(x) \leq R_n^I(x) + C_2 n^{-\frac{1}{4}} (\ln n)^{3/4}. \tag{4}$$

The above result says that the shortfall risk $R^I(x)$ for double barrier options in the BS model can be approximated by a sequence of shortfall risks with an initial capital x for a similar options in the CRR markets and it also provides a one sided error estimate of the approximation.

In order to compare the option prices and the shortfall risks in the BS model with the corresponding quantities in the CRR markets, we will use (a trivial form of) the Skorohod-type embedding which allows us to consider the above objects on the same probability space. Thus, define recursively $\theta_0^{(n)} = 0$, $\theta_{k+1}^{(n)} = \inf\left\{ t > \theta_k^{(n)} : |B_t^* - B_{\theta_k^{(n)}}^*| = \sqrt{\frac{T}{n}} \right\}$. Using the same arguments as in [6] we obtain that for each of the measures $P^B, \tilde{P}^B$ $\left(\theta_{k+1}^{(n)} - \theta_k^{(n)}, B_{\theta_{k+1}^{(n)}}^* - B_{\theta_k^{(n)}}^* \right)$ are independent of $\mathscr{F}_{\theta_k^{(n)}}^B$, $k = 1, 2, \ldots$. The Skorohod embedding also allows us to define mappings (introduced in [3] and [6]) which map hedges in CRR markets to hedges in the BS model and which will play a decisive role in Theorems 3 and 4 below. For the reader's convenience we review the definitions. For any $n \in \mathbb{N}$ set $\mathfrak{b}_i^{(n)} = B_{\theta_i^{(n)}}^* - B_{\theta_{i-1}^{(n)}}^*$, $i = 1, 2, \ldots$ and following [6] introduce for each $k = 1, 2, \ldots$ the finite σ–algebra $\mathscr{G}_k^{B,n} = \sigma\{\mathfrak{b}_1^{(n)}, \ldots, \mathfrak{b}_k^{(n)}\}$ with $\mathscr{G}_0^{B,n} = \{\emptyset, \Omega_B\}$. Let $\mathscr{S}_{0,n}^{B,n}$ be the set of all stopping times with respect to the filtration $\mathscr{G}_k^{B,n}, k = 0, 1, 2 \ldots$ with values in $\{0, 1 \ldots, n\}$. Observe that for any n we have a natural bijection $\Pi_n : L^\infty(\mathscr{F}_n^\xi, P_n^\xi) \to L^\infty(\mathscr{G}_n^{B,n}, P^B)$ which is given by $\Pi_n(Z) = \tilde{Z}$ so that if $Z = f(\xi_1, \ldots, \xi_n)$ for a function f on $\{-1, 1\}^n$ then $\tilde{Z} = f\left(\sqrt{\frac{n}{T}} \mathfrak{b}_1^{(n)}, \ldots, \sqrt{\frac{n}{T}} \mathfrak{b}_n^{(n)} \right)$. Notice that if we restrict Π_n to $\mathscr{T}_{0n}^\xi$ we get a bijection $\Pi_n : \mathscr{T}_{0n}^\xi \to \mathscr{S}_{0,n}^{B,n}$. In addition to the set $\mathscr{S}_{0,n}^{B,n}$ consider also the set $\mathscr{T}_{0,n}^{B,n}$ of stopping times with respect to the filtration $\left\{ \mathscr{F}_{\theta_k^{(n)}}^B \right\}_{k=0}^n$ with values in

$\{0, 1, \dots n\}$. Clearly $\mathscr{S}_{0,n}^{B,n} \subset \mathscr{T}_{0,n}^{B,n}$. Next, we define a function $\phi_n : \mathscr{T}_{0n}^{\xi} \to \mathscr{T}_{0T}^{B}$ which maps stopping times in CRR markets to stopping times in the BS model by

$$\phi_n(\sigma) = T \wedge \theta_{\Pi_n(\sigma)}^{(n)} \text{ if } \Pi_n(\sigma) < n \text{ and } \phi_n(\sigma) = T \text{ if } \Pi_n(\sigma) = n. \tag{5}$$

It is easy to see that $\phi_n(\sigma) \in \mathscr{T}_{0T}^{B}$ (see (2.28) in [3]). For each n and $x > 0$ let $\mathscr{A}^{B,n}(x)$ be the set of all *admissible* self-financing strategies with an initial capital x in the BS model which can be managed only on the set $\{0, \theta_1^{(n)}, \dots, \theta_n^{(n)}\}$, such that the discounted portfolio value remains constant after the moment $\theta_n^{(n)}$ and set $\mathscr{A}^{B,n} = \bigcup_{x \geq 0} \mathscr{A}^{B,n}(x)$. Thus if $\pi = \{(\beta_t, \gamma_t)\}_{t=0}^{\infty} \in \mathscr{A}^{B,n}$ then $\beta_t = \beta_{\theta_k^{(n)}}$ and $\gamma_t = \gamma_{\theta_k^{(n)}}$ for any $k < n$ and $t \in [\theta_k^{(n)}, \theta_{k+1}^{(n)})$. Furthermore, in order to keep the discounted portfolio constant after $\theta_n^{(n)}$ the investor should sell all his stocks at the moment $\theta_n^{(n)}$ and buy bonds with the proceeds, and so $\gamma_t = 0$ for $t \geq \theta_n^{(n)}$. From (1) it follows that for $\pi = \{(\beta_t, \gamma_t)\}_{t=0}^{\infty} \in \mathscr{A}^{B,n}$ the corresponding discounted portfolio value is given by

$$\tilde{V}_t^{\pi} = \tilde{V}_{\theta_k^{(n)}}^{\pi} + \gamma_{\theta_k^{(n)}} \left(\tilde{S}_t^B - \tilde{S}_{\theta_k^{(n)}}^B \right), \ t \in \left[\theta_k^{(n)}, \theta_{k+1}^{(n)}\right] \text{ and } \tilde{V}_t^{\pi} = \tilde{V}_{\theta_n^{(n)}}^{\pi}, \ t > \theta_n^{(n)}. \tag{6}$$

Finally, we define a function $\psi_n : \mathscr{A}^{\xi,n}(x) \to \mathscr{A}^{B,n}(x)$ which maps *admissible* self-financing strategies in the CRR n–step model to the set of the above self-financing strategies in the BS model. For $\pi = \{(\beta_k, \gamma_k)\}_{k=1}^{n} \in \mathscr{A}^{\xi,n}(x)$ define $\psi_n(\pi) \in \mathscr{A}^{B,n}(x)$ by

$$\begin{aligned} \tilde{V}_t^{\psi_n(\pi)} &= \tilde{V}_{\theta_k^{(n)}}^{\psi_n(\pi)} + \Pi_n(\gamma_{k+1})(\tilde{S}_t^B - \tilde{S}_{\theta_k^{(n)}}^B), \ t \in [\theta_k^{(n)}, \theta_{k+1}^{(n)}], \\ \text{and } \tilde{V}_t^{\psi_n(\pi)} &= \tilde{V}_{\theta_n^{(n)}}^{\psi_n(\pi)}, \quad t > \theta_n^{(n)}. \end{aligned} \tag{7}$$

Using the same arguments as in [3] (after (2.30)) it follows that $\psi_n(\pi) \in \mathscr{A}^{B,n}(x)$.

Let $I = (L, R)$ be an open interval and set $L_n = L \exp\left(-n^{-\frac{1}{3}}\right)$, $R_n = R \exp(n^{-\frac{1}{3}})$ (with $R_n = \infty$ if $R = \infty$) and $I_n = (L_n, R_n)$. Consider an investor in the BS market whose initial capital x is less than the option price $\mathscr{V}^I$. A hedge $(\pi, \sigma) \in \mathscr{A}^B(x) \times \mathscr{T}_{0T}^B$ will be called ε-optimal if $R^I(\pi, \sigma) \leq R^I(x) + \varepsilon$. For $\varepsilon = 0$ the above hedge is called an optimal hedge. For the CRR markets we have an analogous definitions. In the next section we will follow [2] and construct optimal hedges $(\pi_n, \sigma_n) \in \mathscr{A}^{\xi,n}(x) \times \mathscr{T}_{0n}^{\xi}$ for double barrier options in the n–step CRR markets with barriers L_n, R_n. By embedding these hedges into the BS model we obtain a simple representation of ε–optimal hedges for the the BS model.

Theorem 3. *For any n let $(\pi_n, \sigma_n) \in \mathscr{A}^{\xi,n}(x) \times \mathscr{T}_{0n}^{\xi}$ be the optimal hedge which is given by (10) with $H = I_n$. Then $lim_{n \to \infty} R^I(\psi_n(\pi_n), \phi_n(\sigma_n)) = R^I(x)$.*

Next, let $(\pi, \sigma) \in \mathscr{A}^{\xi,n}\left(\mathscr{V}_n^{I_n}\right) \times \mathscr{T}_{0n}^{\xi}$ be a perfect hedge for a double barrier option in the n–step CRR market with the barriers L_n, R_n, i.e. a hedge which satisfies

$\tilde{V}^{\pi}_{\sigma\wedge k} \geq Q^{I,n}(\sigma,k)$ for any $k \leq n$. In general the construction of perfect hedges for game options in CRR markets can be done explicitly (see [5], Theorem 2.1). The following result shows that if we embed the perfect hedge (π,σ) into the BS model we obtain a hedge with small shortfall risk for the barrier option with barriers L, R.

Theorem 4. *Let $I = (L, R)$ be an open interval. For any n let $(\pi^p_n, \sigma^p_n) \in \mathscr{A}^{\xi,n}(\mathscr{V}^{I_n}_n) \times \mathscr{T}^{\xi}_{0n}$ be a perfect hedge for a double barrier option in the n–step CRR market with the barriers L_n, R_n. Define $(\pi^B_n, \sigma^B_n) \in \mathscr{A}^B(\mathscr{V}^{I_n}_n) \times \mathscr{T}^B_{0T}$ by $\pi^B_n = \psi_n(\pi^p_n)$ and $\sigma^B_n = \phi_n(\sigma^p_n)$. There exists a constant C_3 such that for any n, $R^I(\pi^B_n, \sigma^B_n) \leq C_3 n^{-\frac{1}{4}}(\ln n)^{\frac{3}{4}}$.*

From (13) and Theorem 1 we will see that there exists a constant $\tilde{C}$ such that $|\mathscr{V}^I - \mathscr{V}^{I_n}_n| \leq \tilde{C} n^{-\frac{1}{4}}(\ln n)^{\frac{3}{4}}$ for any n. Since the above term is small, then in practice a seller of a double barrier game option with the barriers L, R can invest the amount $\mathscr{V}^{I_n}_n$ in the portfolio and use the above hedges facing only small shortfall risk.

Similar results with slightly worse error estimates can be obtained for double barrier knock–in options. In this case for a given open interval $I = (L, R)$ the payoff processes in the BS model and in the n–step CRR market are defined by $\mathbb{X}_t = G_t(S^{(n)})$, $\mathbb{Y}^I_t = F_t(S^B)\mathbb{I}_{t\geq\tau_I}$ and $\mathbb{X}^{(n)}_k = G_{\frac{kT}{n}}(S^B)$, $\mathbb{Y}^{I,n}_k = F_{\frac{kT}{n}}(S^B)\mathbb{I}_{k\geq\tau^{(n)}_I}$, respectively. Notice that the seller will pay for cancellation an amount which does not depend on the barriers. If we define the high payoff process $\mathbb{X}^I_t$, $t \geq 0$ in a way similar to the low payoff process $\mathbb{Y}^I_t$, $t \geq 0$, namely, $\mathbb{X}^I_t = G_t(S^B)\mathbb{I}_{t\geq\tau_I}$ then the seller could cancel the contract at the moment $t = 0$ without paying anything to the buyer, which would make such a contract worthless.

Now, for the BS model we define the option price and the shortfall risks by

$$\tilde{\mathscr{V}}^I = \inf_{\sigma\in\mathscr{T}^B_{0T}} \sup_{\tau\in\mathscr{T}^B_{0T}} \tilde{E}^B\tilde{Q}^{B,I}(\sigma,\tau),\ \tilde{R}^I(\pi,\sigma) = \sup_{\tau\in\mathscr{T}^B_{0T}} E^B(\tilde{Q}^{B,I}(\sigma,\tau) \\ -\tilde{V}^{\pi}_{\sigma\wedge\tau})^+,\ \ \tilde{R}^I(\pi) = \inf_{\sigma\in\mathscr{T}^B_{0T}} \tilde{R}^I(\pi,\sigma) \text{ and } \tilde{R}^I(x) = \inf_{\pi\in\mathscr{A}^B(x)} \tilde{R}^I(\pi)$$

where $\tilde{Q}^I(t,s) = e^{-r(t\wedge s)}(\mathbb{X}_t\mathbb{I}_{t<s} + \mathbb{Y}^I_t\mathbb{I}_{s\leq t})$ is the discounted payoff function. For the n–step CRR market the corresponding definitions are

$$\tilde{\mathscr{V}}^I_n = \min_{\zeta\in\mathscr{T}^{\xi}_{0n}} \max_{\eta\in\mathscr{T}^{\xi}_{0n}} \tilde{E}^{\xi}_n\tilde{Q}^{I,n}(\zeta,\eta),\ \tilde{R}^I_n(\pi,\sigma) = \max_{\tau\in\mathscr{T}^{\xi}_{0n}} E^{\xi}_n(\tilde{Q}^{I,n}(\sigma,\tau) \\ -\tilde{V}^{\pi}_{\sigma\wedge\tau})^+,\ \ \tilde{R}^I_n(\pi) = \min_{\sigma\in\mathscr{T}^{\xi}_{0n}} \tilde{R}^I_n(\pi,\sigma) \text{ and } \tilde{R}^I_n(x) = \inf_{\pi\in\mathscr{A}^{\xi,n}(x)} \tilde{R}^I_n(\pi)$$

where $\tilde{Q}^{I,n}(k,l) = (1+r_n)^{-k\wedge l}\left(\mathbb{X}^{(n)}_k\mathbb{I}_{k<l} + \mathbb{Y}^{I,n}_l\mathbb{I}_{l\leq k}\right)$ is the discounted payoff function. Denote also by $Q^{(n)}(k,l) = (1+r_n)^{-k\wedge l}(G_{\frac{kT}{n}}(S^n)\mathbb{I}_{k<l} + F_{\frac{lT}{n}}(S^n)\mathbb{I}_{l\leq k})$ the regular payoff and let $\mathscr{V}_n = \min_{\zeta\in\mathscr{T}^{\xi}_{0n}} \max_{\eta\in\mathscr{T}^{\xi}_{0n}} \tilde{E}^{\xi}_n Q^{(n)}(\zeta,\eta)$ be the option price for this payoff.

Theorem 5. *Let $I = (L, R)$ be an open interval.*

(i) *For each $\epsilon > 0$ there exists a constant $\tilde{C}_{1,\epsilon}$ such that for any $n \in \mathbb{N}$, $|\tilde{\mathscr{V}}^I - \tilde{\mathscr{V}}_n^I| \le \tilde{C}_{1,\epsilon} n^{-\frac{1}{4}+\epsilon}$.*

(ii) *For each initial capital x, $lim_{n\to\infty} \tilde{R}_n^I(x) = \tilde{R}^I(x)$. Furthermore, for each $\epsilon > 0$ there exists a constant $\tilde{C}_{2,\epsilon}$ such that for any x and $n \in \mathbb{N}$, $\tilde{R}^I(x) \le \tilde{R}_n^I(x) + \tilde{C}_{2,\epsilon} n^{-\frac{1}{4}+\epsilon}$.*

(iii) *For each $n \in \mathbb{N}$ let $(\pi_n^p, \sigma_n^p) \in \mathscr{A}^{\xi,n}(\tilde{\mathscr{V}}_n^I) \times \mathscr{T}_{0n}^{\xi}$ be a perfect hedge for a double barrier knock-in option as above in the n–step CRR market with the barriers L, R. Then for any $\epsilon > 0$ and $n \in \mathbb{N}$, $\tilde{R}^I(\psi_n(\pi_n^p), \phi_n(\sigma_n^p)) \le \tilde{C}_{2,\epsilon} n^{-\frac{1}{4}+\epsilon}$.*

(iv) *For any $n \in \mathbb{N}$ let $(\tilde{\pi}_n^I, \tilde{\sigma}_n^I) \in \mathscr{A}^{\xi,n}(x) \times \mathscr{T}_{0n}^{\xi}$ be an optimal hedge (hedge which minimizes the shortfall risk for an initial capital x in the n–step CRR market with the barriers L, R and can be calculated in a similar way to formula (10)). Then $lim_{n\to\infty} \tilde{R}^I(\psi_n(\tilde{\pi}_n^I), \phi_n(\tilde{\sigma}_n^I)) = \tilde{R}^I(x)$.*

All the constants above do not depend on the interval I.

The proof of this theorem is rather long but it is based on similar arguments as for knock–out options, and so we will give in this paper the detailed proof for the latter case only.

3 Auxiliary Lemmas

We first introduce the machinery which enables us to reduce the optimization of the shortfall risk to optimal stopping problems for Dynkin's games with appropriately chosen payoff processes, so that on the next stage we will be able to employ the Skorohod embedding in order to compare values of the corresponding discrete and continuous time games. This machinery was used in [3] for similar purposes in the case of regular game options. For any n set $a_1^{(n)} = e^{\kappa\sqrt{\frac{T}{n}}} - 1$, $a_2^{(n)} = e^{-\kappa\sqrt{\frac{T}{n}}} - 1$ and observe that for any $m \le n$ the random variable $\frac{\tilde{S}^{(n)}_{\frac{mT}{n}}}{\tilde{S}^{(n)}_{(m-1)T/n}} - 1 = \exp(\kappa(\frac{T}{n})^{1/2}\xi_m) - 1$ takes on only the values $a_1^{(n)}, a_2^{(n)}$. For each $y > 0$ and $n \in \mathbb{N}$ introduce the closed interval $K_n(y) = \left[-\frac{y}{a_1^{(n)}}, -\frac{y}{a_2^{(n)}}\right]$ and for $0 \le k < n$ and a given positive $\mathscr{F}_k^{\xi}$-measurable random variable X define

$$\mathscr{A}_k^{\xi,n}(X) = \left\{ Y | Y = X + \alpha \left(\exp\left(\kappa \left(\frac{T}{n}\right)^{1/2} \xi_{k+1} \right) - 1 \right) \text{ for some } \right.$$
$$\left. \mathscr{F}_k^{\xi} - \text{measurable } \alpha \in K_n(X) \right\}. \tag{8}$$

From (2) it follows that $\mathscr{A}_k^{\xi,n}(X)$ is the set of all possible discounted portfolio values at time $k+1$ provided that the discounted portfolio value at time k is X.

Let H be an open interval. For any $\pi \in \mathscr{A}^{\xi,n}$ define a sequence of random variables $\{W_k^{H,\pi}\}_{k=0}^n$ and a stopping time $\sigma(H,\pi) \in \mathscr{T}_{0n}^{\xi}$

$$W_n^{H,\pi} = (\tilde{Y}_n^{H,n} - \tilde{V}_n^{\pi})^+, \; W_k^{H,\pi} = \min\left(\left(\tilde{X}_k^{H,n} - \tilde{V}_k^{\pi}\right)^+, \right.$$
$$\left. \max\left(\left(\tilde{Y}_k^{H,n} - \tilde{V}_k^{\pi}\right)^+, E_n^{\xi}(W_{k+1}^{H,\pi} | \mathscr{F}_k^{\xi})\right)\right) \text{ for } k < n$$
$$\text{and } \sigma(H,\pi) = \min\left\{k | (\tilde{X}_k^{H,n} - \tilde{V}_k^{\pi})^+ = W_k^{H,\pi}\right\} \wedge n.$$

On the Brownian probability space set $S_t^{B,n} = S_0$, for $t \in [0, T/n]$ and $S_t^{B,n} = S_0 \exp(\sum_{k=1}^{[nt/T]} (\frac{rT}{n} + \kappa \mathfrak{b}_k^{(n)}))$, for $t \in [T/n, T]$. Define $\tau_H^{B,n} = \min\{k \geq 0 | S_{\frac{kT}{n}}^{B,n} \notin H\}$. Clearly $\tau_H^{B,n}$ is a stopping time with respect to the filtration $\mathscr{G}_k^{B,n}$, $k \geq 0$. Consider the new payoffs $Y_k^{B,H,n} = F_{\frac{kT}{n}}(S^{B,n}) \mathbb{I}_{k < \tau_H^{B,n}}$ and $X_k^{B,H,n} = G_{\frac{kT}{n}}(S^{B,n}) \mathbb{I}_{k < \tau_H^{B,n}}$, $k \leq n$. The corresponding payoff function is given by $Q^{B,H,n}(k,l) = \tilde{X}_k^{B,H,n} \mathbb{I}_{k<l} + \tilde{Y}_l^{B,H,n} \mathbb{I}_{l \leq k}$, $k, l \leq n$ where $\tilde{Y}_k^{B,H,n} = (1+r_n)^{-k} Y_k^{B,H,n}$ and $\tilde{X}_k^{B,H,n} = (1+r_n)^{-k} X_k^{B,H,n}$ are the discounted payoffs. For any n we now consider hedges which are elements in $\mathscr{A}^{B,n} \times \mathscr{T}_{0,n}^{B,n}$. Given a positive $\mathscr{F}_{\theta_k^{(n)}}^B$–measurable random variable X define $\mathscr{A}_k^{B,n}(X)$ by (8) with $\sqrt{\frac{T}{n}} \xi_{k+1}$ and $\mathscr{F}_k^{\xi}$ replaced by $\mathfrak{b}_{k+1}^{(n)}$ and $\mathscr{F}_{\theta_k^{(n)}}^B$, respectively. By (6) we have that $\mathscr{A}_k^{B,n}(X)$ consists of all possible discounted values at the time $\theta_{k+1}^{(n)}$ of portfolios managed only at embedding times $\{\theta_i^{(n)}\}$ with the discounted stock evolution $\tilde{S}_t^B$, provided the discounted portfolio value at time $\theta_k^{(n)}$ is X.

Next, define the shortfall risk by

$$R_n^{B,H}(\pi,\zeta) = \sup_{\eta \in \mathscr{T}_{0n}^{B,n}} E^B(Q^{B,H,n}(\zeta,\eta) - \tilde{V}_{\theta_{\zeta\wedge\eta}^{(n)}}^{\pi})^+,$$
$$R_n^{B,H}(\pi) = \inf_{\zeta \in \mathscr{T}_{0n}^{B,n}} R_n^{B,H}(\pi,\zeta) \text{ and } R_n^{B,H}(x) = \inf_{\pi \in \mathscr{A}^{B,n}(x)} R_n^{B,H}(\pi). \quad (9)$$

For any $\pi \in \mathscr{A}^{B,n}$ define a sequence of random variables $\{U_k^{H,\pi}\}_{k=0}^n$ and a stopping time $\zeta(H,\pi) \in \mathscr{T}_{0,n}^{B,n}$

$$U_n^{H,\pi} = (\tilde{Y}_n^{B,H,n} - \tilde{V}_{\theta_n^{(n)}}^{\pi})^+, \; U_k^{H,\pi} = \min\left(\left(\tilde{X}_k^{B,H,n} - \tilde{V}_{\theta_k^{(n)}}^{\pi}\right)^+, \right.$$
$$\left. \max\left(\left(\tilde{Y}_k^{B,H,n} - \tilde{V}_{\theta_k^{(n)}}^{\pi}\right)^+, E^B(U_{k+1}^{H,\pi} | \mathscr{F}_{\theta_k^{(n)}}^B)\right)\right), \; k < n$$
$$\text{and } \zeta(H,\pi) = \min\{k | (\tilde{X}_k^{B,H,n} - \tilde{V}_{\theta_k^{(n)}}^{\pi})^+ = U_k^{H,\pi}\} \wedge n.$$

For $k \le n$ and $x_1, \dots, x_k \in \mathbb{R}$, consider the function $\psi^{x_1,\dots,x_k} \in M[0, \frac{kT}{n}]$ given by

$$\psi^{x_1,\dots,x_k}(t) = S_0 \exp(\tfrac{rjT}{n} + \kappa \textstyle\sum_{i=1}^{j} x_i),\ t \in [jt/n, (j+1)T/n),\ 1 \le j \le k$$
$$\text{and } \psi^{x_1,\dots,x_k}(0) = S_0,\ \ t \in [0, T/n),$$

there exist $f_k^n, g_k^n : \mathbb{R}^k \to \mathbb{R}$ such that for any $x_1, \dots, x_k \in \mathbb{R}$,

$$f_k^n(x_1, \dots, x_k) = (1 + r_n)^{-k} F_{\frac{kT}{n}}(\psi^{x_1,\dots,x_k}) = e^{-rkT/n} F_{\frac{kT}{n}}(\psi^{x_1,\dots,x_k}),$$
$$\text{and } g_k^n(x_1, \dots, x_k) = (1 + r_n)^{-k} G_{\frac{kT}{n}}(\psi^{x_1,\dots,x_k}) = e^{-rkT/n} G_{\frac{kT}{n}}(\psi^{x_1,\dots,x_k}).$$

Set $q_k^{H,n}(x_1, \dots, x_k) = \mathbb{I}_{[\min_{0\le i\le k} \psi^{x_1,\dots,x_i}(\frac{iT}{n}), \max_{0\le i\le k} \psi^{x_1,\dots,x_i}(\frac{iT}{n})] \subset H}$. Finally, define a sequence $\{J_k^{H,n}\}_{k=0}^n$ of functions $J_k^{H,n} : [0,\infty)\times\mathbb{R}^k \to \mathbb{R}$ by the following backward recursion

$$J_n^{H,n}(y, u_1, u_2 \dots, u_n) = (f_n^n(u_1, \dots, u_n) q_n^{H,n}(u_1, \dots, u_n) - y)^+ \text{ and}$$
$$J_k^{H,n}(y, u_1, \dots, u_k) = \min\left(\left(g_k^n(u_1, \dots, u_k) q_k^{H,n}(u_1, \dots, u_k) - y\right)^+, \right.$$
$$\max\left(\left(f_k^n(u_1, \dots, u_k) \times q_k^{H,n}(u_1, \dots, u_k) - y\right)^+, \right.$$
$$\inf_{u\in K_n(y)} \left(p^{(n)} J_{k+1}^{H,n}\left(y + ua_1^{(n)}, u_1, \dots, u_k, \sqrt{\frac{T}{n}}\right) \right.$$
$$\left.\left.\left. + \left(1 - p^{(n)}\right) J_{k+1}^{H,n}\left(y + ua_2^{(n)}, u_1, \dots, u_k, -\sqrt{\frac{T}{n}}\right)\right)\right)\right)$$
$$\text{for } k = n-1, n-2, \dots, 0.$$

Similarly to [3] this dynamic programming relations will enable us to compute shortfall risks defined in (3) and (9).

Lemma 1. *The function $J_k^{H,n}(y, u_1, \dots, u_k)$ is continuous and decreasing with respect to y for any n, $k \le n$ and an open interval H.*

Proof. The proof is the same as the proof of Lemma 3.2 in [3], just replace $J_k^{H,n}$ by J_k^n. □

For a given closed interval $K = [a, b]$ and a function $f : K \times \mathbb{R}^k \to \mathbb{R}$ such that $f(\cdot, v)$ is continuous for all $v \in \mathbb{R}^k$ define $argmin_{a\le u\le b} f(u, v) = \min\{w \in K | f(w, v) = \min_{\beta\in K} f(\beta, v)\}$. Lemma 1 enables us to define the following functions

$$h_k^{H,n}(y, x_1, \dots, x_k) = argmin_{u\in K_n(y)} \left(p^{(n)} J_{k+1}^{H,n}\left(y + ua_1^{(n)}, \right.\right.$$
$$\left.\left. u_1, \dots, u_k, \sqrt{\tfrac{T}{n}}\right) + (1 - p^{(n)}) J_{k+1}^{H,n}\left(y + ua_2^{(n)}, u_1, \dots, u_k, -\sqrt{\tfrac{T}{n}}\right)\right),\ \ k < n.$$

Let x be an initial capital. For any n and an open interval H there exists a hedge$(\pi_n^H, \sigma_n^H) \in \mathscr{A}^{\xi,n}(x) \times \mathscr{T}_{0n}^{\xi}$ such that

$$\tilde{V}_0^{\pi_n^H} = x \text{ and } \tilde{V}_{k+1}^{\pi_n^H} = \tilde{V}_k^{\pi_n^H} + h_k^{H,n}\left(\tilde{V}_k^{\pi_n^H}, e^{\kappa\sqrt{\frac{T}{n}}\xi_1}, \ldots, e^{\kappa\sqrt{\frac{T}{n}}\xi_k}\right)$$
$$\times (e^{\kappa\sqrt{\frac{T}{n}}\xi_{k+1}} - 1) \text{ for } k > 0 \text{ and } \sigma_n^H = \sigma(H, \pi_n^H). \tag{10}$$

From the arguments concerning $\mathscr{A}_k^{\xi,n}(X)$ at the beginning of this section it follows that π_n^H is an *admissible* strategy. Let $(\pi_n^{B,H}, \zeta_n^H) \in \mathscr{A}^{B,n}(x) \times \mathscr{T}_{0,n}^{B,n}$ be a hedge which is given by $\pi_n^{B,H} = \psi_n(\pi_n^H)$ and $\zeta_n^H = \Pi_n(\sigma_n^H)$ where, recall, the maps ψ_n, Π_n were defined in Sect. 2. Namely, we consider a hedge which is determined by

$$\tilde{V}_0^{\pi_n^{B,H}} = x \text{ and } \tilde{V}_{k+1}^{\pi_n^{B,H}} = \tilde{V}_k^{\pi_n^{B,H}} + h_k^{H,n}(\tilde{V}_k^{\pi_n^{B,H}}, e^{\mathfrak{b}_1^{(n)}}, \ldots, e^{\mathfrak{b}_k^{(n)}})$$
$$\times (e^{\mathfrak{b}_{k+1}^{(n)}} - 1) \text{ for } k > 0 \text{ and } \zeta_n^H = \zeta(H, \pi_n^{B,H}). \tag{11}$$

The following lemma enables us to consider all relevant processes on the Brownian probability space and to deal with stopping times with respect to the same filtration.

Lemma 2. *For any initial capital x, $n \in \mathbb{N}$ and an open interval H,*

$$R_n^H(x) = R_n^H(\pi_n^H, \sigma_n^H) = J_0^{H,n}(x) = R_n^{B,H}(\pi_n^{B,H}, \zeta_n^H) = R_n^{B,H}(x).$$

Proof. The proof is the same as in Lemma 3.3 of [3], just replace J_k^n, R_n, $R^{B,n}$, (π_n, σ_n) and $(\tilde{\pi}_n, \zeta_n)$ by $J_k^{H,n}$, R_n^H, $R_n^{B,H}$, (π_n^H, σ_n^H) and $(\pi_n^{B,H}, \zeta_n^H)$, respectively. □

We next deal with estimates for the BS model. Let $H = (L, R)$ be an open interval. For any $\epsilon > 0$ set $H_\epsilon = (Le^{-\epsilon}, Re^{\epsilon})$. Since $H_\varepsilon \supset H$ and both the payoff and the shortfall risk can only grow on a larger interval it is clear that $\mathscr{V}^{H_\epsilon} \geq \mathscr{V}^H$ for any $\epsilon > 0$ and $R^{H_\epsilon}(x) \geq R^H(x)$ for any initial capital x. The following result provides an estimate from above of the term $R^{H_\epsilon}(x) - R^H(x)$.

Lemma 3. *For any $q > 1$ there exists a constant A_q such that for any initial capital x, $\epsilon > 0$ and an open interval H,*

$$R^{H_\epsilon}(x) - R^H(x) \leq A_q \epsilon^{1/q}. \tag{12}$$

Proof. Before proving the lemma observe that if $P = \tilde{P}$ then the option price can be represented as the shortfall risk for an initial capital $x = 0$, i.e. if $\mu = 0$ then $\mathscr{V}^I = R^I(0)$ for any open interval I. Hence, by (12) for any $q > 1$ there exists a constant $\tilde{A}_q$ (which is equal to A_q for the case $\mu = 0$) such that for any open interval H and $\epsilon > 0$,

$$\mathscr{V}^{H_\epsilon} - \mathscr{V}^H \leq \tilde{A}_q \epsilon^{1/q}. \tag{13}$$

Next we turn to the proof of the lemma. Choose an initial capital x, an open interval $H = (L, R)$, some $\epsilon > 0$ and fix $\delta > 0$. There exists a $\pi_1 \in \mathscr{A}^B(x)$ such that $R^H(\pi_1) < R^H(x) + \delta$. It is well known (see [4]) that in a BS model the discounted portfolio process $\{\tilde{V}_t^{\pi_1}\}_{t=0}^T$ has a continuous modification. Observe that $(Q^{B,H}(\sigma,\tau) - \tilde{V}^{\pi_1}_{\sigma\wedge\tau})^+ = (Q^{B,H}(\tau_H \wedge \sigma, \tau) - \tilde{V}^{\pi_1}_{\tau_H\wedge\sigma\wedge\tau})^+$ for all stopping times $\sigma, \tau \in \mathscr{T}_{0T}^B$. Thus, there exists a hedge $(\pi_1, \sigma_1) \in \mathscr{A}^B(x) \times \mathscr{T}_{0T}^B$ such that

$$R^H(\pi_1, \sigma_1) < R^H(x) + \delta \text{ and } \sigma_1 \le \tau_H. \tag{14}$$

Define the stopping time $\sigma_2 \in \mathscr{T}_{0T}^B$ by $\sigma_2 = \sigma_1 \mathbb{I}_{\sigma_1 < \tau_H} + T\mathbb{I}_{\sigma_1 \ge \tau_H}$. Observe that if $\pi_1 = \{(\mathbf{e}_t, \gamma_t)\}_{t=0}^T$ and $\pi_2 = \{(\tilde{\mathbf{e}}_t, \tilde{\gamma}_t)\}_{t=0}^T$ with $\tilde{\gamma}_t = \gamma_t \mathbb{I}_{\sigma_1 \le t}$ and $\tilde{\mathbf{e}}_t = (x + \int_0^t \tilde{\gamma}_u d\tilde{S}_u^B - \tilde{\gamma}_t \tilde{S}_t^B)/b_0$ then π_2 is an admissible self-financing strategy and $\tilde{V}_t^{\pi_2} = \tilde{V}^{\pi_1}_{t\wedge\sigma_1}$. Consider the hedge $(\pi_2, \sigma_2) \in \mathscr{A}^B(x) \times \mathscr{T}_{0T}^B$ then $(Q^{B,H_\epsilon}(\sigma_2, \tilde{\tau}) - \tilde{V}^{\pi_2}_{\sigma_2\wedge\tilde{\tau}})^+ = (Q^{B,H_\epsilon}(\sigma_2, \tilde{\tau}\wedge\tau_{H_\epsilon}) - \tilde{V}^{\pi_2}_{\tau_{H_\epsilon}\wedge\sigma_2\wedge\tilde{\tau}})^+$ for any $\tilde{\tau} \in \mathscr{T}_{0T}^B$. Thus, there exists a stopping time $\tau \in \mathscr{T}_{0T}^B$ such that

$$R^{H_\epsilon}(\pi_2, \sigma_2) < E^B[Q^{B,H_\epsilon}(\sigma_2, \tau) - \tilde{V}^{\pi_2}_{\sigma_2\wedge\tau})^+] + \delta \text{ and } \tau \le \tau_{H_\epsilon}. \tag{15}$$

For any $\alpha > 0$ denote $J_\alpha = (Le^\alpha, Re^{-\alpha})$. Set $U_\alpha = (Q^{B,H}(\sigma_1, \tau\wedge\tau_{J_\alpha}) - \tilde{V}^{\pi_1}_{\sigma_1\wedge\tau\wedge\tau_{J_\alpha}})^+$. Clearly, $\tau \wedge \tau_{J_\alpha} \le \tau \wedge \tau_H$ for any $\alpha > 0$ and $\tau \wedge \tau_{J_\alpha} \uparrow \tau \wedge \tau_H$ as $\alpha \to 0$. Since the process $\{\tilde{V}_t^{\pi_1}\}_{t=0}^T$ is continuous and $\sigma_1 \le \tau_H$ we obtain by the choice of π_2 that

$$\lim_{\alpha\to 0} U_\alpha = (e^{-r\sigma_1} G_{\sigma_1}(S^B)\mathbb{I}_{\sigma_1 < \tau\wedge\tau_H} + e^{-r(\tau\wedge\tau_H)} F_{\tau\wedge\tau_H}(S^B)\mathbb{I}_{\sigma_1 \ge \tau\wedge\tau_H} - \tilde{V}_\tau^{\pi_2})^+. \tag{16}$$

Observe that $R^H(\pi_1, \sigma_1) \ge E^B U_\alpha$ for any α. Thus from (16) and Fatou's lemma we obtain

$$\begin{aligned} R^H(\pi_1, \sigma_1) \ge E^B \lim_{\alpha\to 0} U_\alpha = E^B(&e^{-r\sigma_1} G_{\sigma_1}(S^B)\mathbb{I}_{\sigma_1 < \tau\wedge\tau_H} \\ &+ e^{-r(\tau\wedge\tau_H)} F_{\tau\wedge\tau_H}(S^B)\mathbb{I}_{\sigma_1 \ge \tau\wedge\tau_H} - \tilde{V}_\tau^{\pi_2})^+. \end{aligned} \tag{17}$$

Since $\sigma_2 \ge \sigma_1$ a.s. then from the definition of π_2 it follows that $\tilde{V}^{\pi_2}_{\sigma_2\wedge t} = \tilde{V}^{\pi_1}_{\sigma_1\wedge\sigma_2\wedge t} = \tilde{V}^{\pi_1}_{\sigma_1\wedge t} = \tilde{V}_t^{\pi_2}$ for all t. This together with (15) gives

$$R^{H_\epsilon}(\pi_2, \sigma_2) < E^B(e^{-r\sigma_2} G_{\sigma_2}(S^B)\mathbb{I}_{\sigma_2 < \tau} + e^{-r\tau} F_\tau(S^B)\mathbb{I}_{\sigma_2 \ge \tau} - \tilde{V}_\tau^{\pi_2})^+ + \delta. \tag{18}$$

Observe that if $\sigma_2 < \tau$ then $\sigma_2 = \sigma_1 < \tau \wedge \tau_H$ and if $\sigma_2 \ge \tau$ then $\sigma_1 \ge \tau \wedge \tau_H$. And so from (14), (17) and (18) we obtain that

$$\begin{aligned} &R^{H_\epsilon}(x) - R^H(x) \le R^{H_\epsilon}(\pi_2, \sigma_2) - R^H(\pi_1, \sigma_1) + \delta \le 2\delta \\ &\quad + E^B|e^{-r\tau} F_\tau(S^B) - e^{-r(\tau\wedge\tau_H)} F_{\tau\wedge\tau_H}(S^B)| \le 2\delta + E^B\Gamma_1 + E^B\Gamma_2 \end{aligned} \tag{19}$$

where $\Gamma_1 = |e^{-r\tau} - e^{-r(\tau\wedge\tau_H)}|F_\tau(S^B)$ and $\Gamma_2 = |F_\tau(S^B) - F_{\tau\wedge\tau_H}(S^B)|$. In order to estimate $E^B\Gamma_1$ and $E^B\Gamma_2$ introduce the process $W_t = \frac{\ln S_t^B - \ln S_0}{\kappa} = B_t + (\frac{r+\mu}{\kappa} - \frac{\kappa}{2})t$, $t \geq 0$. From Girsanov's theorem (see [4]) it follows that $\{W_t\}_{t=0}^T$ is a Brownian motion with respect to the measure P_W whose restriction to the σ–algebra $\mathscr{F}_t^B$ satisfies $D_t = \frac{dP^B}{dP_W}|\mathscr{F}_t^B = \exp\left((\frac{r+\mu}{\kappa} - \frac{\kappa}{2})B_t + \frac{(\frac{r+\mu}{\kappa} - \frac{\kappa}{2})^2}{2}t\right)$. Let $q > 1$. Denote the expectation with respect to P_W by E_W. From the assumptions on the functions F, G and the Hölder inequality,

$$\begin{aligned} E^B\Gamma_1 &\leq E_W\left(r(\tau - \tau\wedge\tau_H)(F_0(S_0) + \mathscr{L}(T+2)(1 + \sup_{0\leq t\leq T} S_t^B))D_T\right) \\ &\leq c_1(E_W(\tau - \tau\wedge\tau_H)^q)^{1/q} \end{aligned} \tag{20}$$

for some constant c_1 (which depends on q). Observe that $\Gamma_2 \leq \Gamma_3 + \Gamma_4$ where $\Gamma_3 = \mathscr{L}(\tau - \tau\wedge\tau_H)(1 + \sup_{0\leq t\leq T} S_t^B)$ and $\Gamma_4 = \sup_{\tau\wedge\tau_H\leq t\leq\tau} \mathscr{L}|S_t^B - S_{\tau\wedge\tau_H}^B|$. By the Hölder inequality,

$$E^B\Gamma_3 = E_W(\mathscr{L}(\tau - \tau\wedge\tau_H)(1 + \sup_{0\leq t\leq T} S_t^B)D_T) \leq c_2(E_W(\tau - \tau\wedge\tau_H)^q)^{1/q}$$

for some constant c_2. Set $\Gamma_5 = \sup_{\tau\wedge\tau_H\leq t\leq\tau} \kappa|W_t - W_{\tau\wedge\tau_H}|$. Employing the inequality $|e^x - 1| \leq x$ for $0 \leq x \leq 1$ it follows that $\Gamma_4 \leq \mathscr{L}\sup_{0\leq t\leq T} S_t^B(\mathbb{I}_{\Gamma_5>1} + \Gamma_5)$ and together with the Markov and Hölder inequalities we obtain that there exists a constant c_3 such that

$$\begin{aligned} E^B\Gamma_4 &\leq E_W(D_T\mathscr{L}\sup_{0\leq t\leq T} S_t^B\mathbb{I}_{\Gamma_5>1}) + E_W(D_T\mathscr{L}\sup_{0\leq t\leq T} S_t^B\Gamma_5) \\ &\leq c_3(P_W\{\Gamma_5 > 1\})^{1/q} + c_3(E_W\Gamma_5^q)^{1/q} \leq 2c_3(E_W\Gamma_5^q)^{1/q}. \end{aligned} \tag{21}$$

Using the Burkholder–Davis–Gandy inequality (see [4]) for the martingale $W_t - W_{\tau\wedge\tau_H}$, $t \geq \tau\wedge\tau_H$ we obtain that there exists a constant c_4 such that $E_W\Gamma_5^q \leq c_4E_W(\tau - \tau\wedge\tau_H)^{q/2}$. This together with (20)-(21) gives

$$E^B(\Gamma_1 + \Gamma_2) \leq c_5(E_W(\tau - \tau\wedge\tau_H)^{q/2})^{1/q} \tag{22}$$

for some constant c_5. Finally, we estimate the term $E_W(\tau - \tau\wedge\tau_H)^{q/2}$. First assume that $L > 0$ and $R < \infty$. Set $x_1 = (\ln L - \ln S_0)/\kappa$, $x_2 = (\ln R - \ln S_0)/\kappa$, $y_1 = x_1 - \frac{\epsilon}{\kappa}$ and $y_2 = x_2 + \frac{\epsilon}{\kappa}$. For any $x \in \mathbb{R}$ let $\tau^{(x)} = \inf\{t \geq 0 | W_t = x\}$ be the first time the process $\{W_t\}_{t=0}^\infty$ hits the level x. Clearly $\tau^{(x)}$ is a finite stopping time with respect to P_W. By (15) we obtain that

$$\begin{aligned} \tau - \tau\wedge\tau_H &\leq T\wedge(\tau_{H_\epsilon} - \tau_H) = T\wedge(\tau^{(y_1)}\wedge\tau^{(y_2)} - \tau^{(x_1)}\wedge\tau^{(x_2)}) \\ &\leq T\wedge(\tau^{(y_1)} - \tau^{(x_1)}) + T\wedge(\tau^{(y_2)} - \tau^{(x_2)}). \end{aligned} \tag{23}$$

From the strong Markov property of the Brownian motion it follows that under P_W the random variable $\tau^{(y_1)} - \tau^{(x_1)}$ has the same distribution as $\tau^{(y_1 - x_1)} = \tau^{(-\frac{\epsilon}{\kappa})}$ and the random variable $\tau^{(y_2)} - \tau^{(x_2)}$ has the same distribution as $\tau^{(y_2 - x_2)} = \tau^{(\frac{\epsilon}{\kappa})}$.

Recall, (see [4]) that for any $z \in \mathbb{R}$ the probability density function of $\tau^{(z)}$ (with respect to P_W) is $f_{\tau^{(z)}}(t) = \frac{|z|}{\sqrt{2\pi t^3}} \exp(-\frac{z^2}{2t})$. Hence, using the inequality $(a+b)^{q/2} \leq 2^{q/2}(a^{q/2} + b^{q/2})$ together with (23) we obtain that

$$E_W(\tau - \tau \wedge \tau_H)^{q/2} \leq 2^{q/2} \left[E_W \left(T \wedge \tau^{\left(-\frac{\epsilon}{\kappa}\right)} \right)^{q/2} + E_W(T \wedge \tau^{(\frac{\epsilon}{\kappa})})^{q/2} \right]$$
$$\leq \frac{2^{q/2}\epsilon}{\sqrt{2\pi}\kappa} \left(\int_0^T \frac{1}{t^{3/2-q/2}} dt + T^{q/2} \int_T^\infty \frac{1}{t^{3/2}} dt \right) = 2^{q/2+1} \frac{q}{q-1} T^{(q-1)/2} \frac{\epsilon}{\sqrt{2\pi}\kappa}. \tag{24}$$

Observe that when either $L = 0$ or $R = \infty$ (but not both) we obtain either $\tau - \tau \wedge \tau_H \leq T \wedge \left(\tau^{(y_2)} - \tau^{(x_2)}\right)$ or $\tau - \tau \wedge \tau_H \leq T \wedge (\tau^{(y_1)} - \tau^{(x_1)})$, respectively. Thus for these cases (24) holds true, as well. From (19), (22) and (24) we see that there exists a constant A_q such that $R^{H_\epsilon}(x) - R^H(x) \leq 2\delta + A_q\epsilon^{1/q}$ and since $\delta > 0$ is arbitrary we complete the proof. $\square$

The following result provides an estimate from above of the shortfall risk when one of the barriers is close to the initial stock price S_0.

Lemma 4. *Let $I = (L, R)$ be an open interval which satisfies* $\min(\frac{R}{S_0}, \frac{S_0}{L}) \leq e^\epsilon$. *For any $q > 1$, $\epsilon > 0$ and an initial capital x, $R^I(x) \leq (F_0(S_0) - x)^+ + A_q\epsilon^{1/q}$.*

Proof. We use the same notations as in Lemma 3. Similarly to (22)–(24) (by letting τ_H=0) we obtain that

$$\sup_{\tau \in \mathscr{T}_{0T}^B} E^B |e^{-r(\tau \wedge (\tau^{(\frac{\epsilon}{\kappa})} \vee \tau^{\left(-\frac{\epsilon}{\kappa}\right)}))} F_{\tau \wedge \left(\tau^{\left(\frac{\epsilon}{\kappa}\right)} \vee \tau^{\left(-\frac{\epsilon}{\kappa}\right)}\right)}(S^B) - F_0(S_0)|$$
$$\leq c_5 \left(E_W(T \wedge (\tau^{\left(\frac{\epsilon}{\kappa}\right)} \vee \tau^{\left(-\frac{\epsilon}{\kappa}\right)}))^{q/2} \right)^{1/q} \leq A_q \epsilon^{1/q}. \tag{25}$$

Let x be an initial capital. Consider the constant portfolio $\pi \in \mathscr{A}^B(x)$ which satisfies $\tilde{V}_t^\pi = x$ for all t. Set $\sigma = (\tau^{(\frac{\epsilon}{\kappa})} \vee \tau^{(-\frac{\epsilon}{\kappa})}) \wedge T$. Since $\tau^{(\frac{\epsilon}{\kappa})} \vee \tau^{(-\frac{\epsilon}{\kappa})} \geq \tau_I$ we obtain that $R^I(x) \leq R^I(\pi, \sigma) \leq \sup_{\tau \in \mathscr{T}_{0T}^B} E^B(e^{-r(\tau \wedge (\tau^{(\frac{\epsilon}{\kappa})} \vee \tau^{(-\frac{\epsilon}{\kappa})}))} F_{\tau \wedge (\tau^{(\frac{\epsilon}{\kappa})} \vee \tau^{(-\frac{\epsilon}{\kappa})})}$ $(S^B) - x)^+$ and combining with (25) we complete the proof.

4 Proving the Main Results

In this section we complete the proof of Theorems 1–4. We start with the proof of Theorem 2. Let $x > 0$ be an initial capital and let $I = (L, R)$ be an open interval as before. Fix $\epsilon > 0$ and denote $I_\epsilon = (Le^{-\epsilon}, Re^\epsilon)$. Choose $\delta > 0$. For any z let $\mathscr{A}^{B,C}(z) \subset \mathscr{A}^B(z)$ be the subset consisting of all $\pi \in \mathscr{A}^B(z)$ such that the discounted portfolio process $\{\tilde{V}_t^\pi\}_{t=0}^T$ is a right continuous martingale with respect to the martingale measure $\tilde{P}^B$ and $\tilde{V}_T^\pi = f(B^*_{t_1}, \dots, B^*_{t_k})$ for some smooth function $f \in C_0^\infty(\mathbb{R}^k)$ with a compact support and $t_1, \dots, t_k \in [0, T]$. Using the same arguments as in Lemmas 4.1–4.3 in [3] we obtain that there exists $z < x$ and $\pi \in \mathscr{A}^{B,C}(z)$ such that $R^{I_\epsilon}(\pi) < R^{I_\epsilon}(x) + \delta$. Thus there exist k, $0 < t_1 < t_2 \dots < t_k \le T$ and $0 \le f_\delta \in C_0^\infty(\mathbb{R}^k)$ such that the portfolio $\pi \in \mathscr{A}^B$ with $\tilde{V}_t^\pi = \tilde{E}(f_\delta(B^*_{t_1}, \dots, B^*_{t_k})|\mathscr{F}_t^B)$ satisfies

$$R^{I_\epsilon}(\pi) < R^{I_\epsilon}(x) + \delta \ \text{ and } \ V_0^\pi < x. \tag{26}$$

Set $\Psi_n = f_\delta(B^*_{\theta^{(n)}_{[nt_1/T]}}, \dots, B^*_{\theta^{(n)}_{[nt_k/T]}})$, $u_n = \max_{0 \le k \le n} |\theta_k^{(n)} - \frac{kT}{n}|$ and $w_n = \max_{1 \le k \le n} |\theta_k^{(n)} - \theta_{k-1}^{(n)}| + |T - \theta_n^{(n)}|$. Since $w_n \le 3u_n + \frac{T}{n}$ then from (4.7) in [6] we obtain that for any $m \in \mathbb{R}_+$ there exists a constant $K^{(m)}$ such that for all n,

$$E^B u_n^{2m} \le K^{(m)} n^{-m} \ \text{ and } \ E^B w_n^{2m} \le K^{(m)} n^{-m}. \tag{27}$$

From the exponential moment estimates (4.8) and (4.25) of [6] it follows that there exists a constant K_1 such that for any natural n and a real a,

$$E^B e^{|a|\theta_n^{(n)} \vee T} \le e^{|a|K_1 T} \ \text{ and } \ E^B \sup_{0 \le t \le \theta_n^{(n)} \vee T} \exp(aB_t) \le 2e^{a^2 K_1 T}. \tag{28}$$

Similarly to Sect. 4 in [3] we obtain that for sufficiently large n, $v_n := \tilde{E}^B(\Psi_n) = \tilde{E}_n^\xi f_\delta\left(\sqrt{\frac{T}{n}} \sum_{i=1}^{[nt_1/T]} \xi_i, \dots, \sqrt{\frac{T}{n}} \sum_{i=1}^{[nt_k/T]} \xi_i\right) < x$. Since CRR markets are complete we can find a portfolio $\tilde{\pi}(n) \in \mathscr{A}^{\xi,n}(v_n)$ such that $\tilde{V}_n^{\tilde{\pi}} = f_\delta\left(\sqrt{\frac{T}{n}} \sum_{i=1}^{[nt_1/T]} \xi_i, \dots, \sqrt{\frac{T}{n}} \sum_{i=1}^{[nt_k/T]} \xi_i\right)$. For a fixed n let $\pi' = \psi_n(\tilde{\pi}) \in \mathscr{A}^{B,n}(v_n)$. From (7) it follows that $\tilde{V}^{\pi'}_{\theta_n^{(n)}} = \Psi_n$. Since $R_n^I(\cdot)$ is a non increasing function then by (26) and Lemma 3.2,

$$R_n^I(x) - R^{I_\epsilon}(x) \le R_n^I(v_n) - R^{I_\epsilon}(x) \le \delta + R_n^{B,I}(\pi') - R^{I_\epsilon}(\pi). \tag{29}$$

There exists a stopping time $\sigma \in \mathscr{T}_{0T}^B$ such that

$$R^{I_\epsilon}(\pi) > \sup_{\tau \in \mathscr{T}_{0T}^B} E^B(Q^{B,I_\epsilon}(\sigma, \tau) - \tilde{V}^\pi_{\sigma \wedge \tau})^+ - \delta. \tag{30}$$

Define $\zeta \in \mathscr{T}_{0,n}^{B,n}$ by

$$\zeta = (n \wedge \min\{i | \theta_i^{(n)} \ge \sigma\})\mathbb{I}_{\sigma < T} + n\mathbb{I}_{\sigma = T}. \tag{31}$$

There exists a stopping time $\eta \in \mathscr{T}_{0,n}^{B,n}$ such that

$$E^B\left(Q^{B,I,n}\left(\frac{\zeta T}{n}, \frac{\eta T}{n}\right) - \tilde{V}^{\pi'}_{\theta^{(n)}_{\zeta \wedge \eta}}\right)^+ > R_n^{B,I}(\pi') - \delta. \tag{32}$$

From (30) and (32) we obtain that

$$R_n^{B,I}(\pi') - R^{I_\epsilon}(\pi) < 2\delta + E^B(\Lambda_1 + \Lambda_2 + \Lambda_3) \tag{33}$$

where $\Lambda_1 = |\tilde{V}^{\pi'}_{\theta^{(n)}_{\zeta\wedge\eta}} - \tilde{V}^{\pi}_{\theta^{(n)}_{\zeta\wedge\eta}\wedge T}|$, $\Lambda_2 = |\tilde{V}^{\pi}_{\theta^{(n)}_{\zeta\wedge\eta}\wedge T} - \tilde{V}^{\pi}_{\theta^{(n)}_{\eta}\wedge\sigma}|$ and $\Lambda_3 = (Q^{B,I,n}(\frac{\zeta T}{n}, \frac{\eta T}{n}) - Q^{B,I_\epsilon}(\sigma, \theta^{(n)}_\eta \wedge T))^+$. By using the same arguments as in (5.12)-(5.17) of [3] we obtain that

$$E^B(\Lambda_1 + \Lambda_2) \le C(f_\delta)n^{-1/4} \tag{34}$$

for some constant $C(f_\delta)$ which depends only on f_δ. Next, we estimate Λ_3. Set

$$Q^B(s,t) = e^{-rt}G_t(S^B)\mathbb{I}_{t<s} + e^{-rs}F_s(S^B)\mathbb{I}_{s\le t}, \quad s,t \ge 0 \text{ and}$$
$$Q^{B,n}(k,l) = (1+r_n)^{-k}G_{\frac{kT}{n}}(S^{B,n})\mathbb{I}_{k<l} + (1+r_n)^{-l}F_{\frac{lT}{n}}(S^{B,n})\mathbb{I}_{l\le k}, \quad k,l \le n.$$

From the assumptions on the functions F, G we get

$$\Lambda_3 \le \left(Q^{B,n}\left(\frac{\zeta T}{n}, \frac{\eta T}{n}\right) - Q^B\left(\sigma, \theta^{(n)}_\eta \wedge T\right)\right)^+ \tag{35}$$
$$+ \mathbb{I}_\Theta\left(G_0(S_0) + \mathscr{L}(T+2)\left(1 + \max_{0\le k\le n} S^{B,n}_{\frac{kT}{n}}\right)\right) \tag{36}$$

where $\Theta = \{\zeta\wedge\eta < \tau_I^{B,n}\} \cap \{\sigma\wedge\theta^{(n)}_\eta \ge \tau_{I_\epsilon}\}$. Similarly to Lemmas 3.2 and 3.3 in [6] it follows that there exists a constant $C^{(1)}$ such that

$$\sup_{\tilde{\zeta}\in\mathscr{T}^{B,n}_{0,n}} \sup_{\tilde{\eta}\in\mathscr{T}^{B,n}_{0,n}} E^B\left|Q^B\left(\theta^{(n)}_{\tilde{\zeta}}, \theta^{(n)}_{\tilde{\eta}}\right) - Q^{B,n}\left(\frac{\tilde{\zeta}T}{n}, \frac{\tilde{\eta}T}{n}\right)\right|$$
$$\le C^{(1)}n^{-1/4}(\ln n)^{3/4}. \tag{37}$$

From (28) and the Cauchy-Schwarz inequality it follows that

$$E^B\left(\mathbb{I}_\Theta\left(G_0(S_0) + \mathscr{L}(T+2)\left(1 + \max_{0\le k\le n} S^{B,n}_{\frac{kT}{n}}\right)\right)\right) \le C^{(2)}P(\Theta)^{1/2} \tag{38}$$

for some constant $C^{(2)}$. By (31) we see that $\sigma < \theta^{(n)}_\eta \wedge T$ provided $\zeta < \eta$. This together with (35)–(38) gives

$$E^B\Lambda_3 \le C^{(2)}P(\Theta)^{1/2} + C^{(1)}n^{-1/4}(\ln n)^{3/4} + \alpha_1 + \alpha_2 \tag{39}$$

where $\alpha_1 = E^B|e^{-r\theta^{(n)}_{\zeta\wedge\eta}}G_{\theta^{(n)}_{\zeta\wedge\eta}}(S^B) - e^{-r\sigma\wedge\theta^{(n)}_\eta}G_{\sigma\wedge\theta^{(n)}_\eta}(S^B)|$ and $\alpha_2 = E^B|e^{-r\theta^{(n)}_{\zeta\wedge\eta}}\times F_{\theta^{(n)}_{\zeta\wedge\eta}}(S^B) - e^{-r\sigma\wedge\theta^{(n)}_\eta}F_{\sigma\wedge\theta^{(n)}_\eta}(S^B)|$. From Lemma 4.4 in [3] it follows that there exists a constants $C^{(3)}, C^{(4)}$ such that

$$\alpha_1+\alpha_2 \le C^{(3)}(E^B(\theta^{(n)}_{\zeta\wedge\eta}-\theta^{(n)}_\eta\wedge\sigma)^2)^{1/2}+C^{(4)}(E^B(\theta^{(n)}_{\zeta\wedge\eta}-\theta^{(n)}_\eta\wedge\sigma)^2)^{1/4}.$$

By (31) we obtain that $|\theta^{(n)}_{\zeta\wedge\eta}-\theta^{(n)}_\eta\wedge\sigma|\le|\theta^{(n)}_\zeta-\sigma|\le|T-\theta^{(n)}_n|\le u_n$. Thus by (27),

$$\alpha_1+\alpha_2\le C^{(5)}n^{-1/4} \tag{40}$$

for some constant $C^{(5)}$. Finally, we estimate $P(\Theta)$. Observe that $\sigma\wedge\theta^{(n)}_\eta\le\theta^{(n)}_{\zeta\wedge\eta}$, and so

$$\begin{aligned}\Theta &\subseteq \left\{\frac{\sup_{0\le t\le\sigma\wedge\theta^{(n)}_\eta}S^B_t}{\max_{0\le k\le\zeta\wedge\eta}S^{B,n}_{kT/n}}>e^\epsilon\right\}\bigcup\left\{\frac{\inf_{0\le t\le\sigma\wedge\theta^{(n)}_\eta}S^B_t}{\min_{0\le k\le\zeta\wedge\eta}S^{B,n}_{kT/n}}<e^{-\epsilon}\right\}\\ &\subseteq\left\{\max_{0\le k\le n-1}\sup_{\theta^{(n)}_k\le t\le\theta^{(n)}_{k+1}} r\left|t-\frac{kT}{n}\right|+\kappa\left|B^*_t-B^*_{\theta^{(n)}_k}\right|>\epsilon\right\}.\end{aligned} \tag{41}$$

Since $|B^*_t-B^*_{\theta^{(n)}_k}|\le\sqrt{\frac{T}{n}}$ and $|t-\frac{kT}{n}|\le u_n+\frac{T}{n}$ for any $k<n$ and $t\in[\theta^{(n)}_k,\theta^{(n)}_{k+1}]$ then using the inequality $(a+b)^3\le 4(a^3+b^3)$ for $a,b\ge 0$ we obtain by (27) that $E^B(\max_{0\le k\le n-1}\sup_{\theta^{(n)}_k\le t\le\theta^{(n)}_{k+1}} r|t-\frac{kT}{n}|+\kappa|B^*_t-B^*_{\theta^{(n)}_k}|)^3\le C^{(6)}n^{-3/2}$ for some constant $C^{(6)}$. From (41) and the Markov inequality it follows that $P(\Theta)\le C^{(6)}\frac{n^{-3/2}}{\epsilon^3}$ and together with (29), (33), (34), (39) and (40) we conclude that

$$\begin{aligned}R^I_n(x)-R^{I_\epsilon}(x)&\le 3\delta+(C^{(5)}+C(f_\delta))n^{-1/4}+C^{(1)}n^{-1/4}(\ln n)^{3/4}\\ &\quad+C^{(2)}\sqrt{C^{(6)}\frac{n^{-3/2}}{\epsilon^3}}.\end{aligned} \tag{42}$$

Since the above constants do not depend on n then $R^{I_\epsilon}(x)\ge\limsup_{n\to\infty}R^I_n(x)-3\delta$. By letting $\delta\downarrow 0$ and Lemma 3, $R^I(x)=\lim_{\epsilon\to 0}R^{I_\epsilon}(x)\ge\limsup_{n\to\infty}R^I_n(x)$.

In order to compete the proof of Theorem 2 we should prove (4). Fix an initial capital x, an open interval $I=(L,R)$ and a natural number n. If $\min(\frac{R}{S_0},\frac{S_0}{L})\le e^{n^{-1/3}}$ then from Lemma 4 and the inequality $R^I_n(x)\ge(F_0(S_0)-x)^+$ it follows

$$R^I(x)-R^I_n(x)\le R^I(x)-(F_0(S_0)-x)^+\le A_{\frac{4}{3}}n^{-1/4}. \tag{43}$$

Next, we deal with the case where $\min(\frac{R}{S_0},\frac{S_0}{L})>e^{n^{-1/3}}$ (which is true for sufficiently large n). Introduce the open interval $J_n=(L\exp(n^{-1/3}),R\exp(-n^{-1/3}))$. Set $(\pi,\sigma)=(\psi_n(\pi^I_n),\phi_n(\sigma^I_n))$ where (π^I_n,σ^I_n) is the optimal hedge given by (10) and the functions ψ_n,ϕ_n were defined in Sect. 2. We can consider the portfolio $\pi=\psi_n(\pi_n)$ not only as an element in $\mathscr{A}^{B,n}(x)$ but also as an element in $\mathscr{A}^B(x)$ if we restrict the above portfolio to the interval $[0,T]$. From Lemma 2 we obtain that

$$R^{J_n}(\pi,\sigma)-R^I_n(x)=R^{J_n}(\pi,\sigma)-R^{B,I}_n(\pi,\zeta^I_n) \tag{44}$$

where, recall, ζ_n^I was defined in (11). Since I and n are fixed we denote $\zeta = \zeta_n^I$. Recall that $\Pi_n(\sigma_n^I) = \zeta$ and so from (5) we get $\sigma = (T \wedge \theta_\zeta^{(n)})\mathbb{I}_{\zeta<n} + T\mathbb{I}_{\zeta=n}$. For a fixed $\delta > 0$ choose a stopping time τ such that

$$R^{J_n}(\pi,\sigma) < \delta + E^B[(Q^{B,J_n}(\sigma,\tau) - \tilde{V}^\pi_{\sigma\wedge\tau})^+]. \tag{45}$$

Set $\eta = n \wedge \min\{k|\theta_k^{(n)} \geq \tau\} \in \mathcal{T}_{0,n}^{B,n}$. Denote $\Gamma_1 = (Q^{B,J_n}(\sigma,\tau) - Q^{B,J_n}(\sigma \wedge \theta_n^{(n)}, \tau \wedge \theta_n^{(n)}))^+$ and $\Gamma_2 = (Q^{B,J_n}(\sigma \wedge \theta_n^{(n)}, \tau\wedge\theta_n^{(n)}) - Q^{B,I,n}(\frac{\zeta T}{n}, \frac{\eta T}{n}))^+$. From (45) it follows that

$$\begin{aligned} R^{J_n}(\pi,\sigma) - R_n^{B,I}(\pi,\zeta) &< E^B(Q^{B,J_n}(\sigma\wedge\theta_n^{(n)}, \tau\wedge\theta_n^{(n)}) - \tilde{V}^\pi_{\sigma\wedge\tau})^+ \\ &-E^B(Q^{B,J_n}(\sigma\wedge\theta_n^{(n)}, \tau\wedge\theta_n^{(n)}) - \tilde{V}^\pi_{\theta^{(n)}_{\zeta\wedge\eta}})^+ + \delta + E^B(\Gamma_1+\Gamma_2). \end{aligned} \tag{46}$$

Since $\pi \in \mathcal{A}^{B,n}(x)$ then by (6), $\tilde{V}^\pi_{\sigma\wedge\tau} = \tilde{V}^\pi_{\sigma\wedge\tau\wedge\theta_n^{(n)}} = \tilde{E}^B(\tilde{V}^\pi_{\theta^{(n)}_{\zeta\wedge\eta}}|\mathcal{F}^B_{\sigma\wedge\tau\wedge\theta_n^{(n)}})$. This together with the Jensen inequality yields that

$$\begin{aligned} &E^B(Q^{B,J_n}\left(\sigma\wedge\theta_n^{(n)}, \tau\wedge\theta_n^{(n)}\right) - \tilde{V}^\pi_{\sigma\wedge\tau})^+ \\ &\leq E^B\left(\frac{Z_{\sigma\wedge\tau\wedge\theta_n^{(n)}}}{Z_{\theta^{(n)}_{\zeta\wedge\eta}}}\left(Q^{B,J_n}(\sigma\wedge\theta_n^{(n)}, \tau\wedge\theta_n^{(n)}) - \tilde{V}^\pi_{\theta^{(n)}_{\zeta\wedge\eta}}\right)^+\right). \end{aligned} \tag{47}$$

By (46) and (47) we obtain that

$$R^{J_n}(\pi,\sigma) - R_n^{B,I}(\pi,\zeta) < \delta + E^B(\Gamma_1+\Gamma_2) + \alpha_3 \tag{48}$$

where $\alpha_3 = E^B\left(\frac{Z_{\sigma\wedge\tau\wedge\theta_n^{(n)}} - Z_{\theta^{(n)}_{\zeta\wedge\eta}}}{Z_{\theta^{(n)}_{\zeta\wedge\eta}}} Q^{B,J_n}(\sigma\wedge\theta_n^{(n)}, \tau\wedge\theta_n^{(n)})\right)$. Since $\Gamma_1 \leq (Q^B(\sigma,\tau) - Q^B(\sigma\wedge\theta_n^{(n)}, \tau\wedge\theta_n^{(n)}))^+$, then by using the same arguments as in (5.25)-(5.29) of [3] we obtain that

$$\alpha_3 + E^B\Gamma_1 \leq C^{(7)} n^{-1/4} \tag{49}$$

for some constant $C^{(7)}$. From the assumptions on the functions F, G it follows that

$$\begin{aligned} \Gamma_2 \leq &\left(Q^B(\sigma\wedge\theta_n^{(n)}, \tau\wedge\theta_n^{(n)}) - Q^{B,n}\left(\frac{\zeta T}{n}, \frac{\eta T}{n}\right)\right)^+ \\ &+\mathbb{I}_{\tilde{\Theta}}\left(G_0(S_0) + \mathcal{L}(T+2)\left(1+\sup_{0\leq t\leq T} S_t^B\right)\right) \end{aligned} \tag{50}$$

where $\tilde{\Theta} = \{\eta\wedge\zeta \geq \tau_I^{B,n}\} \cap \{\sigma\wedge\tau\wedge\theta_n^{(n)} < \tau_{J_n}\}$. By the Cauchy-Schwarz inequality,

$$E^B\mathbb{I}_{\tilde{\Theta}}\left(G_0(S_0) + \mathcal{L}(T+2)\left(1+\sup_{0\leq t\leq T} S_t^B\right)\right) \leq C^{(8)}\left(P\left(\tilde{\Theta}\right)\right)^{1/2} \tag{51}$$

for some constant $C^{(8)}$. From (37), (50),(51)

$$\begin{aligned}E^B\Gamma_2 &\le E^B\left(Q^B\left(\sigma\wedge\theta_n^{(n)},\tau\wedge\theta_n^{(n)}\right)-Q^B(\theta_\zeta^{(n)},\theta_\eta^{(n)})\right)^+\\&\quad+C^{(1)}n^{-1/4}(\ln n)^{3/4}+C^{(8)}(P(\tilde\Theta))^{1/2}\\&\le C^{(1)}n^{-1/4}(\ln n)^{3/4}+C^{(8)}(P(\tilde\Theta))^{1/2}+C^{(9)}n^{-1/4}\end{aligned}\tag{52}$$

for some constant $C^{(9)}$. The last inequality follows from (5.29)–(5.30) in [3]. Finally, we estimate $P(\tilde\Theta)$. Observe that $\sigma\wedge\tau\wedge\theta_n^{(n)}\ge\theta^{(n)}_{(\zeta\wedge\eta-1)^+}$. Thus

$$\begin{aligned}\tilde\Theta &\subseteq\left\{\frac{\max_{0\le k\le\zeta\wedge\eta}S^{B,n}_{kT/n}}{\sup_{0\le t\le\sigma\wedge\tau\wedge\theta_n^{(n)}}S_t^B}>e^{n^{-1/3}}\right\}\bigcup\left\{\frac{\min_{0\le k\le\zeta\wedge\eta}S^{B,n}_{kT/n}}{\inf_{0\le t\le\sigma\wedge\tau\wedge\theta_n^{(n)}}S_t^B}<e^{-n^{-1/3}}\right\}\\&\subseteq\left\{\max_{0\le k\le n-1}\left(r\left|\theta_{k+1}^{(n)}-\frac{kT}{n}\right|+\kappa\left|B^*_{\theta^{(n)}_{k+1}}-B^*_{\theta^{(n)}_k}\right|\right)>n^{-1/3}\right\}\\&\subseteq\left\{r(u_n+w_n)+\kappa\sqrt{\frac{T}{n}}>n^{-1/3}\right\}.\end{aligned}\tag{53}$$

From (27), (53) and the Markov inequality it follows that

$$P(\tilde\Theta)\le nE^B\left(r(u_n+w_n)+\kappa\sqrt{\frac{T}{n}}\right)^3\le C^{(10)}n^{-1/2}\tag{54}$$

for some constant $C^{(10)}$. Since δ is arbitrary then combining (44), (48), (49), (52) and (54) we conclude that there exists a constant $C^{(11)}$ such that

$$R^{J_n}(\pi,\sigma)-R_n^I(x)=R^{J_n}(\pi,\sigma)-R_n^{B,I}(\pi,\zeta)\le C^{(11)}n^{-1/4}(\ln n)^{3/4}.\tag{55}$$

By (55) and Lemma 3 it follows that for n which satisfies $\min(\frac{R}{S_0},\frac{S_0}{L})>e^{n^{-1/3}}$ we have

$$\begin{aligned}R^I(x)-R_n^I(x)&\le R^I(x)-R^{J_n}(x)+R^{J_n}(\pi,\sigma)-R_n^I(x)\le A_{\frac{4}{3}}n^{-1/4}\\&\quad+C^{(11)}n^{-1/4}(\ln n)^{3/4}.\end{aligned}\tag{56}$$

From (43) and (56) we derive (4) and complete the proof of Theorem 2. □

Next, we prove Theorem 3. Let $H=(L,R)$ be an open interval as before and for any n set $H_n=(L\exp(-n^{-1/3}),R\exp(n^{-1/3}))$. Fix n and let $(\pi_n^{H_n},\sigma_n^{H_n})\in\mathscr{A}^{\xi,n}(x)\times\mathscr{T}^\xi_{0n}$ be the optimal hedge given by (10). For any $\epsilon>0$ denote $J_\epsilon=(Le^{-\epsilon},Re^{\epsilon})$. Since $H_n\subseteq J_\epsilon$ for sufficiently large n, then from (55) (for $I=H_n$), Lemma 3 and Theorem 2 we obtain

$$\begin{aligned}\limsup_{n\to\infty}R^H(\psi_n(\pi_n^{H_n}),\phi_n(\sigma_n^{H_n}))&\le\limsup_{n\to\infty}R_n^{H_n}(x)\\&\le\lim_{\epsilon\to0}\limsup_{n\to\infty}R_n^{J_\epsilon}(x)=\lim_{\epsilon\to0}R^{J_\epsilon}(x)=R^H(x)\end{aligned}$$

which completes the proof of Theorem 3. □

Next, we prove Theorem 1. Let $I=(L,R)$ be an open interval as before. Assume that $\mu=0$. In this case $\mathcal{V}^I=R^I(0)$ and $\mathcal{V}_n^I=R_n^I(0)$. Hence by using (13), (42) for $(\epsilon=n^{-1/3}, q=\frac{4}{3})$ and taking into account that the value of the portfolios π,π' is zero (which means that $C(f_\delta)=0$) we obtain $\mathcal{V}_n^I-\mathcal{V}^I\leq C^{(12)}n^{-1/4}(\ln n)^{3/4}$ for some constant $C^{(12)}$. This together with (4) completes the proof of Theorem 1. □

Finally, we prove Theorem 4. Let $H=(L,R)$ be an open interval and n be a natural number. Set $H_n=(L\exp(-n^{-1/3}), R\exp(n^{-1/3}))$ and let $(\pi_n,\sigma_n)\in\mathcal{A}^{\xi,n}(\mathcal{V}_n^{H_n})\times\mathcal{T}_{0n}^{\xi}$ be a perfect hedge for a double barrier option in the n–step CRR market with the barriers $L\exp(-n^{-1/3})$, $R\exp(n^{-1/3})$. Set $(\pi,\zeta)=(\psi_n(\pi_n), \Pi_n(\sigma_n))\in\mathcal{A}^{B,n}(V_n^{H_n})\times\mathcal{T}_{0,n}^{B,n}$. From the definition of a perfect hedge and Π_n it follows that for any $k\leq n$, $\tilde{V}^{\pi}_{\zeta\wedge k}=\Pi_n(\tilde{V}^{\pi_n}_{\sigma_n\wedge k})\geq\Pi_n(Q^{H_n,n}(\sigma_n,k))=Q^{B,H_n,n}(\zeta,k)$, implying that $R_n^{B,H_n}(\pi,\zeta)=0$. Set $\sigma=\phi_n(\sigma_n)\in\mathcal{T}_{0T}^{B}$ then $\sigma=(T\wedge\theta_\zeta^{(n)})\mathbb{I}_{\zeta<n}+T\mathbb{I}_{\zeta=n}$. Hence, using (55) for $I=H_n$ we obtain that $R^H(\pi,\sigma)\leq R_n^{B,H_n}(\pi,\zeta)+C^{(11)}n^{-1/4}(\ln n)^{3/4}=C^{(11)}n^{-1/4}(\ln n)^{3/4}$ completing the proof. □

Acknowledgements This work was partially supported by the ISF grant no. 130/06.

References

1. Cox, J.C., Ross, R.A., Rubinstein, M.: Option pricing: a simplified approach. J. Financ. Econom. **7**, 229–263 (1976)
2. Dolinsky, Y., Kifer, Y.: Hedging with risk for game options in discrete time. Stochastics **79**, 169–195 (2007)
3. Dolinsky, Y., Kifer, Y.: Binomial approximations of shortfall risk for game options. Ann. Appl. Probab. **18**, 1737–1770 (2008)
4. Karatzas, I., Shreve, S.E.: Brownian motion and Stochastic calculus. Springer, New York (1991)
5. Kifer, Y.: Game options. Finance and Stoch. **4**, 443–463 (2000)
6. Kifer, Y.: Error estimates for binomial approximiations of game options. Ann. Appl. Probab. **16**, 984–1033 (2006); Correction: **18**, 1271–1277 (2008)
7. Ohtsubo,: Optimal stopping in sequential games with or without a constraint of always terminating. Math. Oper Res. **11**, 591–607 (1986)
8. Rogers, L.C.G., Stapelton, E.J.: Fast accurate binomial pricing. Finance and Stoch. **2**, 3–17 (1998)
9. Shiryaev, A.N., Kabanov, Y.M., Kramkov, D.O., Melnikov, A.B.: To the theory of computations of European and American options, discrete case. Th. Probab. Appl. **39**, 14–60 (1994); Continuous case. Th. Probab. Appl. **39**, 61–102 (1994)

A Game of International Climate Policy Solved by a Homogeneous Oracle-Based Method for Variational Inequalities

Laurent Drouet, Alain Haurie, Jean-Philippe Vial, and Marc Vielle

Abstract This paper presents a game-theoretic model for the international negotiations that should take place to renew or extend the Kyoto protocol beyond 2012. These negotiations should lead to a self-enforcing agreement on a burden sharing scheme to realize the necessary global emissions abatement that would preserve the world against irreversible ecological impacts. The model assumes a noncooperative behavior of the parties except for the fact that they will be collectively committed to reach a target on cumulative emissions by the year 2050. The concept of normalized equilibrium, introduced by J.B. Rosen for concave games with coupled constraints, is used to characterize a family of dynamic equilibrium solutions in an m-player game where the agents are (groups of) countries and the payoffs are the welfare gains obtained from a Computable General Equilibrium (CGE) model. The model is solved using an homogeneous version of the oracle-based optimization engine (OBOE) permitting an implicit definition of the payoffs to the different players, obtained through simulations performed with the global CGE model GEMINI-E3.

L. Drouet (✉)
Resource Centre for Environmental Technologies, Public Research Centre Henri Tudor, Technoport Schlassgoart, 66 rue de Luxembourg, L-4002 Esch-sur-Alzette, Luxembourg
e-mail: laurent.drouet@tudor.lu

A. Haurie
ORDECSYS, Chêne-Bougeries, Switzerland
e-mail: ahaurie@ordecsys.com

J.-P. Vial
University of Geneva and ORDECSYS, Switzerland
e-mail: jpvial@ordecsys.com

M. Vielle
École Polytechnique Fédérale de Lausanne and C-ORDEE, Switzerland and Toulouse School of Economics (LERNA), TSE - Manufacture de Tabacs, 21 allées de Brienne - 31000 Toulouse, France
e-mail: marc.vielle@epfl.ch

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_23,

1 Introduction

This paper is a continuation of the work already reported in [8]. In this research, we study the strategic interactions of groups of countries when they negotiate the sharing of burden to stabilize the long-term concentration of greenhouse gases (GHG). We propose a dynamic game model where the strategies of each player (state or group of countries) refer to the timing of supply of emission rights (also called quotas) on an international emissions trading scheme with full banking and borrowing. The payoffs are obtained in terms of welfare gains (or losses) compared with the business as usual (BAU) situation. In our approach, these welfare gains are obtained through simulations performed with a Computable General Equilibrium (CGE) model. In this game, a coupled constraint is imposed on all players together to limit the cumulative emissions over the whole planning horizon. Games with coupled constraints have been first studied by Rosen [23] who showed that a whole family (manifold) of equilibrium outcomes should be expected, indexed over a set of weights attributed to the players. The appropriateness of the normalized equilibrium concept to deal with environmental games has been first recognized by Haurie [11] and further explored by Haurie and Zaccour [14], Haurie and Krawczyk [12], and Krawczyk [15], [16].

In this paper, we explore the manifold of normalized equilibria in a deterministic framework. In [8] a stochastic version of this game is also proposed, using the concept of S-adapted equilibria (see [13]). The added contribution of this paper compared with [8] is twofold: (i) we show that the Rosen normalized equilibria can be represented as Nash equilibria associated with different sharing schemes of the cumulative emission budget; (ii) we resort to optimization theory for general convex programming to solve the challenging equilibrium problem formulated as a system of variational inequalities. The algorithm we use is a homogeneous version of ACCPM[1] devised by Nesterov and Vial and studied in-depth in [20, 22]. The method is *oracle-based*, in that it uses a black-box mechanism, named *oracle*, to collect first-order information (function values and subgradients) of the functions entering the problem definition. This approach provides a complexity bound on the required number of iterations to reach a given level of precision. We shall also report on the most recent simulations performed with this model.

This paper is organized as follows: we present the structure of the coupled game model in Sect. 2. Then, in Sect. 3, we briefly present the multisector and multi-country CGE model of the world economy that is used for getting the payoffs; We implement the model within an oracle-based optimization framework with the homogeneous version of the Analytic Center Cutting Plane Method (ACCPM); in Sect. 5, we give the numerical results obtained for a case study where countries have to decide on their own abatement level under a global target on cumulative

[1] Acronym for Analytic Center Cutting Plane Method [19]

GHG emissions by 2050 which is consistent with a commitment to limit global temperature rise to 2 degrees Celsius above pre-industrial levels.

2 The Model

In this section, we present a deterministic game model for GHG emissions abatement, in a simple formulation on a time span comprising four periods between 2010 and 2050. We summarize the negotiation as a two-stage process: in stage 1, the negotiators determine a cumulative emission budget over the period 2010–2050 and define a sharing of this emission budget, called respective total allowances, over the set of players (each player is a group of countries); in stage 2, the countries use their total allowance to supply dynamically the international emissions trading market. In this second stage, the players behave as oligopolists supplying a sequence of markets that change over time, due to economic growth and technological progress. We do not represent stage 1 as a noncooperative game but we assume that the equilibrium outcome of stage 2 will be taken into consideration for designing a fair sharing of the cumulative emission budget. We assume that, in stage 1, the cumulative emission budget is dictated by a precautionary principle to keep the radiative forcing within a tolerable limit.

2.1 *Players, Moves, and Payoffs in the Stage 2 Game*

The game is played over 4 periods $t = 1, \ldots, 4$. M is a set of m groups of countries hereafter called players which must decide on the emission quotas they supply in each period on an international GHG emissions trading market.

We denote $\bar{e}_j(t)$ the supply of quotas decided by player j for period t. A global limit $\bar{E}$ is imposed on the cumulative emissions over the four periods $t = 1, \ldots, 4$. Therefore, the following coupled constraints are binding all players together

$$\sum_{j \in M} \sum_{t=1}^{4} \bar{e}_j(t) \le \bar{E}. \tag{1}$$

Let $\bar{\mathbf{e}}(t) = \{\bar{e}_j(t)\}_{j \in M}$ denote the vector of emissions quotas for all players in period t. Given these quotas a general economic equilibrium is computed for the m-countries which determines a welfare gain for each player, hereafter called its payoff at t and denoted $W_j(t; \bar{e}(t))$.. Given a choice of emission quotas $\bar{\mathbf{e}} = \{\bar{\mathbf{e}}(t) \mid t = 1, \ldots, 4\}$ the payoff to player j is given by

$$J_j(\bar{\mathbf{e}}) = \sum_{t=1}^{4} \beta^{t-1} W_j(t; \bar{\mathbf{e}}(t)) \quad j \in M,$$

where β is a common discount factor.

Remark 1. The game is not defined on a dynamic system, e.g., a differential game.[2] The dynamic effect will be essentially associated with the supply over time of the respective total allowances. Therefore, there is no end-of-period effect. The limitation to four periods corresponds to the current situation where an international GHG-emission agreement is envisioned for a period extending roughly from 2010 to 2050; whence the consideration of four 10-year periods.

2.2 Normalized Equilibrium Solutions

We assume that the players behave in a noncooperative way but are bound to satisfy the global cumulative emissions constraints (1). The solution concept that we propose to use is related to the concept of normalized equilibrium introduced by Rosen [23] to deal with games where the players are bound by a coupled constraint.

Let us call $\mathscr{E}$ the set of emissions $\bar{\mathbf{e}}$ that satisfy the constraints (1). Denote also $[\bar{\mathbf{e}}^{*j}, \bar{e}_j]$ the emission program obtained from $\bar{\mathbf{e}}^*$ by replacing only the emission program $\bar{\mathbf{e}}^*_j$ of player j by $\bar{e}_j$.

Definition 1. The emission program $\bar{\mathbf{e}}^*_j$ is an equilibrium under the coupled constraints defined in (1) if the following holds for each player $j \in M$.

$$\bar{\mathbf{e}}^* \in \mathscr{E}$$
$$\forall \bar{e}_j \text{s.t.} [\bar{\mathbf{e}}^{*j}, \bar{e}_j] \in \mathscr{E} \quad J_j(\bar{\mathbf{e}}^*) \geq J_j([\bar{\mathbf{e}}^{*j}, \bar{e}_j])$$

Therefore, in this equilibrium, each player replies optimally to the emission program chosen by the other players, under the constraint that the global cumulative emission limits must be respected.

It is possible to characterize a class of such equilibria through a fixed point condition for a best reply mapping defined as follows. Let $r = (r_j)_{j \in M}$ with $r_j > 0$ and $\sum_{j \in M} r_j = 1$ be a given weighting of the different players. Then introduce the combined response function

$$\theta(\bar{\mathbf{e}}^*, \bar{\mathbf{e}}; r) = \sum_{j \in M} r_j J_j([\bar{\mathbf{e}}^{*j}, \bar{e}_j]). \tag{2}$$

It is easy to verify that, if $\bar{\mathbf{e}}^*$ satisfies the fixed point condition

$$\theta(\bar{\mathbf{e}}^*, \bar{\mathbf{e}}^*; r) = \max_{\bar{\mathbf{e}} \in \mathscr{E}} \theta(\bar{\mathbf{e}}^*, \bar{\mathbf{e}}; r), \tag{3}$$

then it is an equilibrium under the coupled constraint.

[2] The economic model used to define the payoffs of the game, GEMINI-E3, is a time-stepped model, which computes an equilibrium at each period and defines investments through an exogenously defined saving function. Therefore, there is no hereditary effect of the allocation of emission quotas and, as a consequence there is no need for a scrap value function.

Definition 2. The emission program $\bar{\mathbf{e}}^*$ is a normalized equilibrium if it satisfies (3) for a weighting r and a combined response function defined as in (2).

The RHS of (3) defines an optimization problem under constraint. Assuming the required regularity we can introduce a Kuhn–Tucker multiplier λ^0 for the constraint $\sum_{t=0}^{1} \bar{e}_j(t) \leq \bar{E}$ and form the Lagrangian

$$L = \theta(\bar{\mathbf{e}}^*, \bar{\mathbf{e}}; r) + \lambda^0 \left(\bar{E} - \sum_{j \in M} \sum_{t=1}^{4} \bar{e}_j(t) \right).$$

Therefore, by applying the standard K-T optimality conditions we can see that the normalized equilibrium is also the Nash equilibrium solution for an auxiliary game with a payoff function defined for each player j by

$$J_j(\bar{\mathbf{e}}) + \lambda^j \left(\bar{E} - \sum_{j \in M} \sum_{t=1}^{4} \bar{e}_j(t) \right),$$

where

$$\lambda^j = \frac{1}{r_j} \lambda^0.$$

This characterization has an interesting interpretation in terms of negotiation for a climate change policy. A common "tax" λ^0 is defined and applied to each player with an intensity $\frac{1}{r_j}$ that depends on the weight given to this player in the global response function.

2.3 An Interpretation as a Distribution of a Cumulative Budget

Consider an m-player concave game à la Rosen with payoff functions

$$\psi_j(x_1, x_2, \ldots, x_m), \quad x_j \in X_j \quad j = 1, \ldots, m,$$

and a coupled constraint. When the coupled constraint is scalar and separable among players, i.e. when it takes the form

$$\sum_{j=1}^{m} \varphi_j(x_j) = e,$$

the coupled equilibrium can be interpreted in an interesting way. Consider e as being a global allowance and call $\varpi_j \geq 0$ the fraction of this allowance given to player j, with $\sum \varpi_j = 1$. Then define the game with payoffs and decoupled constraints

$$\psi_j(x_1, x_2, \ldots, x_m), \quad x_j \in X_j \quad \varphi_j(x_j) \leq \varpi_j e \quad j = 1, \ldots, m.$$

A Nash equilibrium for this game is characterized, under the usual regularity conditions, by the following conditions

$$\begin{array}{rcl} & \max_j & \psi_j(x_1,\ldots,x_j,\ldots,x_m) - \lambda_j \varphi_j(x_j) \\ \lambda_j & \geq & 0 \\ 0 & = & \lambda_j(\varphi_j(x_j) - \varpi_j e). \end{array}$$

Now assume that at the equilibrium solution all the constraints are active and hence one may expect (in the absence of degeneracy) that all the λ_j are > 0. Since the multipliers are scalars they can be written in the form

$$\lambda_j = \frac{\lambda_0}{r_j},$$

by taking

$$\lambda_0 = \sum_{j=1}^{m} \lambda_j$$

and defining

$$r_j = \frac{\lambda_0}{\lambda_j}, \quad j = 1,\ldots,m.$$

The assumption of active constraints at equilibrium leads to

$$\lambda_0 > 0 \quad \text{and} \quad \sum_{j=1}^{m} \varphi_j(x_j) - e = 0.$$

Therefore, the conditions for a normalized coupled equilibrium are met.

Notice that the assumption that all constraints are active at the Nash equilibrium is crucial. However, in a climate game where the constraints are on emissions quotas this assumptions is very likely to hold. So, in the special situation described here a normalized equilibrium is obtained by defining first a sharing of the common global allowance and then by playing the noncooperative game with the distributed constraints.

2.4 Subgame Perfectness

The game is played in a dynamic setting and the question of subgame perfectness arises naturally. The equilibrium solutions we consider are open-loop and therefore do not possess the subgame perfectness property. In fact, even though the model involves four periods, the equilibrium concept is "static" in spirit. It is used as a way to value the outcome of some sharing of a global emission budget, negotiated in stage 1.

2.5 Stage 1 Negotiations

We view the negotiations in stage 1, not as a noncooperative game, but rather as a search for equity in a "pie sharing process". More precisely, a successful international climate negotiation could involve the following agreements among different groups of countries, each group sharing similar macroeconomic interest:

- the total level of cumulative GHG emissions allowed over a part of the 21^{st} century (typically 2010–2050), for instance based on a precautionary principle involving the overall temperature increase not to be exceeded;
- the distribution of this cumulative emission budget among the different groups, for instance using some concepts of equity;
- the fairness of the sharing will be evaluated on the basis of the equilibrium solution obtained in stage 2.

In the rest of this paper, we shall concentrate on the way to evaluate the equilibrium outcomes of the stage 2 noncooperative game.

3 Getting the Payoffs Via GEMINI-E3

This section, largely reproduced from [8], shows how the payoffs of the game are obtained from economic simulations performed with a computable general equilibrium (CGE) model.

3.1 General Overview

The payoffs of the game are computed using the GEMINI-E3 model. We use the fifth version of GEMINI-E3 describing the world economy in 28 regions with 18 sectors, and which incorporates a highly detailed representation of indirect taxation [5]. This version of GEMINI-E3 is formulated as a Mixed Complementarity Problem (MCP) using GAMS with the PATH solver [9, 10]. GEMINI-E3 is built on a comprehensive energy-economy data set, the GTAP-6 database [7], that expresses a consistent representation of energy markets in physical units as well as a detailed Social Accounting Matrix (SAM) for a large set of countries or regions and bilateral trade flows. It is the fifth GEMINI-E3 version that has been especially designed to calculate the social marginal abatement costs (MAC), i.e. the welfare loss of a unit increase in pollution abatement [4]. The different versions of the model have been used to analyze the implementation of economic instruments for GHG emissions in a second-best setting [3], to assess the strategic allocation of GHG emission allowances in the EU-wide market [6] and to analyze the behavior of Russia in the Kyoto Protocol [2, 4].

For each sector, the model computes the demand on the basis of household consumption, government consumption, exports, investment, and intermediate uses. Total demand is then divided between domestic production and imports, using the Armington assumption [1]. Under this convention, a domestically produced good is treated as a different commodity from an imported good produced in the same industry. Production technologies are described using nested CES functions.

3.2 Welfare Cost

Household's behavior consists of three interdependent decisions: 1) labor supply; 2) savings; and 3) consumption of the different goods and services. In GEMINI-E3, we suppose that labor supply and the rate of saving are exogenously fixed. The utility function corresponds to a Stone–Geary utility function [25] which is written as:

$$u_r = \sum_i \beta_{ir} \ln(HC_{ir} - \phi_{ir}),$$

where HC_{ir} is the household consumption of product i in region[3] r, ϕ_{ir} represents the minimum necessary purchases of good i, and β_{ir} corresponds to the marginal budget share of good i. Maximization under budgetary constraint:

$$HCT_r = \sum_i PC_{ir} HC_{ir}$$

yields

$$HCI_{ir} = \phi_{ir} + \frac{\beta_{ir}}{PC_{ir}} \left[HCT_r - \sum_k (PC_{kr}\phi_{kr}) \right],$$

where PC_{ir} is the price of household consumption for product i in region r.

The welfare cost of climate policies is measured comprehensively by changes in households' welfare since final demand of other institutional sectors is supposed unchanged in scenarios. Measurement of this welfare change is represented by the sum of the change in income and the "Equivalent Variation of Income" (EVI) of the change in prices, according to the classical formula. In the case of a Stone–Geary utility function, the EVI for a change from an initial situation defined by the price system ($\overline{PC_{ir}}$) to a final situation (PC_{ir}) is such as

$$\frac{\overline{HCT}_r - \sum_i \overline{PC}_{ir}\phi_{ir}}{\prod_i (\overline{PC}_{ir})^{\beta_{ir}}} = \frac{\overline{HCT}_r + EVI_r - \sum_i PC_{ir}\phi_{ir}}{\prod_i (PC_{ir})^{\beta_{ir}}}.$$

[3] The attentive reader should not be confused with the use of r as the index of regions, whereas r_j has been used before as a weight in the definition of a normalized equilibrium. The notations in this section, devoted to a brief presentation of GEMINI-E3, are not exactly the same as in the rest of the paper which describes the game and the solution method used to solve it. The same remark applies to the use of the symbols ϕ and i in this section vs. the rest of the paper.

The households' surplus is then given by

$$S_r = \left(HCT_r - \sum_i PC_{ir}\phi_{ir}\right) - \prod_i \left(\frac{PC_{ir}}{\overline{PC}_{ir}}\right)^{\beta_{ir}} \left(\overline{HCT}_r - \sum_i \overline{PC}_{ir}\phi_{ir}\right).$$

In summary, the CGE model associates a welfare gain (cost) for each country and each period with a given emissions program $\bar{e}$, which defines quotas for all countries at each period. It is important to notice that these welfare gains are obtained under the assumption that an international emissions trading scheme is put in place.

3.3 Getting the Payoffs of the Emission Quota Game

To summarize this game model, the players (i.e. the groups of regions) strategies correspond to allocations of their respective total allowances among the time periods. This determines quotas for each period. A general economic equilibrium problem is solved by GEMINI-E3 at each period. On the basis of the quotas chosen by the players, at each period, supplies and demands of emission permits are balanced and a permit price is obtained on an international carbon market. Countries which are net suppliers of permits receive revenue and in contrary countries buying permits must pay for it.[4] These financial transfers are collected by governments and are redistributed to households. These transfers influence the whole balance in the economy and determine, at the end, the terms of trade and the surpluses of each countries. The payoff $W_j(t, \bar{\mathbf{e}}(t))$ of Player (group of region) j is obtained as the discounted value of the sum of households surpluses S_j of player j on the period 2005–2050, using a 5% discount rate.

4 Oracle-Based Optimization Framework

There are different numerical methods to compute a Nash or Rosen-normalized equilibrium. Indeed one may rely to the methods solving variational inequalities [21]. Another approach is proposed in [27] and [26]. In our problem, the payoffs to players are obtained from a comprehensive economic simulation and, therefore, we do not have an explicit and analytical description of the payoff functions. This precludes the use of the various methods proposed for solving variational inequalities or concave games which exploit the analytical form of the payoff functions and their derivatives. Oracle-based methods can be used in our context, since they require only to obtain at each step a numerical information of the oracle and of a sub-(pseudo) gradient.

---

[4] Indeed, the buyer prefers to pay the permit price rather than support its own marginal abatement cost, while the seller can abate at a marginal cost which is lower than the market price.

In this section, we describe the implementation of an oracle-based optimization method to compute a solution to the variational inequality which characterizes the equilibrium solution. Since this implementation of an oracle-based optimization to compute Rosen normalized equilibrium solutions seems to be one of the first applications of the method, we shall develop this description in enough details to give the reader a precise idea of the mathematics involved.

4.1 Normalized Equilibrium and Variational Inequality

For concave games with differentiable payoff functions $J_j(\cdot)$, $\bar{\mathbf{e}}^*$ is a normalized equilibrium if and only if it is a solution of the following variational inequality problem

$$\langle F(\bar{\mathbf{e}}^*), \bar{\mathbf{e}}^* - \bar{\mathbf{e}} \rangle \geq 0 \quad \forall \bar{\mathbf{e}} \in \mathscr{E}, \tag{4}$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product and the pseudogradient $F(\cdot)$ is defined by

$$F(\bar{\mathbf{e}}) = \begin{pmatrix} r_1 \nabla_{\bar{e}_1} J_1(\bar{\mathbf{e}}) \\ \vdots \\ r_j \nabla_{\bar{e}_j} J_j(\bar{\mathbf{e}}) \\ \vdots \\ r_m \nabla_{\bar{e}_m} J_m(\bar{\mathbf{e}}) \end{pmatrix}.$$

It has been proved in [23] that a normalized equilibrium exists if the payoff functions $J_j(\cdot)$ are continuous in $\bar{\mathbf{e}}$ and concave in $\bar{e}_j$ and if $\mathscr{E}$ is compact. In the same reference, it is proved also that the normalized equilibrium is unique if the function $-F(\cdot)$ is strictly monotone, *i.e.* if the following holds

$$\langle F(\bar{\mathbf{e}}^2) - F(\bar{\mathbf{e}}^1), \bar{\mathbf{e}}^1 - \bar{\mathbf{e}}^2 \rangle > 0 \quad \forall \bar{\mathbf{e}}^1 \in \mathscr{E}, \forall \bar{\mathbf{e}}^2 \in \mathscr{E}.$$

Remark 2. In the case where one computes a Nash equilibrium with decoupled constraints, the formulation remains the same, except that the same solution would be obtained for different weights $r_j > 0; j = 1, \ldots, m$. In that case one takes $r_j \equiv 1; j = 1, \ldots, m$.

Solving (4) is a most challenging problem. If $-F$ is monotone, the above variational inequality implies the weaker one

$$\langle F(\bar{\mathbf{e}}), \bar{\mathbf{e}}^* - \bar{\mathbf{e}} \rangle \geq 0 \quad \forall \bar{\mathbf{e}} \in \mathscr{E}. \tag{5}$$

The converse is not true in general, but it is known to hold under the assumption that the monotone operator $-F$ is continuous or maximal monotone.

The weak variational inequality (5) can be formulated as a convex optimization problem. To this end, one defines the so-called *dual-gap* function

$$\phi_D(\bar{\mathbf{e}}) = \min_{\bar{\mathbf{e}}' \in \mathscr{E}} \langle F(\bar{\mathbf{e}}'), \bar{\mathbf{e}} - \bar{\mathbf{e}}' \rangle.$$

This function is concave and nonpositive. Unfortunately, computing the value of the function at some $\bar{\mathbf{e}}$ amounts to solving a nonlinear, nonconvex problem. This disallows the use of standard convex optimization techniques. However, the definition of the dual gap function provides an easy way to compute a piece-wise linear outer approximation. Indeed, let $\bar{\mathbf{e}}^k$ be some point, we have for all optimal $\bar{\mathbf{e}}^* \in \mathscr{E}$

$$\langle F(\bar{\mathbf{e}}^k), \bar{\mathbf{e}}^* - \bar{\mathbf{e}}^k \rangle \geq 0. \tag{6}$$

This property can be used in a cutting plane scheme that will be described in the next section. Figure 1 illustrates this property.

Prior to presenting the solution method, we point out that we cannot expect to find an exact solution. Therefore, we must be satisfied with an approximate solution. We say that $\bar{\mathbf{e}}$ is an ε-approximate weak solution, in short an ε-solution, if

$$\phi_D(\bar{\mathbf{e}}) \geq -\varepsilon. \tag{7}$$

As it will be stated in the next section, it is possible to give a bound on the number of iterations of the cutting plane method to reach a weak ε-solution. Of course, the algorithm may reach an ε-solution earlier, but it cannot be checked directly since the dual gap is not computable. In practice, we use another function, the so-called primal gap function

$$\phi_P(\bar{\mathbf{e}}) = \min_{\bar{\mathbf{e}}' \in \mathscr{E}} \langle F(\bar{\mathbf{e}}), \bar{\mathbf{e}} - \bar{\mathbf{e}}' \rangle.$$

Note that the pointwise computation of ϕ_P amounts to solving a linear programming problem when $\mathscr{E}$ is linear, an easy problem. Since $-F$ is monotone, we have

$$\phi_P(\bar{\mathbf{e}}) \leq \phi_D(\bar{\mathbf{e}}).$$

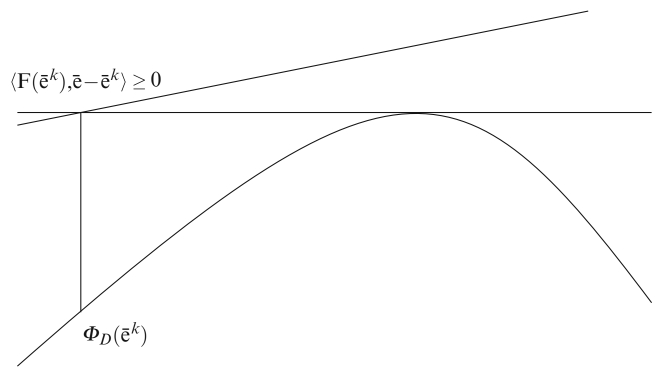


Fig. 1 Outer approximation scheme of ϕ_D

Therefore, one can substitute to the inoperative criterion (7), the more practical one

$$\phi_P(\bar{\mathbf{e}}) \geq -\varepsilon.$$

The reader must be warned that the convex optimization problem of maximizing the concave dual gap function is very peculiar. The oracle that would compute the dual gap function value would consist in solving a nonlinear, nonconvex minimization problem, a computationally intractable problem. On the other hand, the dual gap is bounded above by 0 and this bound is the optimal value. The challenge is to find a point that approximates this optimal value. To this end, one could use (6) to build a linear functional that dominates the dual gap. This piece of information constitutes the output of an oracle on which the algorithm must be built to solve the variational inequality.

In the next section, we describe the algorithm HACCPM that guarantees convergence in a number of iterations of the order of 1 over epsilon squared, where epsilon is precision achieved on the variational inequality. (That is the dual gap function at the epsilon-solution is somewhere between 0, the known optimal value, and minus epsilon.) This result is obtained under the mild assumption of monotonicity of F and a compact convex feasible set. The epsilon solution solves the so-called epsilon-weak solution of the variational inequality. As a by-product of the theoretical convergence result, one obtains that monotone variational inequalities always admit a weak solution. This property is not true for strong solutions which do not always exist under the simple monotonicity assumption. However, weak solutions are also strong solutions if F is continuous of if F is strongly convex.

Variational inequalities are reputedly hard to solve. They are significantly harder than general convex optimization problem, even though they can be formulated as the convex optimization problem of minimizing the dual gap function. The key point is that the dual gap function value is not known (computable), except at the optimum, which is 0. The remarkable theoretical fact is that one can build algorithms with a complexity bound of the same order as the best one for general convex programming.

4.2 The Homogeneous Analytic Center Cutting Plane Method

We now describe the method Homogeneous Analytic Center Cutting Plane Method (HACCPM) that solves

$$\text{Find } \bar{\mathbf{e}} \in \mathscr{E}^*_\varepsilon = \{\bar{\mathbf{e}} \in Q \mid \phi_D(\bar{\mathbf{e}}) \geq -\varepsilon\}.$$

Since ϕ_D is concave, $\mathscr{E}^*_\varepsilon$ is convex. If $\varepsilon = 0$, the set $\mathscr{E}^* = \mathscr{E}^*_0$ is the set of solutions to the weak variational inequality.

HACCPM [19] is a cutting plane method in a conic space with a polynomial bound on the number of iterations. To apply it to our problem of interest, we need to embed the original problem in an extended space. Let us describe the embedding first.

4.2.1 Embedding in an Extended Space

Let us introduce a projective variable[5] $t > 0$ and denote

$$K = \left\{ x = (y,t) \ | \ t > 0, \ \bar{\mathbf{e}} = \frac{y}{t} \in \mathscr{E} \right\}.$$

In this problem, $\mathscr{E}$ is linear and takes the general form

$$\mathscr{E} = \{\bar{\mathbf{e}} \ | \ \langle a_i, \bar{\mathbf{e}} \rangle \leq b_i. \ i = 1 \dots m\}.$$

Its conic version is

$$K_{\mathscr{E}} = \{x = (y,t) \ | \ \langle a_i, y \rangle \leq t b_i. \ i = 1 \dots m, \ t \geq 0\}.$$

We associate with it the logarithmic barrier function

$$B(x) = -\sum_{i=1}^{m} \ln(t b_i - \langle a_i, y \rangle) - \ln t.$$

B is a so-called ν-logarithmically homogeneous self-concordant function, with $\nu = m + 1$. (See definition 2.3.2 in the book by Nesterov and Nemirovski [18].) In this paper $K_{\mathscr{E}}$ is a simplex; thus $m = n + 1$, where n is the dimension of $\bar{\mathbf{e}}$.

The embedding of the valid inequality

$$\langle F(\bar{\mathbf{e}}), \bar{\mathbf{e}}^* - \bar{\mathbf{e}} \rangle \geq 0, \ \forall \bar{\mathbf{e}}^* \in \mathscr{E}^*$$

is done similarly. For any $x = (y,t) \in \operatorname{int} K_{\mathscr{E}}$, define $\bar{\mathbf{e}}(x) = y/t \in \mathscr{E}$ and

$$\hat{G}(x) = (F(\bar{\mathbf{e}}(x)), -\langle F(\bar{\mathbf{e}}(x)), \bar{\mathbf{e}}(x) \rangle).$$

It is easy to check that for $x^* \in X^*$ with

$$X^* = \{x = (y,t) \mid t > 0, \bar{\mathbf{e}}(x) \in \mathscr{E}^*\},$$

the following inequality holds

$$\langle \hat{G}(x), x - x^* \rangle \geq 0.$$

Finally, we define $G(x) = \hat{G}(x)/||\hat{G}(x)||$. We also associate to $\langle G(x), x - x^* \rangle > 0$ the logarithmic barrier $-\log\langle G(x), x - x^* \rangle$.

[5] The reader should not be confused by the use of t as a projective variable, whereas it was used as a time index in the game definition.

4.2.2 The Algorithm and Convergence Properties

The homogeneous cutting plane scheme, in its abstract form, can be briefly described as follows.

0) Set $B_0(x) = \frac{1}{2}||x||^2 + B(x)$.
1) kth iteration ($k \geq 0$).
   a) Compute $x_k = \arg\min_x B_k(x)$,
   b) Set $B_{k+1}(x) = B_k(x) - \ln\langle G(x_k), x_k - x\rangle$.

The algorithm is an abstract one as it assumes that the minimization of $B_k(x)$ in step (1-a) is carried out with full precision. This restriction has been removed in [20].

It is shown in [19] that $\bar{\mathbf{e}}(x^k)$ does not necessarily converge to $\mathscr{E}^*$. The correct candidate solution $\bar{\mathbf{e}}_k$ is as follows. Assume $\{x_i\}_{i=0}^{\infty}$ is a sequence generated by the algorithm. Define

$$\pi_{ik} = \frac{1}{||\hat{G}(x_i)||} \frac{1}{\langle G(x_i), x_i - x_k\rangle}, \quad P_k = \sum_{i=0}^{k-1} \pi_{ik},$$

and

$$\bar{\mathbf{e}}_k = \frac{1}{P_k} \sum_{i=0}^{k-1} \pi_{ik} \bar{\mathbf{e}}(x_i).$$

Assumption 1

1. *$\mathscr{E}$ is bounded and R is a constant such that for all $\bar{\mathbf{e}} \in \mathscr{E}$, $||\bar{\mathbf{e}}|| \leq R$.*
2. *The mapping $-F$ is uniformly bounded on $\mathscr{E}$ and is monotone, i.e., $||F(\bar{\mathbf{e}})|| \leq L$, for all $\bar{\mathbf{e}} \in \mathscr{E}$.*

The main convergence result is given by the following theorem.

Theorem 1. *HACCPM yields an ε-approximate solution after k iterations, with k satisfying*

$$\frac{k}{\sqrt{k+\nu}} \leq \frac{L(1+R^2)}{\varepsilon\theta_3} e^{\theta_2\sqrt{\nu}}. \tag{8}$$

The parameters in this formula are defined by

$$\begin{aligned} \nu &= n + 2 \\ \theta_1 &= (\sqrt{5} - 1)/2 - \log(\sqrt{5} + 1)/2 \\ \theta_2 &= (\sqrt{5} + 1)/2 \\ \theta_3 &= \tfrac{1}{\theta_2} \exp(\theta_1 - \tfrac{1}{2}). \end{aligned}$$

Note that $\theta_1, \theta_2, \theta_3$ are absolute constant, and ν is the dimension of the conic space plus 1.

To make this algorithm operational, one needs to work with an approximate minimizer of $B_k(x)$. More precisely, one can define a neighborhood of the exact minimizer with the following properties: i) checking whether the current iterate in the process of minimizing $B_k(x)$ belongs to this neighborhood is a direct byproduct of the computation; *ii*) the number of iterations to reach an approximate minimizer of $B_{k+1}(x)$ is bounded by a (small) absolute constant when the minimization starts from an approximate minimizer of $B_k(x)$. Finally, it is shown in [20] that the number of main iterations in the HACCPM algorithm with approximate centers, i.e., the number of generated cutting planes, is bounded by an expression similar to (8), but with slight different (absolute) values for the parameters.

In practice, the bound on the number of iterations is not used as a stopping criterion. Rather, we use the $-\varepsilon$ threshold for the primal gap ϕ_P.

5 Implementation

In the implementation, we use the method to compute the Nash equilibrium solution resulting from an allocation of a cumulative emission budget, as discussed in Sect. 2. We have used different rules of equity, based on population, Gross Domestic Product (GDP), grandfathering (i.e. historical emissions) to propose different splits of the cumulative emission budget among different groups of nations which then play a noncooperative game in the allocation of these emission allowances into quotas for each period.

During the optimization process, we compute an approximate gradient by calling the model $x+1$ times to obtain the sensitivity information (x corresponds to the number of the variables). Let $G(\bar{\mathbf{e}})$, a response of a GEMINI-E3 model to emission quotas $\bar{\mathbf{e}}$; an approximate evaluation of a pseudosubgradient at $\bar{\mathbf{e}}$ is given by $\frac{G(\bar{\mathbf{e}})-G(\bar{\mathbf{e}}+\Delta)}{\Delta}$, where Δ is an arbitrary perturbation on the emission quotas. This introduces another source of imprecision in the procedure. However, in practice, the approach shows convergence for most of the tested instances.

5.1 Case Study

We describe here the case study developed in the EU FP7 project TOCSIN.[6] The players are 4 regions of the World:

- NAM: North American countries;
- OEC: Other OECD countries;
- DCS: Developing countries (in particular Brazil, China, India);
- EEC: Oil and gas exporting countries (in particular Russia and Middle East).

[6] Technology-Oriented Cooperation and Strategies in India and China: Reinforcing the EU dialogue with Developing Countries on Climate Change Mitigation.

Table 1 Regional GHG emission in 2005

Region	GHG emission in 2005 (MtC-eq*)
NAM	2'491
OEC	2'187
DCS	4'061
EEC	1'067

*MtC-eq: millions tons of carbon equivalent

The economy is described by the GEMINI-E3 model during the years 2004–2050. The decision variables are the emissions of the 4 regions in each of the 4 periods, denoted $\bar{\mathbf{e}}_j(t)$ where $t = 2020, 2030, 2040, 2050$ and $j \in M = \{\text{NAM,OEC,DCS, EEC}\}$. Emissions in 2005 are known and exogenous (see Table 1). The yearly emissions are interpolated linearly within each period.

The coupled constraint on global emission is then

$$8 \times \sum_{j \in M} e_j(2005) + \sum_{j \in M} [13, 10, 10, 5] \begin{bmatrix} \bar{\mathbf{e}}_j(2020) \\ \bar{\mathbf{e}}_j(2030) \\ \bar{\mathbf{e}}_j(2040) \\ \bar{\mathbf{e}}_j(2050) \end{bmatrix} \leq 519\,\mathrm{GtC} - \mathrm{eq}$$

Notice that here the notations have changed a little, since $\bar{\mathbf{e}}_j(t)$ refers to the yearly emission quotas of country j at the beginning of period t. The coefficients in the row matrix [13, 10, 10, 5] serve to define the total emissions per period, using linear interpolation. The 519 GtC-eq amount is the cumulative emission budget for an emission path that is compatible with a global warming of 2°C in 2100 for an average climate sensitivity of 3.5. This value (519 GtC-eq) has been obtained from simulation performed with a bottom-up model (TIAM [17]) which includes a climate module. This threshold has been established as critical by the European Union and the consequence of an overtaking will be irreversible for our ecosystem [24].

5.2 Results

We have tested six different splits of the emission budget called scenario 1 to scenario 6 that are presented in Table 2 . Due to the lengthy economic simulations performed by GEMINI-E3 at each call, the resolution time of the problem is important [7] and the computation of subgradients has been parallelized to reduce the oracle call time.

[7] We used a Dual 2.6 GHz Intel Xeon computer for the simulations, thus we had four available CPUs.

Table 2 Percent split of cumulative emission budget per scenario

Scenario	NAM	OEC	DCS	EEC
1	27	19	40	14
2	22	17	48	13
3	19	14	54	13
4	18	13	56	13
5	14	11	60	15
6	15	14	60	11

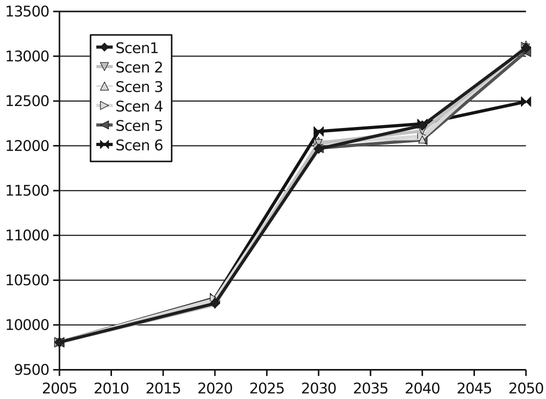


Fig. 2 World GHG emission in MtC-eq

One first interesting result concerns the World GHG emission, as it is shown in Fig. 2 the trajectory does not vary much with the rule adopted to allocate the cumulative emission budget. In Fig. 3, we display the equilibrium quotas for the different splits of a cumulative emission budget amounting to 519 GtC-eq. Of course, the equilibrium quotas at each period depend on the total allowances given to the players; however, we can note that industrialized countries (NAM and OEC) and energy exporting countries tend to allocate more quotas to the first periods of the Game. In contrary, developing countries tend to allocate their quotas to the last periods of the Game. There is a clear dichotomy between DCS and the Rest of the World which could be due to the weight given to this region in the split of the cumulative emissions budget. In all the scenarios tested, DCS represents at least 40% of the World quota and is therefore a central player.

The resulting payoffs, expressed in welfare variation in comparison with the BAU situation, are shown in Fig. 4. The welfare of the region depends on the emission budget initially given (see Table 2). Scenarios 3 and 4 are the most acceptable rules, because in all regions the welfare loss is limited to 0.5% of the household consumption. In contrary, in scenarios 1 and 6 the nonindustrialized countries bear an important welfare loss, respectively, DCS and EEC. Scenario 5 is the most likely acceptable scenario in the context of the post-Kyoto negotiation, because it leads to

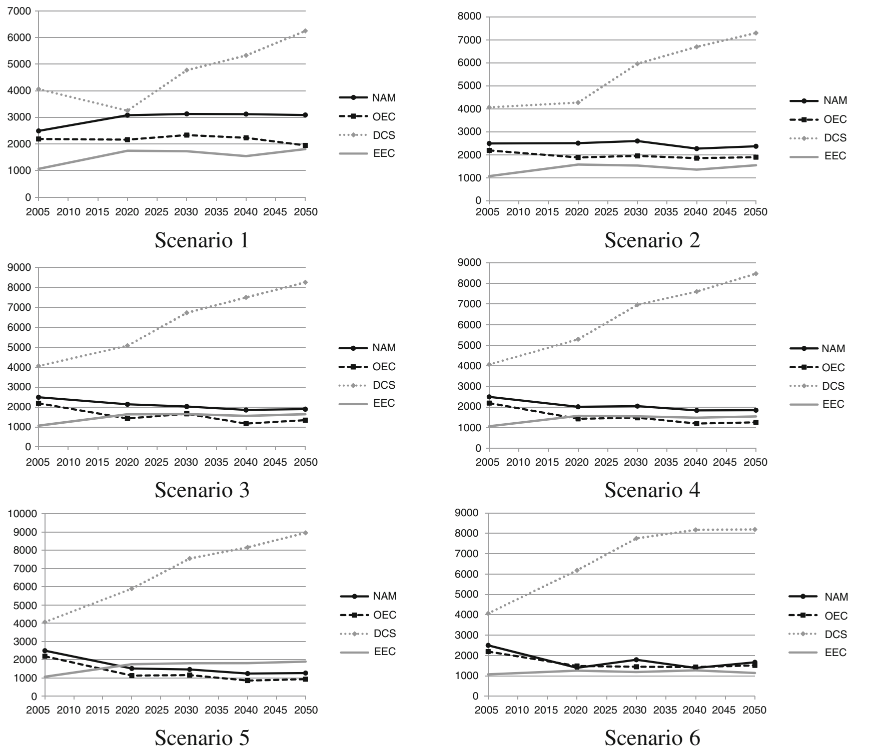


Fig. 3 Equilibrium quotas for each scenario (MtC-eq)

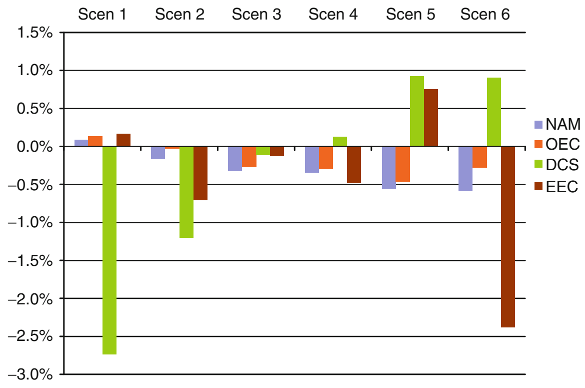


Fig. 4 Equilibrium payoffs for each scenario – sum of actualized surplus in % of households consumption

a welfare gain coming from the selling of permits by DCS and EEC. In the Kyoto Protocol, these two regions were effectively always reluctant to join the coalition of countries which are committed in constrained target of GHG.

6 Conclusion

In this paper, we have shown how an oracle-based optimization method could be implemented to compute Nash equilibrium solutions or Rosen normalized equilibrium solutions in a game where the payoffs are obtained from a large-scale macro-economic simulation model. This approach permits the use of game theoretic concepts in the analysis of economic and climate policies through the use of detailed models of economic general equilibrium. We have shown that the (always difficult) interpretation of the weights in the Rosen normalized equilibrium concept could be simplified in the case of a scalar and separable coupled constraint. In our case, the splitting of the cumulative emission budget over the planning period would be equivalent to a particular weighting in a normalized equilibrium.. We have applied this approach to a realistic description of the World economy and obtained a set of simulations showing the possible "fair" burden sharing that could emerge from negotiation for the post-2012 climate policies.

Acknowledgements Research supported by the EU-FP6 TOCSIN project and the French GICC research program. The authors thank two anonymous referees whose remarks and comments have permitted them to improve substantially this paper.

References

1. Armington, P.S.: A theory of demand for products distinguished by place of production. IMF Staff Papers **16**, 159–178 (1969)
2. Bernard, A., Haurie, A., Vielle, M., Viguier, L.: A two-level dynamic game of carbon emission trading between Russia, China, and Annex B countries. Journal of Economic Dynamics and Control **32**(6), 1830–1856 (2008)
3. Bernard, A., Vielle, M.: Comment allouer un coût global d'environnement entre pays : permis négociables versus taxes ou permis négociables et taxes? Économie Internationale **82**, 103–135 (2000)
4. Bernard, A., Vielle, M.: Measuring the welfare cost of climate change policies: A comparative assessment based on the computable general equilibrium model GEMINI-E3. Environmental Modeling & Assessment **8**(3), 199–217 (2003)
5. Bernard, A., Vielle, M.: GEMINI-E3, a General Equilibrium Model of International National Interactions between Economy, Energy and the Environment. Computational Management Science **5**(3), 173–206 (2008)
6. Bernard, A., Vielle, M., Viguier, L.: Carbon tax and international emissions trading: A swiss perspective. In Haurie, A., Viguier, L. (eds.) Coupling Climate and Economic Dynamics. Springer (2005)
7. Dimaranan, B.V.: Global Trade, Assistance, and Production: The GTAP 6 Data Base. Center for Global Trade Analysis Purdue University, Center for Global Trade Analysis, Purdue University (2006)
8. Drouet, L., Haurie, A., Moresino, F., Vial, J.-P., Vielle, M., Viguier, L..: An oracle based method to compute a coupled equilibrium in a model of international climate policy. Computational Management Science **5**(1), 119–140 (2008)
9. Ferris, M.C., Munson, T.S.: Complementarity problems in GAMS and the PATH solver. Journal of Economic Dynamics and Control **24**, 165–188 (2000)

10. Ferris, M.C., Pang, J.S.: Complementarity and Variational Problems: State of the Art. IAM Publications, Philadelphia (1997)
11. Haurie, A.: Environmental coordination in dynamic oligopolistic markets. Group Decision and Negotiation **4**, 46–67 (1995)
12. Haurie, A., Krawczyk, J.: Optimal charges on river effluent from lumped and distributed sources. Environmental Modeling and Assessment **2**, 177–199 (1997)
13. Haurie, A., Smeers, Y., Zaccour, G.: Stochastic equilibrium programming for dynamic oligopolistic markets. Journal of Optimmization Theory and Applications **66**(2) (1990)
14. Haurie, A., Zaccour, G.: Game-theoretic models of the environment. In: Annals of the International Society of Dynamic Games, volume 3, chapter Differential game models of global environmental management, pages 3–23. Birkhäuser, Boston (1995)
15. Krawczyk, J.B.: Coupled constraint Nash equilibria in environmental games. Resource and Energy Economics **27**(2), 157–181 (2005)
16. Krawczyk, J.B.: Numerical solutions to coupled-constraint (or generalised Nash) equilibrium problems. Computational Management Science **2**(1), 183–204 (2007)
17. Loulou, R.: ETSAP-TIAM: the TIMES integrated assessment model. part ii: mathematical formulation. Computational Management Science **5**(3), 41–66 (2008)
18. Nesterov, Y., Nemirovsky, A.: Interior Point Polynomial Methods in Convex Programming. SIAM, Philadelphia (1994)
19. Nesterov, Y., Vial, J.-P.: Homogeneous Analytic Center Cutting Plane Methods for convex problems and variational inequalities. SIAM J. Optim. **9**(3), 707–728 (1999)
20. Nesterov, Y., Péton, O., Vial, J.-P.: Homogeneous Analytic Center Cutting Plane Methods with approximate centers. Optimization Methods and Software **11&12**, 243–273 (1999)
21. Pang, J.S., Fukushima, M.: Quasi-variational inequalities, generalized Nash equilibria and multi-leader-follower games. Computational Management Science **2**(1), 21–56 (2005)
22. Péton, O. The homogeneous analytic center cutting plane method. Ph.D. Thesis, Université de Genève (2002)
23. Rosen, J.B.: Existence and uniqueness of equilibrium points for concave n-person games. Econometrica **33**(3), 520–534 (1965)
24. Stern, N.: Stern Review: The Economics of Climate Change. HM Treasury UK, London (2006)
25. Stone, J.R.N.: Linear expenditure systems and demand analysis: An application to the pattern of british demand. Economic Journal (1983)
26. von Heusinger, A., Kanzow, C.: SC optimization reformulations of the generalized Nash equilibrium problem. Optimization Methods and Software **23**, 953–973 (2008)
27. von Heusinger, A., Kanzow, C.: Optimization reformulations of the generalized Nash equilibrium problem using Nikaido-Isoda-type functions. Computational Optimization and Applications **43**, 353–377 (2009)

Optimal Stopping of a Risk Process with Disruption and Interest Rates

Elżbieta Z. Ferenstein and Adam Pasternak-Winiarski

Abstract It is a standard approach in classical risk theory to assume a claim process which does not change throughout the whole observation period. Most commonly, encountered models deal with compound Poisson processes. It would be beneficial to investigate more general classes of claim processes with arbitrary distributions of random variables governing inter-occurrence times between losses and loss severities. Further generalization of such framework would be a model allowing for disruptions i.e. changes of such distributions according to some unobservable random variables, representing fluctuating environmental conditions. The question of providing the company with tools allowing for detection of such change and maximizing the returns leads to an optimal stopping problem which we solve explicitly to some extent. Moreover, we provide references to previously examined models as well as numerical examples emphasizing the efficiency of the suggested method.

1 Introduction

The following model has been often investigated in collective risk theory. An insurance company with a given initial capital u_0 receives premiums, which flow at a constant rate $c > 0$, and has to pay for claims which occur according to some point process at times $T_1 < T_2 < \dots$, $\lim T_n = \infty$. The risk process $(U_t)_{t \in \mathrm{R}_+}$ is defined as the difference between the income and the total amount of claims up to time t.

Many articles have been investigating the issue of maximizing company's gains in such models, leading to interesting risk process optimal stopping problems. An illuminating example of such research is an article by Jensen [7], where the method of finding an optimal stopping time maximizing the expected net gain $\mathbb{E}(U_{\tau^*}) = \sup\{\mathbb{E}U_\tau : \tau \in \mathscr{C}\}$, with $\mathscr{C}$ being a class of feasible stopping times, is provided.

E.Z. Ferenstein (✉) and A. Pasternak-Winiarski
Warsaw University of Technology, Faculty of Mathematics and Information Science,
Pl. Politechniki 1, 00-661 Warszawa, Poland
e-mail: efer@mini.pw.edu.pl; A.Pasternak-Winiarski@mini.pw.edu.pl

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_24,

Moreover, through the application of a model based on Markov-modulated Poisson claim processes, the author dealt with the situation of deteriorating market conditions. The solution was then extended in [11] onto more economically relevant models, allowing for the reinvestment of the surplus.

The reasoning applied in the articles mentioned above is based on the smooth semimartingale representation of the risk process and fails when some utility function of the risk process has to be considered. Hence, based on the methodology applied in [1, 2], Ferenstein and Sierociński [5] solved the problem of maximizing the expected utility of a risk process by applying a dynamic programming methodology and proposed an effective method for determining the optimal stopping times in such situations. Muciek [9] extended this model and adapted it to insurance practice. He investigated the optimal stopping times under the assumption that claims increase at some given rate, and that the capital of the company can be invested.

However, both models proved to be still quite restrictive as they enforced throughout the entire period of observation only one distribution for the random variable describing interoccurrence times between losses, and for the random variable describing amounts of subsequent losses. Although Karpowicz & Szajowski [8] presented a solution of the problem of double optimal stopping, their model still did not take into consideration the possibility of disruption. Allowing for such disruption could make the stopping rule more interesting in terms of insurance practice, based on the observed realization of the process and not only on arbitrary management decisions. Hence, Pasternak-Winiarski [10] introduced a model in which the distributions changed according to some unobservable random variable and solved the corresponding optimal stopping problem for finite number of claims.

The main motivation for the research described in this article is to extend the findings from [10] by solving the problem for an infinite number of claims and by introducing a simple market environment, as in [9]. As a result, the stopping rules are more versatile and interesting, for instance in terms of insurance practice.

2 The Model and the Optimal Stopping Problem

Let $(\Omega, \mathscr{F}, P)$ be a probability space. On this space we introduce the following random variables and processes:

(i) Random variable Θ with values in $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and having a geometric distribution with parameters $p, \pi_0 \in [0, 1]$:

$$P(\Theta = 0) = \pi_0,$$

$$P(\Theta = n) = (1 - \pi_0)p(1 - p)^{n-1}, \qquad n \in \mathbb{N}.$$

Θ is unobservable and triggers the disruption. This approach is not a typical one in change-point process literature, as Θ – instead of describing the real time – provides us only with information about the number of first claims with a changed distribution.

(ii) Claim counting process $N_t, t \in \mathbb{R}_+$, with jumps at times $0 < T_1 < T_2 < \ldots, \lim_{n\to\infty} T_n = \infty$, which form a nonexplosive point process $N_t = \sum_{i=1}^{\infty} \mathbb{I}_{\{T_i \le t\}}$.

(iii) A sequence of random variables $S_n = T_n - T_{n-1}, n = 1, 2, \ldots, T_0 = 0$. S_n represents the interoccurrence time between the $n-1$th and the nth loss. S_n depends on the unobservable random variable Θ (as in the disorder problem considered in [12]) and is defined as follows:

$$S_n = W_n' \mathbb{I}_{\{n \le \Theta\}} + W_n'' \mathbb{I}_{\{n > \Theta\}}.$$

$W_n', n \in \mathbb{N}$, is a sequence of i.i.d. random variables with cumulative distribution function (c.d.f.) F_1 (satisfying the condition $F_1(0) = 0$) and density function f_1. Similarly $W_n'', n \in \mathbb{N}$, forms a sequence of i.i.d. random variables with c.d.f. F_2 ($F_2(0) = 0$) and density function f_2. We assume additionally that f_1 and f_2 are commonly bounded by a constant $C \in \mathbb{R}_+$. Furthermore, we impose that W_i' and W_j'' are independent for all $i, j \in \mathbb{N}_0$.

(iv) A sequence $X_n, n \in \mathbb{N}_0$, of random variables representing successive losses. They also depend on the random variable Θ:

$$X_n = X_n' \mathbb{I}_{\{n < \Theta\}} + X_n'' \mathbb{I}_{\{n \ge \Theta\}},$$

where $X_n', n \in \mathbb{N}_0$, is a sequence of i.i.d. random variables with c.d.f. H_1 ($H_1(0) = 0$) and density function h_1 whereas $X_n'', n \in \mathbb{N}_0$, forms a sequence of i.i.d. random variables with c.d.f. H_2 ($H_2(0) = 0$) and density function h_2. X_i' and X_j'' are independent for all $i, j \in \mathbb{N}_0$.

We assume that random variables $W_n', W_n'', X_n', X_n'', \Theta$ are independent.

Let $u_0 > 0$ represent the initial capital and $c > 0$ be a constant rate of income from the insurance premium. We take into account the dynamics of the market situation by introducing the interest rate at which we can invest accrued capital (constant $\alpha \in [0, 1]$). We assume that the claims increase at rate $\beta \in [0, 1]$ as a consequence of, say, inflation.

As a capital assets model for the insurance company, we take the risk process

$$U_t = u_0 e^{\alpha t} + \int_0^t c e^{\alpha(t-s)} ds - \sum_{i=0}^{N_t} X_i e^{\beta T_i}, \qquad X_0 = 0.$$

The return at time t is defined by the process

$$Z_t = g_1(U_t) \mathbb{I}_{\{U_s > 0, s < t\}} \mathbb{I}_{\{t < t_0\}},$$

where g_1 is a utility function and the constant t_0 is a fixed time which denotes the end of the investment period. For simplicity, we define

$$g(u, t) = g_1(u) \mathbb{I}_{\{t \ge 0\}}.$$

We then have

$$Z_t = g(U_t, t_0 - t) \prod_{i=1}^{N_t} \mathbb{I}_{\{U_{T_i} > 0\}}. \tag{1}$$

We fix a number $K \in \mathbb{N}$. We will need the following family of $\sigma-$fields generated by all events up to time $t > 0$:

$$\mathscr{F}_t^U \triangleq \sigma(U_s, s \leq t).$$

Applying the arguments analogical to the ones used in [3], one can easily show that

$$\mathscr{F}_{T_n}^U = \sigma(X_1, T_1, \ldots, X_n, T_n).$$

Hence, let us now denote

$$\mathscr{F}_n \triangleq \mathscr{F}_{T_n}^U \text{ and } \mathscr{G}_n \triangleq \mathscr{F}_n \vee \sigma(\Theta).$$

In our calculations, we will extensively make use of conditional probabilities $\pi_n \triangleq P(\Theta \leq n | \mathscr{F}_n), n \in \mathbb{N}_0$, as well as $\widehat{\pi}_n \triangleq P(\Theta = n+1 | \mathscr{F}_n), n \in \mathbb{N}_0$.

We will now define the optimization problem, which will be solved in the subsequent sections. Let $\mathscr{T}$ be the set of all stopping times with respect to the family $\{\mathscr{F}_t^U\}_{t>0}$. For $n = 0, 1, 2, \ldots k < K$ we denote by $\mathscr{T}_{n,K}$ such subset of $\mathscr{T}$ that satisfies the condition

$$\tau \in \mathscr{T}_{n,K} \Leftrightarrow T_n \leq \tau \leq T_K \text{ a.s.}$$

We will be seeking the optimal stopping time τ_K^* such that

$$\mathbb{E}(Z_{\tau_K^*}) = \sup\{\mathbb{E}(Z_\tau) : \tau \in \mathscr{T}_{0,K}\}.$$

In order to find τ_K^* we first consider optimal stopping times $\tau_{n,K}^*$ such that

$$\mathbb{E}(Z_{\tau_{n,K}^*} | \mathscr{F}_n) = \text{ ess sup } \{\mathbb{E}(Z_\tau | \mathscr{F}_n) : \tau \in \mathscr{T}_{n,K}\}.$$

Then using the standard methods of dynamic programming we will obtain $\tau_K^* = \tau_{0,K}^*$.

After finding τ_K^* for fixed K we concentrate on solving the optimal stopping problem in the situation where an unlimited number of claims is attainable, that is we find τ^* such that

$$\mathbb{E}(Z_{\tau^*}) = \sup \{\mathbb{E}(Z_\tau) : \tau \in \mathscr{T}\}.$$

τ^* will be defined as a limit of finite horizon stopping times τ_K^*.

To clarify in more details the structure of the stopping rule derived in this paper, we illustrate it briefly below. First, we find a special set of functions $R_i^* (\cdot, \cdot, \cdot)$,

$i = 0, \dots, K$. Then, at time $T_0 = 0$, with $U_0 = u_0$, we calculate $R_0^*(U_0, T_0, \pi_0)$. If at time $T_0 + R_0^*(U_0, T_0, \pi_0)$, a first claim has not yet been observed, we stop. Otherwise, when the first claim occurs at time $T_1 < T_0 + R_0^*(U_0, T_0, \pi_0)$, we calculate the value $R_1^*(U_1, T_1, \pi_1)$ and wait for the next claim till the time $T_1 + R_1^*(U_1, T_1, \pi_1)$, etc. In other words, the optimal stopping times derived in this model can be interpreted as constituting a threshold rule.

3 Solution of the Finite Horizon Problem

In this section, we will find the form of the optimal stopping rule in the finite horizon case, i.e. optimal in the class $\mathscr{T}_{n,K}$, where K – representing the maximal number of claims – is finite and fixed. First, in Theorem 1, we will derive dynamic programming equations satisfied by

$$\Gamma_{n,K} = \operatorname{ess\,sup} \{\mathbb{E}(Z_\tau | \mathscr{F}_n) : \tau \in \mathscr{T}_{n,K}\}.$$

Then, in Theorem 2, we will find optimal stopping times $\tau_{n,K}^*$ and τ_K^* and corresponding optimal conditional mean rewards and optimal mean rewards, respectively.

For notational simplicity we define

$$\mu_t = \prod_{i=1}^{N_t} \mathbb{I}_{\{U_{T_i} > 0\}}, \qquad \mu_0 = 1. \tag{2}$$

A simple consequence of these notations and formula (1) is that

$$\Gamma_{K,K} = Z_{T_K} = \mu_{T_K} g(U_{T_K}, t_0 - T_K). \tag{3}$$

A crucial role in the subsequent reasoning is played by Lemmas 1 and 2 given below. Lemma 1 defines the recursive relation between conditional probabilities π_n as well as $\widehat{\pi}_n$, essential in our further considerations. Lemma 2 is a representation theorem for stopping times (to be found in [3, 4]).

Lemma 1. *There exist functions* $\xi_1 : [0, 1] \times (0, \infty) \times (0, \infty) \to [0, 1]$, $\xi_2 : [0, 1] \to [0, 1]$ *such that*

$$\pi_n = \xi_1(\pi_{n-1}, X_n, S_n), \; \widehat{\pi}_n = \xi_2(\pi_n),$$

and

$$\xi_1(t, x, s) = \frac{f_2(s)h_2(x)t + pf_1(s)h_2(x)(1-t)}{f_2(s)h_2(x)t + pf_1(s)h_2(x)(1-t) + f_1(s)h_1(x)(1-p)(1-t)},$$

$$\xi_2(t) = \frac{\eta_1(t) + \eta_3(t)t - (1 + \eta_1(t))t}{1 + \eta_1(t) - \eta_2(t)},$$

where

$$\eta_1(t) = \int_0^\infty \int_0^\infty \xi_1(t,x,w)\mathrm{d}H_1(x)\mathrm{d}F_1(w),$$

$$\eta_2(t) = \int_0^\infty \int_0^\infty \xi_1(t,x,w)\mathrm{d}H_2(x)\mathrm{d}F_1(w),$$

$$\eta_3(t) = \int_0^\infty \int_0^\infty \xi_1(t,x,w)\mathrm{d}H_2(x)\mathrm{d}F_2(w).$$

Proof. The nature of the proof is purely technical so it will be omitted here. The reasoning is based upon the Bayes formula. Similar recursive relation is presented in [12], Part 4, Theorem 7. □

Lemma 2. *If* $\tau \in \mathscr{T}_{n,K}$, *then there exists a positive* $\mathscr{F}_n$ *- measurable random variable* R_n *such that* $\min(\tau, T_{n+1}) = \min(T_n + R_n, T_{n+1})$.

To simplify further calculations, we define following auxiliary functions

$$d : \mathbb{R}_+^2 \to \mathbb{R}_+ \qquad \text{by} \qquad d(t,r) := \left(\frac{c}{\alpha} + u_0\right)\left(\mathrm{e}^{\alpha(t+r)} - \mathrm{e}^{\alpha t}\right),$$

$$D : \mathbb{R}_+^3 \to \mathbb{R} \qquad \text{by} \qquad D(t,r,x) := d(t,r) - x\mathrm{e}^{\beta(t+r)} \qquad \text{and}$$

$$\hat{D} : \mathbb{R}_+^2 \times \mathbb{R} \to \mathbb{R} \qquad \text{by} \qquad \hat{D}(t,r,u) := \mathrm{e}^{-\beta(t+r)}(u + d(t,r)).$$

Theorem 1.

(a) *For* $n = K-1, K-2, \ldots, 0$ *we have*

$$\begin{aligned}\Gamma_{n,K} = \ & ess\ sup\{\mu_{T_n} g(U_{T_n} + d(T_n, R_n)), t_0 - T_n - R_n)(\overline{F_2}(R_n)\pi_n \\ & +\overline{F_1}(R_n)(1-\pi_n)) + \mathbb{E}(\mathbb{I}_{\{R_n \geq S_{n+1}\}}\Gamma_{n+1,K}|\mathscr{F}_n) : \ R_n \geq 0 \\ & \text{and } R_n \text{ is } \mathscr{F}_n\text{-measurable}\},\end{aligned}$$

where $\overline{F} = 1 - F$

(b) *For* $n = K, K-1, \ldots, 0$ *we have*

$$\Gamma_{n,K} = \mu_{T_n}\gamma_{K-n}(U_{T_n}, T_n, \pi_n) \text{ a.s. }, \tag{4}$$

where the sequence of functions $\{\gamma_j : \mathbb{R} \times [0,\infty) \times [0,1] \to \mathbb{R}\}_{j=0,\ldots,K}$ *is defined recursively as follows:*

$$\gamma_0(u,t,\pi) = g(u, t_0 - t),$$

$$\gamma_j(u,t,\pi) = \sup_{r\geq 0}\{g(u+d(t,r),t_0-t-r)(\overline{F_2}(r)\pi + \overline{F_1}(r)(1-\pi))$$

$$+\xi_2(\pi)\int_0^r \int_0^{\hat{D}(t,w,u)} \gamma_{j-1}(u+D(t,w,x),t+w,\xi_1(\pi,x,w))\mathrm{d}H_2(x)\mathrm{d}F_1(w)$$

$$+(1-\pi-\xi_2(\pi))\int_0^r \int_0^{\hat{D}(t,w,u)} \gamma_{j-1}(u+D(t,w,x),t+w,\xi_1(\pi,x,w))\mathrm{d}H_1(x)\mathrm{d}F_1(w)$$

$$+\pi\int_0^r \int_0^{\hat{D}(t,w,u)} \gamma_{j-1}(u+D(t,w,x),t+w,\xi_1(\pi,x,w))\mathrm{d}H_2(x)\mathrm{d}F_2(w)\},$$

where ξ_1, ξ_2 are the functions defined in Lemma 1.

Proof. (a) Let $\tau \in \mathscr{T}_{n,K}$ and $0 \leq n < K < \infty$. Lemma 2 implies that

$$\begin{aligned} A_n :&= \{\tau < T_{n+1}\} = \{T_n + R_n < T_{n+1}\} = \{R_n < S_{n+1}\} \\ &= \left(\left\{R_n < W'_{n+1}\right\} \cap \{\Theta > n\}\right) \cup \left(\left\{R_n < W''_{n+1}\right\} \cap \{\Theta \leq n\}\right) = A_n^1 \cup A_n^2. \end{aligned}$$

Thus,

$$\mathbb{E}(Z_\tau|\mathscr{F}_n) = \mathbb{E}(Z_\tau \mathbb{I}_{A_n^1}|\mathscr{F}_n) + \mathbb{E}(Z_\tau \mathbb{I}_{A_n^2}|\mathscr{F}_n) + \mathbb{E}(Z_\tau \mathbb{I}_{\overline{A_n}}|\mathscr{F}_n) = a_n^1 + a_n^2 + b_n. \tag{5}$$

We will now calculate a_n^1. First, we transform the given form of Z_τ using (1) and (2)

$$a_n^1 = \mu_{T_n}\mathbb{E}(\mathbb{I}_{\left\{R_n < W'_{n+1}\right\}}\mathbb{I}_{\{\Theta > n\}} g(U_\tau, t_0-\tau)|\mathscr{F}_n).$$

As $R_n < W'_{n+1} = S_{n+1}$, it is obvious that from T_n until τ no loss had been observed. Therefore, as $U_{t+R_n} - U_t = d(t, R_n)$ we can rewrite U_τ and get

$$\begin{aligned} a_n^1 &= \mu_{T_n}\mathbb{E}(\mathbb{I}_{\left\{R_n < W'_{n+1}\right\}}\mathbb{I}_{\{n<\Theta\}} g(U_{T_n} + d(T_n, R_n), t_0 - T_n - R_n)|\mathscr{F}_n) \\ &= \mu_{T_n} g(U_{T_n} + d(T_n, R_n), t_0 - T_n - R_n)\mathbb{E}(\mathbb{I}_{\left\{R_n < W'_{n+1}\right\}})\mathbb{E}(\mathbb{I}_{\{\Theta > n\}}|\mathscr{F}_n) \\ &= \mu_{T_n} g(U_{T_n} + d(T_n, R_n), t_0 - T_n - R_n)\overline{F_1}(R_n)(1-\pi_n). \end{aligned}$$

Similarly, one can show that

$$a_n^2 = \mu_{T_n} g(U_{T_n} + d(T_n, R_n), t_0 - T_n - R_n)\overline{F_2}(R_n)\pi_n.$$

If we additionally define $\tau' := \max(\tau, T_{n+1})$, then it is easy to see that $\tau' \in \mathscr{T}_{n+1,K}$ and:

$$b_n = \mathbb{E}(\mathbb{E}(Z_{\tau'}\mathbb{I}_{\{S_{n+1}\le R_n\}}|\mathscr{F}_{n+1})|\mathscr{F}_n) = \mathbb{E}(\mathbb{I}_{\{S_{n+1}\le R_n\}}\mathbb{E}(Z_{\tau'}|\mathscr{F}_{n+1})|\mathscr{F}_n). \quad (6)$$

The formulas (5)-(6) imply that

$$\begin{aligned}&\mathbb{E}(Z_\tau|\mathscr{F}_n)\\ &\quad = \mu_{T_n} g(U_{T_n} + d(T_n, R_n), t_0 - T_n - R_n)(\overline{F_2}(R_n)\pi_n + \overline{F_1}(R_n)(1-\pi_n))\\ &\qquad + \mathbb{E}(\mathbb{I}_{\{S_{n+1}\le R_n\}}\mathbb{E}(Z_{\tau'}|\mathscr{F}_{n+1})|\mathscr{F}_n).\end{aligned}$$

Now, following the standard reasoning of the optimal stopping theory, we get the dynamic programming equation for $\Gamma_{n,K}, n = K, K-1, \ldots, 0$, given in (4), with $\Gamma_{K,K} = \mu_{T_K} g(U_{T_K}, t_0 - T_K)$.

(b) We will prove (b) using the backward induction method. First, one should note that (4) is satisfied for $n = K$, as (3) gives

$$\Gamma_{K,K} = \mu_{T_K} g(U_{T_K}, t_0 - T_K) = \mu_{T_K}\gamma_0(U_{T_K}, T_K, \pi_K). \quad (7)$$

Let $n = K-1$. It is easy to observe that

$$\begin{aligned}\{R_{K-1} \ge S_K\} &= \left(\left\{R_{K-1} \ge W'_K\right\} \cap \{\Theta \le K\}\right) \cup \left(\left\{R_{K-1} \ge W''_K\right\} \cap \{\Theta > K\}\right)\\ &= \left(\left\{R_{K-1} \ge W'_K\right\} \cap \{\Theta = K\}\right) \cup \left(\left\{R_{K-1} \ge W''_K\right\} \cap (\{\Theta > K\})\right)\\ &\qquad \cup \left(\left\{R_{K-1} \ge W'_K\right\} \cap \{\Theta \le K-1\}\right) := B^1_K \cup B^2_K \cup B^3_K.\end{aligned}$$

The above equality implies

$$\begin{aligned}\Gamma_{K-1,K} = \mu_{T_{K-1}} \operatorname{ess\,sup}\Big\{& g(U_{T_{K-1}} + d(T_{K-1}, R_{K-1}), t_0 - T_{K-1} - R_{K-1})\\ &\overline{F_2}(R_{K-1})\pi_{K-1} + g(U_{T_{K-1}} + d(T_{K-1}, R_{K-1}), t_0 - T_{K-1} - R_{K-1})\\ &\overline{F_1}(R_{K-1})(1-\pi_{K-1}) + \mathbb{E}\left(\left(\mathbb{I}_{B^1_K} + \mathbb{I}_{B^2_K} + \mathbb{I}_{B^3_K}\right)\Gamma_{K,K}|\mathscr{F}_{K-1}\right):\\ &R_{K-1} \ge 0 \text{ and } R_K \text{ is } \mathscr{F}_K\text{-measurable}\Big\}.\end{aligned}$$

We present the calculations only for the set B^1_K - the remaining summands under the conditional expectation can be transformed in a similar way. Taking (7) into consideration, rewriting μ_{T_K} as $\mu_{T_{K-1}}\mathbb{I}_{\{U_{T_{K-1}}+d(T_{K-1},S_K)-X_K e^{\beta(T_{K-1}+S_K)}>0\}}$ and applying the definitions of random variables S_K, X_K and the process U_{T_K} we get that

$$\begin{aligned}\mathbb{E}\left(\mathbb{I}_{B^1_K}\Gamma_{K,K}|\mathscr{F}_{K-1}\right) = \mu_{T_{K-1}} &\times \mathbb{E}(\mathbb{I}_{B^1_K}\mathbb{I}_{\{U_{T_{K-1}}+D(T_{K-1},W'_K,X''_K)>0\}}\\ &g(U_{T_{K-1}}+D(T_{K-1}, W'_K, X''_K), t_0 - T_{K-1} - W'_K)|\mathscr{F}_{K-1}).\end{aligned}$$

The independence of random variables W'_K and X''_K from $\mathscr{F}_{K-1}$ along with the definition of conditional probability $\widehat{\pi}_K$ implies that

$$\mathbb{E}\left(\mathbb{I}_{B^1_K}\Gamma_{K,K}|\mathscr{F}_{K-1}\right) = \mu_{T_{K-1}}\xi_2(\pi_{K-1}) \times \int_0^{R_{K-1}} \int_0^{\hat{D}(T_{K-1},w,U_{T_{K-1}})} g(U_{T_{K-1}} + D(T_{K-1}, w, x), t_0 - T_{K-1} - w)\mathrm{d}H_2(x)\mathrm{d}F_1(w).$$

Analogous calculations for B^2_K and B^3_K complete the backward induction step for $n = K-1$.

Let $1 \le n < K-1$ and suppose that $\Gamma_{n,K} = \mu_{T_n}\gamma_{K-n}(U_{T_n}, T_n, \pi_n)$.
From (a) we have

$$\begin{aligned}\Gamma_{n-1,K} = \operatorname{ess\,sup}\Big\{&\mu_{T_{n-1}} g(U_{T_{n-1}} + d(T_{n-1}, R_{n-1}), t_0 - T_{n-1} - R_{n-1})\\ &\overline{F_2}(R_{n-1})\pi_{n-1} + \mu_{T_{n-1}} g(U_{T_{n-1}} + d(T_{n-1}, R_{n-1}), t_0 - T_{n-1} - R_{n-1})\\ &\overline{F_1}(R_{n-1})(1-\pi_{n-1}) + \mathbb{E}\left(\left(\mathbb{I}_{B^1_n\cup B^2_n\cup B^3_n}\right)\mu_{T_n}\gamma_{K-n}(U_{T_n}, T_n, \pi_n)|\mathscr{F}_{n-1}\right):\\ &R_{n-1} \ge 0,\ R_{n-1} \text{ is } \mathscr{F}_{n-1}\text{-measurable}\Big\}.\end{aligned} \tag{8}$$

Since $\mu_{T_n} = \mu_{T_{n-1}}\mathbb{I}_{\{U_{T_{n-1}}+D(T_{n-1},S_n,X_n)>0\}}$, similarly to the calculations presented above we derive the formula for conditional expectation from (8) related to $\mathbb{I}_{B^1_n}$. We get

$$\begin{aligned}&\mathbb{E}\left(\mathbb{I}_{B^1_n}\mu_{T_n}\gamma_{K-n}(U_{T_n}, T_n, \pi_n)|\mathscr{F}_{n-1}\right)\\ &= \mu_{T_{n-1}}\mathbb{E}\left(\mathbb{I}_{B^1_n}\mathbb{I}_{\{U_{T_{n-1}}+D(T_{n-1},W'_n,X''_n)>0\}}\gamma_{K-n}(U_{T_n}, T_n, \pi_n)|\mathscr{F}_{n-1}\right)\\ &= \mu_{T_{n-1}}\mathbb{E}(\mathbb{I}_{\{\Theta=n\}}\mathbb{E}(\mathbb{I}_{\{R_{n-1}\ge W'_n\}}\mathbb{I}_{\{U_{T_{n-1}}+D(T_{n-1},W'_n,X''_n)>0\}}\\ &\quad\times\gamma_{K-n}(U_{T_{n-1}} + D(T_{n-1}, W'_n, X''_n), T_{n-1}\\ &\quad+W'_n, \xi_1(\pi_{n-1}, X''_n, W'_n))|\mathscr{G}_{n-1})|\mathscr{F}_{n-1}) = \xi_2(\pi_{n-1})\\ &\quad\times\int_0^{R_{n-1}}\int_0^{\hat{D}(T_{n-1},w,U_{T_{n-1}})}\gamma_{K-n}(U_{T_{n-1}} + D(T_{n-1}, w, x), T_{n-1}\\ &\quad+w, \xi_1(\pi_{n-1}, x, w))\mathrm{d}H_2(x)\mathrm{d}F_1(w).\end{aligned}$$

Analogous calculations for B^2_n and B^3_n complete the proof of the theorem. □

We will now concentrate on the problem of finding the optimal stopping time τ^*_K. To this end, as it was proved in [5], we have to analyze properties of the sequence of functions γ_n.

Let $B = B[(-\infty, \infty) \times [0, \infty) \times [0, 1]]$ be the space of all bounded continuous functions on $(-\infty, \infty) \times [0, \infty) \times [0, 1]$, $B^0 = \{\delta : \delta(u, t, \pi) = \delta_1(u, t, \pi)\mathbb{I}_{\{t \le t_0\}}, \delta_1 \in B\}$. On B^0 we define a norm:

$$\| \delta \|_\alpha = \sup_{u, 0 \le t \le t_0, \pi} \left\{ \left(\frac{t}{t_0} \right)^\alpha \mid \delta(u, t, \pi) \mid \right\},$$

where $\alpha > 1$ is an arbitrary constant, such that $\chi = \frac{Ct_0}{\alpha - 1} \in (0, 1)$ (the properties of similar norms were considered in [6]).

For any $\delta \in B^0$, $u \in \mathbb{R}$, $t, r \ge 0$ and $\pi \in [0, 1]$ we define:

$$\begin{aligned}
\phi_\delta(r, u, t, \pi) &= g(u + d(t, r), t_0 - t - r)(\overline{F_2}(r)\pi + \overline{F_1}(r)(1 - \pi)) \\
&+ \xi_2(\pi) \int_0^r \int_0^{\hat{D}(t,w,u)} \delta(u + D(t, w, x), t + w, \xi_1(\pi, x, w))\mathrm{d}H_2(x)\mathrm{d}F_1(w) \\
&+ (1 - \pi - \xi_2(\pi)) \int_0^r \int_0^{\hat{D}(t,w,u)} \delta(u + D(t, w, x), t + w, \xi_1(\pi, x, w))\mathrm{d}H_1(x)\mathrm{d}F_1(w) \\
&+ \pi \int_0^r \int_0^{\hat{D}(t,w,u)} \delta(u + D(t, w, x), t + w, \xi_1(\pi, x, w))\mathrm{d}H_2(x)\mathrm{d}F_2(w).
\end{aligned}$$

We will now make use of the properties of the c.d.f.'s F_1 and F_2. They imply that if g_1 is continuous and $t \ne t_0 - r$, then $\phi_\delta(r, u, t, \pi)$ is continuous with respect to (r, u, t, π). Therefore, we make the following

Assumption 1. *The function $g_1(\cdot)$ is bounded and continuous.*

For any $\delta \in B^0$ we define an operator Φ as follows:

$$(\Phi\delta)(u, t, \pi) = \sup_{r \ge 0}\{\phi_\delta(r, u, t, \pi)\}. \tag{9}$$

Lemma 3. *For $\pi \in [0, 1]$ and for every $\delta \in B^0$ we have:*

$$(\Phi\delta)(u, t, \pi) = \max_{0 \le r \le t_0 - t}\{\phi_\delta(r, u, t, \pi)\}.$$

Furthermore, there exists a function r_δ satisfying $(\Phi\delta)(u, t, \pi) = \phi_\delta(r_\delta(u, t, \pi), u, t, \pi)$.

Proof. When $r > t_0 - t$ and $\delta \in B^0$ then $g(u + cr, t_0 - t - r) = 0$ and the equality (9) can be rewritten in the following way

$$\phi_\delta(r, u, t, \pi) = \xi_2(\pi) \int_0^{t_0 - t} \int_0^{\hat{D}(t,w,u)} \delta(u + D(t, w, x), t + w, \xi_1(\pi, x, w))\mathrm{d}H_2(x)\mathrm{d}F_1(w)$$

$$+(1-\pi-\xi_2(\pi))\int_0^{t_0-t}\int_0^{\hat{D}(t,w,u)}\delta(u+D(t,w,x),t+w,\xi_1(\pi,x,w))\mathrm{d}H_1(x)\mathrm{d}F_1(w)$$

$$+\pi\int_0^{t_0-t}\int_0^{\hat{D}(t,w,u)}\delta(u+D(t,w,x),t+w,\xi_1(\pi,x,w))\mathrm{d}H_2(x)\mathrm{d}F_2(w).$$

Hence, Assumption 1 and the fact that F_1 and F_2 are continuous functions in the compact interval $[0,t_0]$ imply the form of Φ. □

It is easy to note that for $i=0,1,2,\ldots,K-1$, $u\in\mathbb{R}$, $t\geq 0$, $\pi\in[0,1]$ the sequence $\gamma_i(u,t,\pi)$ can be rewritten according to the following pattern

$$\gamma_{i+1}(u,t,\pi)=\begin{cases}(\Phi\gamma_i)(u,t,\pi) & \text{if } u\geq 0,\, t\leq t_0,\\ 0 & \text{otherwise.}\end{cases}$$

From Lemma 3 we know that there exist functions $r_{K-1-i}:=r_{\gamma_i}$, such that

$$\gamma_{i+1}(u,t,\pi)=\begin{cases}\phi_{\gamma_i}(r_{K-1-i}(u,t,\pi),u,t,\pi) & \text{if } u\geq 0,\, t\leq t_0,\\ 0 & \text{otherwise.}\end{cases}$$

To determine the form of optimal stopping times $\tau^*_{n,K}$ we define following random variables:

$$R^*_i=r_i(U_{T_i},T_i,\pi_i)$$

and

$$\sigma_{n,K}=\min\{K,\inf\{i\geq n: R^*_i<S_{i+1}\}\}.$$

Theorem 2. *Let*

$$\tau^*_{n,K}=\begin{cases}T_{\sigma_{n,K}}+R^*_{\sigma_{n,K}} & \text{if } \sigma_{n,K}<K,\\ T_K & \text{if } \sigma_{n,K}=K,\end{cases}\qquad \text{and}\qquad \tau^*_K=\tau^*_{0,K}.$$

Then for any $0\leq n\leq K$ *we have*

$$\Gamma_{n,K}=\mathbb{E}(Z_{\tau^*_{n,K}}|\mathscr{F}_n)\ a.s.\qquad \text{and}\qquad \Gamma_{0,K}=\mathbb{E}(Z_{\tau^*_K}).$$

Proof. This is a straightforward consequence of the definition of random variables R^*_i, $\sigma_{n,K}$ and Theorem 1. □

4 Solution of the Infinite Horizon Problem

Lemma 4. *The operator $\Phi : B^0 \to B^0$ defined by (9) is a contraction in the norm $\|\cdot\|_\alpha$*

Proof. Let $\delta_1, \delta_2 \in B^0$. Then by Lemma 3 there exist $\varrho_i := r_{\delta_i}(u,t,\pi) \le t_0 - t$, $i = 1,2$, such that $(\Phi\delta_i) = \phi_{\delta_i}(\varrho_i, u, t, \pi)$. It is obvious that $\phi_{\delta_2}(\varrho_2, u, t, \pi) \ge \phi_{\delta_2}(\varrho_1, u, t, \pi)$. Hence, applying an auxiliary notation

$$\overline{\delta}_i(u,t,w,x,\pi) := \delta_i(u + D(t,w,x), t + w, \xi_1(\pi,x,w)), \qquad i = 1,2,$$

we get

$$\begin{aligned}
&(\Phi\delta_1)(u,t,\pi) - (\Phi\delta_2)(u,t,\pi) \le \phi_{\delta_1}(\varrho_1,u,t,\pi) - \phi_{\delta_2}(\varrho_1,u,t,\pi)\\
&\quad = \xi_2(\pi)\int_0^{\varrho_1}\int_0^{\hat{D}(t,w,u)} (\overline{\delta}_1 - \overline{\delta}_2)(u,t,w,x,\pi)\left(\frac{t+w}{t_0}\right)^\alpha\left(\frac{t_0}{t+w}\right)^\alpha \mathrm{d}H_2(x)\mathrm{d}F_1(w)\\
&\qquad + (1-\pi-\xi_2(\pi))\int_0^{\varrho_1}\int_0^{\hat{D}(t,w,u)} (\overline{\delta}_1 - \overline{\delta}_2)(u,t,w,x,\pi)\left(\frac{t+w}{t_0}\right)^\alpha\left(\frac{t_0}{t+w}\right)^\alpha\\
&\qquad\quad \times \mathrm{d}H_1(x)\mathrm{d}F_1(w)\\
&\qquad + \pi\int_0^{\varrho_1}\int_0^{\hat{D}(t,w,u)} (\overline{\delta}_1 - \overline{\delta}_2)(u,t,w,x,\pi)\left(\frac{t+w}{t_0}\right)^\alpha\left(\frac{t_0}{t+w}\right)^\alpha \mathrm{d}H_2(x)\mathrm{d}F_2(w)\\
&\quad \le (t_0)^\alpha \xi_2(\pi)\int_0^{t_0-t} \|\delta_1 - \delta_2\|_\alpha \left(\frac{1}{t+w}\right)^\alpha \mathrm{d}F_1(w)\\
&\qquad + (t_0)^\alpha(1-\pi-\xi_2(\pi))\int_0^{t_0-t} \|\delta_1 - \delta_2\|_\alpha \left(\frac{1}{t+w}\right)^\alpha \mathrm{d}F_1(w)\\
&\qquad + (t_0)^\alpha \pi\int_0^{t_0-t} \|\delta_1 - \delta_2\|_\alpha \left(\frac{1}{t+w}\right)^\alpha \mathrm{d}F_2(w).
\end{aligned}$$

An analogous estimation can be carried through for $(\Phi\delta_2)(u,t,\pi) - (\Phi\delta_1)(u,t,\pi)$. Since we have assumed that the density functions f_1 and f_2 are commonly bounded by a constant C we get:

$$\begin{aligned}
&|(\Phi\delta_2)(u,t,\pi) - (\Phi\delta_1)(u,t,\pi)|\\
&\quad \le (t_0)^\alpha C\|\delta_1 - \delta_2\|_\alpha \int_0^{t_0-t}\left(\frac{1}{t+w}\right)^\alpha dw < (t_0)^\alpha \frac{C}{\alpha - 1}||\delta_1 - \delta_2||_\alpha\left(\frac{t_0}{t^\alpha}\right).
\end{aligned}$$

As a straightforward consequence

$$\|(\Phi\delta_2)(u,t,\pi) - (\Phi\delta_1)(u,t,\pi)\|_\alpha < \frac{Ct_0}{\alpha - 1}||\delta_1 - \delta_2||_\alpha \leq \chi||\delta_1 - \delta_2||_\alpha,$$

where $\chi < 1$. □

Applying Banach's Fixed Point Theorem, we get Lemma 5:

Lemma 5. *There exists $\gamma \in B^0$ such that*

$$\gamma = \Phi\gamma \qquad \textit{and} \qquad \lim_{K\to\infty} ||\gamma_K - \gamma||_\alpha = 0.$$

Lemma 5 will turn out to be useful in the crucial part of the proof of Theorem 3, describing the optimal stopping rule in the case of infinite horizon.

Theorem 3. *Assume that the utility function g_1 is differentiable and nondecreasing and F_i have commonly bounded density functions f_i for $i = 1, 2$. Then:*

(a) for $n = 0, 1, \ldots$ the limit $\hat{\tau}_n := \lim_{K\to\infty} \tau^*_{n,K}$ exists and $\hat{\tau}_n$ is an optimal stopping rule in $\mathscr{T} \cap \{\tau \geq T_n\}$.
(b) $\mathbb{E}(Z_{\hat{\tau}_n}|\mathscr{F}_n) = \mu_{T_n}\gamma(U_{T_n}, T_n, \pi_n)$ a.s.

Proof. (a) It is obvious that $\tau^*_{n,K} \leq \tau^*_{n,K+1}$ a.s. for $n \geq 0$. Hence, the stopping rule $\hat{\tau}_n$: $T_n \leq \hat{\tau}_n = \lim_{K\to\infty} \tau^*_{n,K} \leq t_0$ exists. We only have to prove the optimality of $\hat{\tau}_n$, to which aim we will apply arguments similar to those used in [1,2], and [5].

Let $\xi_t = (t, U_t, Y_t, V_t, \pi_{N_t}, N_t)$, where $Y_t = t - T_{N_t}$ and $V_t = \mu_t = \prod_{i=1}^{N_t} \mathbb{I}_{\{U_{T_i}>0\}}$, $t \geq 0$. Then, one can show that $\xi = \{\xi_t : t \geq 0\}$ is a Markov process with the state space $\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+ \times \{0,1\} \times [0,1] \times \mathbb{N}_0$. One can see that the return Z_t can be described as a function, say $\tilde{g}$, of ξ_t. Then, we can calculate a strong generator of ξ in the form

$$(A\tilde{g})(t,u,y,v,\pi,n)$$

$$= v\left\{\pi_n\left(e^{\alpha t}(u_0\alpha + c)g'(u) - \frac{f_2(y)}{\overline{F_2}(y)}\left(g_1(u) - \int_0^{ue^{-\beta t}} g_1(u - e^{\beta t}x)\mathrm{d}H_2(x)\right)\right)\right.$$

$$\left.+(1-\pi_n)\left(e^{\alpha t}(u_0\alpha + c)g'(u) - \frac{f_1(y)}{\overline{F_1}(y)}\left(g_1(u) - \int_0^{ue^{-\beta t}} g_1(u - e^{\beta t}x)\mathrm{d}H_1(x)\right)\right)\right.$$

$$+ \xi_2(\pi_n)\frac{f_1(y)}{\overline{F_1}(y)}\left(\int\limits_0^{u\mathrm{e}^{-\beta t}} g_1(u-\mathrm{e}^{\beta t}x)\mathrm{d}H_2(x) - \int\limits_0^{u\mathrm{e}^{-\beta t}} g_1(u-\mathrm{e}^{\beta t}x)\mathrm{d}H_1(x)\right)\Bigg\}, \tag{10}$$

where the expression above is well defined as we can assume $f_i(\tilde{t}_0^i) = 0$ for $\tilde{t}_0^i := \sup\limits_{t \le t_0}\{F_i(t) < 1\}$.

Obviously $M_t := \tilde{g}(\xi_t) - \tilde{g}(0) - \int\limits_0^t (A\tilde{g})(\xi_s)\mathrm{d}s$, $t \ge 0$ is a martingale with respect to the filtration $\sigma(\xi_s, s \le t)$, which is the same as $\mathscr{F}_t^U$. As T_n and $\tau^*_{n,K}$ are stopping times satisfying the condition $T_n \le \tau^*_{n,K}, a.s.$, and $\mathscr{F}_n = \mathscr{F}^U_{T_n}$ we can apply optional sampling theorem and get

$$E(M_{\tau^*_{n,K}}|\mathscr{F}_n) = M_{T_n} \text{ a.s.},$$

$$\mathbb{E}\left(\tilde{g}(\xi_{\tau^*_{n,K}}) - \tilde{g}(0) - \int\limits_0^{\tau^*_{n,K}} (A\tilde{g})(\xi_s)\mathrm{d}s|\mathscr{F}_n\right)$$
$$= \tilde{g}(\xi_{T_n}) - \tilde{g}(0) - \int\limits_0^{T_n}(A\tilde{g})(\xi_s)\mathrm{d}s \text{ a.s.},$$

and finally

$$\mathbb{E}\left(\tilde{g}(\xi_{\tau^*_{n,K}})|\mathscr{F}_n\right) - \tilde{g}(\xi_{T_n}) = \mathbb{E}\left(\int\limits_{T_n}^{\tau^*_{n,K}} (A\tilde{g})(\xi_s)\mathrm{d}s|\mathscr{F}_n\right) \text{ a.s.} \tag{11}$$

We will now calculate the limit of the expression from the right-hand side of the equality (11) with $K \to \infty$. First, applying the form of the generator from (10) and denoting

$$\mathscr{I}_i := \int\limits_0^{U_s\mathrm{e}^{-\beta s}} g_1(U_s - \mathrm{e}^{\beta s}x)\mathrm{d}H_i(x), \qquad i = 1, 2$$

we get

$$(A\tilde{g})(\xi_s) = \mu_s\Bigg\{\pi_{N_s}\left(\mathrm{e}^{\alpha s}(u_0\alpha + c)g'(U_s) + \frac{f_2(s-T_{N_s})}{\overline{F_2}(s-T_{N_s})}\left(\mathscr{I}_2 - g_1(U_s)\right)\right)$$
$$+(1-\pi_{N_s})\left(\mathrm{e}^{\alpha s}(u_0\alpha + c)g'(U_s) + \frac{f_1(s-T_{N_s})}{\overline{F_1}(s-T_{N_s})}\left(\mathscr{I}_1 - g_1(U_s)\right)\right)$$
$$+ \xi_2(\pi_{N_s})\frac{f_1(s-T_{N_s})}{\overline{F_1}(s-T_{N_s})}\left(\mathscr{I}_2 - \mathscr{I}_1\right)\Bigg\}. \tag{12}$$

Inserting in (11) the formula for infinitesimal generator given in (12) we get

$$\mathbb{E}\left(\tilde{g}(\xi_{\tau^*_{n,K}})|\mathscr{F}_n\right) - \tilde{g}(\xi_{T_n}) = \mathbb{E}\left(J^1_{n,K}|\mathscr{F}_n\right) - \mathbb{E}\left(J^2_{n,K}|\mathscr{F}_n\right) \text{ a.s.,}$$

where

$$J^1_{n,K} = \int_{T_n}^{\tau^*_{n,K}} \left(\left(\mathrm{e}^{\alpha s}(u_0\alpha + c)g'(U_s) + \left(\frac{f_2(s-T_{N_s})}{\overline{F_2}(s-T_{N_s})} + \frac{\xi_2(\pi_{N_s})}{\pi_{N_s}} \frac{f_1(s-T_{N_s})}{\overline{F_1}(s-T_{N_s})} \right) \mathscr{I}_2 \right) \pi_{N_s} \right.$$
$$\left. + \left(\mathrm{e}^{\alpha s}(u_0\alpha + c)g'(U_s) + \frac{f_1(s-T_{N_s})}{\overline{F_1}(s-T_{N_s})} \left(1 - \frac{\xi_2(\pi_{N_s})}{1-\pi_{N_s}}\right) \mathscr{I}_1 \right) (1-\pi_{N_s}) \right) \mu_s \mathrm{d}s,$$

$$J^2_{n,K} = \int_{T_n}^{\tau^*_{n,K}} \left(\pi_{N_s} \frac{f_2(s-T_{N_s})}{\overline{F_2}(s-T_{N_s})} + (1-\pi_{N_s}) \frac{f_1(s-T_{N_s})}{\overline{F_1}(s-T_{N_s})} \right) g_1(U_s)\mu_s \mathrm{d}s.$$

Note that $J^2_{n,K}$ is a nonnegative random variable. As $\pi_{N_s} + \xi_2(\pi_{N_s}) \leq 1$, $J^1_{n,K}$ is also a nonnegative random variable. Moreover,

$$\begin{aligned} J^2_{n,K} &\leq \int_{T_n}^{\tau^*_{n,K}} \left(\pi_{N_s} \frac{f_2(s-T_{N_s})}{\overline{F_2}(s-T_{N_s})} + (1-\pi_{N_s}) \frac{f_1(s-T_{N_s})}{\overline{F_1}(s-T_{N_s})} \right) g_1(u + d(0,t_0))\mathrm{d}s \\ &\leq g_1(u+d(0,t_0)) \left(\int_{T_n}^{t_0} \sum_{k=n}^{\infty} \left(\frac{f_2(s-T_{N_s})}{\overline{F_2}(s-T_{N_s})} + \frac{f_1(s-T_{N_s})}{\overline{F_1}(s-T_{N_s})} \right) \mathbb{I}_{\{N_s=k\}} \mathrm{d}s \right) \\ &\leq g_1(u+d(0,t_0)) \sum_{k=1}^{N_{t_0}} \int_0^{S_k} \left(\frac{f_2(y)}{\overline{F_2}(y)} + \frac{f_1(y)}{\overline{F_1}(y)} \right) \mathrm{d}y \end{aligned}$$

Applying the Monotone Convergence Theorem, the properties of conditional expectation and the fact that $\lim_{K\to\infty} \tau^*_{n,K} = \hat{\tau}_n$ we get

$$\lim_{K\to\infty} \mathbb{E}\left(\int_{T_n}^{\tau^*_{n,K}} (A\tilde{g})(\xi_s)\mathrm{d}s|\mathscr{F}_n \right) = \mathbb{E}\left(\int_{T_n}^{\hat{\tau}_n} (A\tilde{g})(\xi_s)\mathrm{d}s|\mathscr{F}_n \right) \text{ a.s.}$$

On the other hand, Dynkin formula also implies that

$$\mathbb{E}\left(\int_{T_n}^{\hat{\tau}_n} (A\tilde{g})(\xi_s)\mathrm{d}s|\mathscr{F}_n \right) = \mathbb{E}\left(\tilde{g}(\xi_{\hat{\tau}_n})|\mathscr{F}_n)\right) - \tilde{g}(\xi_{T_n}) \text{ a.s.}$$

Hence,

$$\lim_{K\to\infty} \mathbb{E}\left(\tilde{g}(\xi_{\tau^*_{n,K}})|\mathscr{F}_n\right) = \mathbb{E}\left(\tilde{g}(\xi_{\hat{\tau}_n})|\mathscr{F}_n\right) \text{ a.s.} \tag{13}$$

To complete this part of the proof, we only have to show that $\hat{\tau}_n$ is an optimal stopping time in the class $\mathscr{T} \cap \{\tau : \tau \geq T_n\}$. To that end let us assume that τ is some other stopping rule in $\mathscr{T} \cap \{\tau : \tau \geq T_n\}$. Then, as $\tau^*_{n,K}$ is optimal in $\mathscr{T}_{n,K}$, we have for any K,

$$\mathbb{E}\left(\tilde{g}(\xi_{\tau^*_{n,K}})|\mathscr{F}_n\right) \geq \mathbb{E}\left(\tilde{g}(\xi_{\tau\wedge T_K})|\mathscr{F}_n\right) \text{ a.s.}$$

A similar argument to the one leading to formula (13) reveals that

$$\mathbb{E}\left(\tilde{g}(\xi_{\hat{\tau}_n})|\mathscr{F}_n\right) \geq \mathbb{E}\left(\tilde{g}(\xi_{\tau})|\mathscr{F}_n\right) \text{ a.s.}$$

The proof of (a) is now complete.
(b) As it was shown in part (b) of Theorem 1, the following equality stands

$$\mathbb{E}\left(\tilde{g}(\xi_{\tau^*_{n,K}})|\mathscr{F}_n\right) = \mu_{T_n}\gamma_{K-n}(U_{T_n}, T_n, \pi_n).$$

Then Lemma 5 and (13) imply

$$\begin{aligned}\mathbb{E}\left(Z_{\hat{\tau}_n}|\mathscr{F}_n\right) &= \lim_{K\to\infty} \mathbb{E}\left(\tilde{g}(\xi_{\tau^*_{n,K}})|\mathscr{F}_n\right)\\ &= \lim_{K\to\infty} \mu_{T_n}\gamma_{K-n}(U_{T_n}, T_n, \pi_n) = \mu_{T_n}\gamma(U_{T_n}, T_n, \pi_n) \text{ a.s.}\end{aligned}$$

which completes the proof of the theorem. □

5 Numerical Examples

We now propose an example to provide the reader with a complete overview of the stopping rules in the suggested model. For the sake of simplicity, to focus the attention of the reader on the method itself rather than on strenuous calculations, we will assume that the rates of inflation α as well as claim severity growth β are equal to zero. The extension of the example below to a model with nonzero rates is straightforward.

Let us assume that the observation period is equal to 1 year ($t_0 = 1$), $c = 1$ is the constant rate of income from the insurance premium and the fixed number K of claims that may occur in our model is 1. Moreover, let us assume that the probability p defining the distribution of Θ is equal to $\frac{1}{2}$.

We impose that W_1' and W_1'' are uniformly distributed over the interval $[0, t_0]$. The claim severity distribution changes from $H_1(x) = \frac{e^x-1}{4}\mathbb{I}_{\{x\in[0,ln5]\}}$ for X_1' to $H_2(x) = \frac{e^x-1}{8}\mathbb{I}_{\{x\in[0,ln9]\}}$ for X_1''. As $\mathbb{E}X' < \mathbb{E}X''$ it is obvious that the example reflects a situation of deteriorating market conditions. Three models will be investigated:

Model 1. In this scenario, the insurance company does not apply any optimal stopping rule. Hence, the observation ends along with first claim or when the time t_0 is reached, whichever happens first.

Model 2. In this scenario, the company applies the optimal stopping rule suggested in the articles [5] and [9]. Hence, having no possibility to anticipate potential distribution changes the company assumes that $S_1 = W_1'$ and $X_1 = X_1'$. In such case

$$\phi_\delta(r,u,t) = \mathrm{e}^{u+r}(1-r)\mathbb{I}_{\{t+r<1\}} + \frac{1}{4}\int_0^r \int_0^{u+s} \delta(u+s-x,t+s)\mathrm{e}^x \mathrm{d}x\mathrm{d}s.$$

As $K = 1$ we execute only one step of iterative procedure and we get that

$$\Phi_{\gamma_0}(u,t) = \max_{0\le r\le 1}\left\{\mathrm{e}^{u+r}(1-r)\mathbb{I}_{\{t+r<1\}} + \frac{1}{4}\int_0^r \mathrm{e}^{u+s}(u+s)\mathbb{I}_{\{t+s<1\}}\mathrm{d}s\right\}.$$

Standard calculations reveal that in such case

$$\tau^*_{0,1} = R^*_0\mathbb{I}_{\{R^*_0<T_1\}} + T_1\mathbb{I}_{\{R^*_0\ge T_1\}}, \text{ where } R^*_0 = \min\left\{\frac{1}{3}u_0, 1\right\}.$$

Model 3. In the third scenario, the company can use the suggested model with disruptions. It is easy to see that under the assumptions stated above we have

$$\begin{aligned}
\xi_1(\pi,x,s) &= \frac{\pi+1}{3-\pi}, \qquad \xi_2(\pi) = \frac{(1-\pi)^2}{3-\pi},\\
\phi_\delta(r,u,t,\pi) &= \mathrm{e}^{u+r}(1-r)\mathbb{I}_{\{t+r<1\}}\\
&\quad + \frac{5-3\pi}{24-8\pi}\int_0^r\int_0^{u+s}\delta\left(u+s-x,t+s,\frac{\pi+1}{3-\pi}\right)\mathrm{e}^x\mathrm{d}x\mathrm{d}s,\\
\Phi_{\gamma_0}(u,t,\pi) &= \max_{0\le r\le 1}\Big\{\mathrm{e}^{u+r}(1-r)\mathbb{I}_{\{t+r<1\}}\\
&\quad + \frac{5-3\pi}{24-8\pi}\int_0^r \mathrm{e}^{u+s}(u+s)\mathbb{I}_{\{t+s<1\}}\mathrm{d}s\Big\}.
\end{aligned}$$

It is not difficult to see that

$$\tau^*_{0,1} = R^*_0\mathbb{I}_{\{R^*_0<T_1\}} + T_1\mathbb{I}_{\{R^*_0\ge T_1\}}, \text{ where } R^*_0 = \min\left\{\frac{5-3\pi_0}{19-5\pi_0}u_0, 1\right\}.$$

The results of simulations for different values of initial capital u and probability π_0 are presented below. We record the average returns for the company over 10,000 trajectories generated in each scenario (Tables 1 and 2).

It is easy to note that the scenarios where the optimal stopping rules have been applied offer significantly higher returns regardless of the initial probability π_0. The model with disruption is better in the situations when the probability of distribution

Table 1 Returns for $\pi_0 = 0.5$ and $p = 0.5$

Initial capital	Model 1	Model 2	Model 3
0,2	0,311	1,222	1,222
0,4	0,474	1,496	1,497
0,6	0,687	1,831	1,836
0,8	0,978	2,252	2,258
1	1,323	2,759	2,777
1,2	1,741	3,408	3,437
1,4	2,248	4,176	4,243
1,6	2,808	5,055	5,219
1,8	3,486	6,041	6,387
2	4,295	7,151	7,817
2,2	5,26	8,335	9,51
2,4	6,41	9,484	11,483
2,6	7,832	10,595	13,83
2,8	9,539	11,353	16,622
3	11,645	11,645	19,858
3,2	14,254	14,254	23,596
3,4	17,431	17,431	27,857
3,6	21,212	21,212	32,588
3,8	26,084	26,084	38,373
4,2	38,81	38,81	50,297
4,6	57,844	57,844	62
4,8	71,016	71,016	71,016

Table 2 Returns for $u_0 = 1$ and $p = 0.5$

π_0	Model 1	Model 2	Model 3
0,04	1,699	2,842	2,837
0,12	1,619	2,82	2,819
0,2	1,561	2,806	2,809
0,28	1,506	2,801	2,806
0,36	1,434	2,783	2,792
0,44	1,374	2,777	2,79
0,52	1,306	2,764	2,785
0,6	1,238	2,754	2,77
0,68	1,178	2,732	2,762
0,76	1,119	2,723	2,764
0,84	1,054	2,712	2,751
0,92	0,988	2,694	2,746
1	0,919	2,687	2,742

change at $t = 0$ is higher than 0.2, whereas it somewhat fails when π_0 is low. However, this can easily be justified. As Model 3 is bound to deal with more complex market situations and covers a wider spectrum of processes, the simplicity of Model 2 prevails in the environment for which it was originally designed (as low probability of disruption in the first step in a model with only one claim observed in fact reflects the environment with no disruption). However, we believe that with greater values of K this effect would not be observed.

The analysis of company's returns for different values of initial capital confirms the considerable advantage of the models with optimal stopping rules over "passive" models. As the stopping moment in Model 2 and Model 3 is defined by a minimum of observation period and an increasing function of initial capital u_0, it is obvious that from some initial capital forth ($u_0 = 3$ in case of Model 2 and $u_0 = 4\frac{5}{7}$ in case of Model 3) all the models will give the same returns. This fact underlines the advantages of optimal stopping models especially in the situation when company's initial capital is low and the probability of becoming insolvent is substantial.

6 Final Remarks

A diligent reader will notice the fact that this work does not only introduce a new model and solve the problem of optimal stopping of the risk process within this scheme. It also widens the spectrum of the processes for which we can apply stopping rules derived in the models from [5] and [9]. Authors assumed there that distribution of the random variable describing interoccurrence times between losses, F, fulfills the condition $F(t_0) < 1$. As it was shown in this article, such assumption is irrelevant, provided that $F'(t)$ is bounded. A simple change of norm allows to apply the Fixed Point Theorem.

References

1. Boshuizen, F.A., Gouweleeuw, J.M.: General Optimal Stopping Theorems for Semi-Markov Processes. Adv. Appl. Probab. **25**, 825–846 (1993)
2. Boshuizen, F.A., Gouweleeuw, J.M.: A Continuous-time Job Search Model: General Renewal Process. Commun. Stat. Stoch. Model **11**, 349–369 (1995)
3. Brémaud, P.: Point Processes and Queues. Martingale Dynamics. Springer (1995)
4. Davis M.H.A.: Markov Models and Optimization. Chappman and Hall (1993)
5. Ferenstein, E.Z., Sierocinski, A.: Optimal Stopping of a Risk Process. Appl. Math. **24**, 335–342 (1997)
6. Goebel, K., Kirk, W.A.: Topics in Metric Fixed Point Theory. Cambridge University Press (1990)
7. Jensen, U.: An Optimal Stopping Problem in Risk Theory. Scand. Actuar. J. **1997**, 149–159 (1997)
8. Karpowicz, A., Szajowski, K.: Double Optimal Stopping of a Risk Process. Int. J. Probab. Stoch. Process. **79**, 155–167 (2007)
9. Muciek, B.K.: Optimal Stopping of a Risk Process: Model with Interest Rates. J. Appl. Probab. **39**, 261–270 (2002)
10. Pasternak-Winiarski, A.: Optimal Stopping of a Risk Process. Master Thesis (in Polish) (2007)
11. Schöttl, A.: Optimal Stopping of a Risk Reserve Process with Interest and Cost Rates. J. Appl. Probab. **35**, 115–123 (1998)
12. Shiryayev, A.: Optimal Stopping Rules. Springer (1978)

Examples in Dynamic Optimal Taxation

Mikhail Krastanov and Rossen Rozenov

Abstract One famous result in the theory of capital income taxation is that the optimal tax is zero in equilibrium (Chamley, Econometrica 54(3):607–622, 1986; Judd, J. Public Econ. 28:59–83, 1985) . This result has been derived as an open-loop Stackelberg solution to an appropriate differential game. In this paper, we consider specific feedback solutions to three dynamic models of taxation and find that the optimal tax is generally different from zero.

1 Introduction

Dynamic game theory provides a natural setup for the analysis of optimal taxation problems as it allows to model explicitly the strategic interaction between the government and the taxpayers. Taxes affect economic behavior and the reaction of firms and individuals to the decrease in their disposable incomes should be taken into account when designing the tax system. Chamley [1] and Judd [5] were among the first to explore the implications of capital income taxation on welfare in a dynamic framework. In these studies, as well as in most later works, the problem of choosing an optimal tax policy rule is stated formally as a Stackelberg game with the government acting as a leader. The solutions to this game are typically sought in the open-loop strategy space. This, however, raises the important question of the time consistency of optimal policy since open-loop Stackelberg equilibria are generally known to be time inconsistent [2]. Feedback Stackelberg solutions, on the other hand, are time consistent but they are much more difficult to calculate [3].

M. Krastanov (✉)
Institute of Mathematics and Informatics, Bulgarian Academy of Sciences,
Acad. G. Bonchev Str. Bl.8, 1113 Sofia, Bulgaria
e-mail: krast@math.bas.bg

R. Rozenov
Bulgarian National Bank, 1 King Alexander Sq., 1000 Sofia, Bulgaria
e-mail: rossen_rozenov@yahoo.com

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_25,

Alternatively, the solution of the differential game could be obtained as a feedback Nash equilibrium. It has been shown recently by Rubio [8] (assuming differentiability of the value function) that for certain two-player games the stationary feedback Stackelberg and Nash equilibria coincide. A sufficient condition for that is the following: the mixed partial derivatives of the instantaneous utility functions and the right-hand side of the differential equation with respect to the controls of the two players must be zero. Games, for which this condition holds true, are said to have "orthogonal reaction functions" [8]. The coincidence of equilibria result is important since for many differential games with hierarchical structure and orthogonal reaction functions, finding the feedback Nash equilibrium is equivalent to finding the respective feedback Stackelberg equilibrium. The one-dimensional taxation models presented below are examples of such differential games, provided that the respective value functions are C^1.

In this paper, we consider three models of dynamic optimal taxation and calculate explicitly their feedback Nash equilibria for some special cases. The derivation of equilibria for these models is based on the solution of an optimal control problem (OCP) with a specific structure. The solution to this problem is obtained using sufficient conditions for optimality (see [9]). The main finding is that whenever the government and the taxpayers employ state feedback strategies, the optimal tax on capital income can be different from zero.

2 An Auxiliary Optimal Control Problem

The following OCP is essential for the derivation of solutions to the dynamic optimal taxation models presented in the next section:

$$\int_0^{\infty} \mathrm{e}^{-\rho t} L(u(t))\mathrm{d}t \to \max \tag{1}$$

$$\dot{x}(t) = f(x(t))g(x(t)) - u(t) \tag{2}$$

$$x(0) = x_0 \geq 0, x(t) \geq 0,\ u(t) \geq 0 \tag{3}$$

Under certain conditions, the solution of OCP can be found explicitly using the result established in Proposition 1 below.

Proposition 1. *Assume that the following conditions hold true:*

(A1) The function $L(\cdot)$ is differentiable with $L'(u) > 0$ and $L''(u) \leq 0$ for each $u \geq 0$

(A2) The functions $g(\cdot)$ and $f(\cdot)$ are differentiable with $f(0) = 0$ and $f'(x) > 0$ for each $x \geq 0$

(A3) The function $f(\cdot)g(\cdot)$ is concave

(A4) $\displaystyle\limsup_{x\to+\infty} \frac{(f(x) - xf'(x))((f(x)g(x))' - \rho)}{xf'(x)} < \rho$

(A5) $L'(u^*(x))f(x) = \alpha$ *for each* $x \geq 0$*, where* α *is a positive number and*

$$u^*(x) = \frac{f(x)(\rho - f(x)g'(x))}{f'(x)}.$$

Then the feedback control $u^*(\cdot)$ *is the solution of OCP (1)–(3) whenever the corresponding solution* $x^*(\cdot)$ *of (2) is defined on the interval* $[0, +\infty)$.

Proof. First, we note that when $u(t) = u^*(x(t))$, (2) is an autonomous differential equation with continuous right-hand side. Because it has a unique solution $x^*(\cdot)$ and because $x(t) \equiv 0$ is an integral curve of this differential equation, whenever $x_0 \geq 0$ condition (3) is always satisfied.

Define the current value Hamiltonian for the problem (1)–(3) as

$$H(x, u, \pi) = L(u) + \pi(f(x)g(x) - u).$$

The sufficient optimality conditions for this problem (see [9]) require that the maximized Hamiltonian be concave and the following relations be satisfied:

$$\pi = L'(u^*) \tag{4}$$
$$\dot{\pi} = \pi(\rho - f'(x)g(x) - f(x)g'(x)) \tag{5}$$
$$\lim_{t\to\infty} \mathrm{e}^{-\rho t}\pi(t)x(t) = 0.$$

We set $\pi(t) := L'(u^*(x^*(t)), t \in [0, +\infty)$. Then (4) holds true. Moreover, by (A5) and (4) it follows that

$$\pi(t)f(x^*(t)) = \alpha. \tag{6}$$

By taking the time derivative of (6), we obtain

$$\dot{\pi}(t)f(x^*(t)) + \pi(t)f'(x^*(t))\dot{x}^*(t) = 0.$$

Substituting $\dot{x}^*(t)$ with the right-hand side of (2) with $u(t) = u^*(x^*(t))$, the above equation is equivalent to (5).

The maximized Hamiltonian is concave because of the concavity of $f(x)g(x)$ and the fact that $\pi(t) = L'(u^*(x^*(t))) > 0$.

Clearly, if the solution $x^*(t)$, $t \in [0, +\infty)$, is bounded, then the transversality condition will be satisfied. Let us assume that $\lim\limits_{t\to+\infty} x^*(t) = +\infty$. Then, inequality (A4) implies the existence of a real number $\xi < \rho$ and $\bar{x} > 0$ such that

$$\frac{(f(x) - xf'(x))((f(x)g(x))' - \rho)}{xf'(x)} < \xi \quad \text{for each } x > \bar{x}. \tag{7}$$

We set $y(t) := x^*(t)/f(x^*(t)) > 0, t \in [0, +\infty)$. Then for all sufficiently large t we have that

$$\dot{y}(t) = \frac{(f(x^*(t)) - x^*(t)f'(x^*(t)))}{f(x^*(t))^2}\dot{x}^*(t)$$
$$= \frac{(f(x^*(t)) - x^*(t)f'(x^*(t)))((f(x^*(t))g(x^*(t)))' - \rho)}{x^*(t)f'(x^*(t))}y(t)$$

Therefore, $\dot{y}(t) \leq \xi y(t)$ according to (7). This ensures that the transversality condition holds since $\xi < \rho$. The proof is completed. □

In the next section, we show how the above result can be applied to solve specific differential games arising from models of optimal taxation.

3 Three Models of Optimal Taxation

3.1 Redistributive Capital Income Taxation

This model was originally proposed by Judd [5]. He considered an economy inhabited by three types of agents – capitalists, workers, and a government. Capitalists do not work and receive income only from their capital ownership. Capital is used for the production of a single good, which can be either consumed or saved and invested. Workers supply labor inelastically and receive wages. In addition, the government grants them a lump-sum transfer or subtracts a lump-sum tax from their labor income. Production technology is described by the function $F(k)$. Capital depreciates at a constant rate $\delta > 0$ so that $F(k) - \delta k$ is output net of depreciation. Further, capital is assumed to be bounded. The static profit maximization conditions for the competitive firm imply the relations $r = F'(k)$ and $w = F(k) - kF'(k)$, where r denotes the interest rate and w is the wage paid to the worker. The workers' income x is composed of wage earnings plus a government transfer T (or minus a lump-sum tax, depending on the sign of T) which is determined as the tax rate τ times the capital income:

$$x(k, \tau) = w + T = F(k) - kF'(k) + \tau k(F'(k) - \delta).$$

The problem that the representative capitalist faces is to maximize his lifetime utility $U(c)$ subject to the capital accumulation constraint:

$$\int_0^\infty \mathrm{e}^{-\rho t} U(c(t))\mathrm{d}t \to \max$$
$$\dot{k}(t) = (1 - \tau^*(k(t)))f(k(t)) - c(t)$$
$$c(t) \geq 0,\ k(t) \geq 0,\ k(0) = k_0 > 0, \tag{8}$$

where we have set $f(k) := (F'(k) - \delta)k$. The problem that government faces is to find a tax rule τ^* that maximizes welfare. In this setup, welfare is represented by the weighted sum of the instantaneous utilities of the capitalist $U(c)$ and the worker $V(x)$, with γ being the weight. Formally, government solves:

$$\int_0^{\infty} \mathrm{e}^{-\rho t}[\gamma V(x(k(t),\tau(t))) + U(c^*(k(t)))]\mathrm{d}t \to \max$$

with respect to τ and subject to

$$\begin{aligned} \dot{k}(t) &= (1-\tau(t))f(k(t)) - c^*(k(t)) \\ k(t) &\geq 0,\ k(0) = k_0 > 0. \end{aligned} \tag{9}$$

Judd [5] considered the open-loop Stackelberg version of the differential game (8)–(9) and arrived at the interesting conclusion that the optimal tax on capital income tends to zero as time goes to infinity. Although this finding has become broadly accepted in public economics, the validity of the zero limiting capital income tax has been questioned in different contexts. For instance, Lansing [7] argued that if the utility of capitalists is logarithmic the optimal tax is nonzero. However, he claimed that this result is due to the properties of the logarithmic function and will not be true in the general case. Below we show that under more general assumptions about the utility and production functions the optimal tax is different from zero whenever the solution is sought in the class of feedback Nash equilibria.

Proposition 2. *Let*

$$U'(\bar{c}(k))f(k) = \alpha \tag{10}$$

for some positive constant α*, where*

$$\bar{c}(k) = \frac{f(k)(\rho + f(k)\bar{\tau}'(k))}{f'(k)} \tag{11}$$

and $f(k)$ *and* $g(k) = 1-\bar{\tau}(k)$ *satisfy the conditions stated in Proposition 1. Define*

$$H^{g*}(k,\lambda) := \max_{\tau} H^g(k,\tau,\lambda) = \max_{\tau}[U(\bar{c}(k)) + \gamma V(x(\tau,k)) + \lambda((1-\tau)f(k) - \bar{c}(k))].$$

We assume that the trajectory $k(\cdot)$ *corresponding to* $\bar{\tau}(k(\cdot))$ *and* $\bar{c}(k(\cdot))$ *is well defined on the interval* $[0,+\infty)$*, the performance criteria for both players are also well defined and the following hold true:*

(a) $H^{g*}(\cdot,\lambda)$ *is concave for each* $\lambda \geq 0$
(b) $H^{g*}(k(t),\lambda(t)) = H^g(k(t),\bar{\tau}(k(t)),\bar{\lambda}(t))$ *for all* $t \in [0,\infty)$ *where* $\lambda(t) := \gamma V'(x(\bar{\tau}(k(t)),k(t))) \geq 0$ *for each* $t \geq 0$
(c)

$$\begin{aligned} &\gamma V''(x)x'_k(k,\bar{\tau}(k))((1-\bar{\tau}(k))f(k) - \bar{c}(k)) = \\ &\quad = \gamma V'(x)[\rho + \delta - F'(k) + \bar{c}'(k)] - U'(\bar{c}(k))\bar{c}'(k) \end{aligned} \tag{12}$$

(d) $\lim_{t\to\infty} \mathrm{e}^{-\rho t}\gamma V'(x(k(t),\bar{\tau}(k(t))))k(t) = 0$

Then the pair $(\bar{c}(k), \bar{\tau}(k))$ *constitutes a feedback Nash equilibrium for the differential game (8)–(9).*

Proof. The solution of the problem of the capitalist follows directly from the result in Proposition 1. After substituting (11) in the condition $U'(\bar{c}(k))f(k) = \alpha$, we obtain a differential equation for $\bar{\tau}'(k)$ from which we find the function $\bar{\tau}(k)$. It remains to be verified if this function solves the problem of the government.

The sufficient optimality conditions for the OCP of the government are as follows:

$$\lambda(t) := \gamma V'(x(\tau(t), k(t))) \tag{13}$$

$$\dot{\lambda}(t) = \lambda(\rho - F'(k(t)) + \delta + \bar{c}'(k(t))) - U'(\bar{c}(k(t)))\bar{c}'(k(t)) \tag{14}$$

together with conditions (a), (b), and (d) in the proposition statement. If we set

$$\lambda(t) := \gamma V'(x(\bar{\tau}(k(t)), k(t))) \geq 0,$$

then clearly (13) holds true. Taking the time derivative of both sides, we obtain

$$\dot{\lambda}(t) = \gamma V''(x)x'_k(k(t), \bar{\tau}(k(t))((1 - \bar{\tau}(k))f(k) - \bar{c}(k)).$$

By condition (12), this is equivalent to (14). Therefore, $\bar{\tau}(k)$ solves the problem of the government. $\square$

Example 1. Let $U(c) = \dfrac{c^{1-\sigma}}{1-\sigma}$, $F(k) = \dfrac{k^\sigma}{\sigma} + \delta k$ and $V(x) = \dfrac{x^{1-\sigma}}{1-\sigma}$ for $\sigma \in (0, 1)$. Then, it is straightforward to see that $f(k) = k^\sigma$. The first step is to find $\bar{\tau}(k)$ using the result of the proposition. From condition (10), which for this particular example, takes the form $c^{-\sigma}k^\sigma = \alpha$, we obtain that $\bar{c}(k) = \beta k$, where $\beta = \alpha^{-1/\sigma}$. Substituting this in (11), we find the tax function as the solution of the following differential equation

$$\bar{\tau}'(k) = \frac{\sigma\beta - \rho}{k^\sigma}.$$

Therefore,

$$\bar{\tau}(k) = \frac{\sigma\beta - \rho}{1-\sigma}k^{1-\sigma} + \tau_0.$$

Suppose we are given that $\bar{\tau}(0) = \dfrac{\sigma - 1}{\sigma}$. With this initial condition, the workers' income is $x = \dfrac{\sigma\beta - \rho}{1-\sigma}k$.

Having found the candidate $\bar{\tau}(k)$ we need to verify condition (12) as well as to ensure that the maximized Hamiltonian is concave. Using the fact that $\dfrac{V''(x)x}{V'(x)} = -\sigma$, condition (12) simplifies to

$$\gamma(1-\sigma)^{\sigma-1}\left(1 - 2\sigma + \frac{\rho}{\beta}\right) = \left(\frac{\sigma\beta - \rho}{\beta}\right)^\sigma.$$

The latter equation gives a relationship between the parameters of the model. Solving it for β would allow us to determine the functions $\bar{c}(k)$ and $\bar{\tau}(k)$. Finally, the verification of the concavity of the maximized Hamiltonian condition is straightforward when $\sigma \in (0, 1)$.

Once we have derived the optimal controls, we are interested in exploring the behavior of the capital income tax. The solution of the differential equation for capital with $\bar{c}(k)$ and $\bar{\tau}(k)$ can be calculated explicitly:

$$k(t)^{1-\sigma} = \frac{1 - \sigma + e^{(\rho-\beta)t}(\sigma(\beta - \rho)k_0^{1-\sigma} - 1 + \sigma)}{\sigma(\beta - \rho)}.$$

The equilibrium capital income tax $\bar{\tau}_\infty$, which is defined as the limit of $\bar{\tau}(\bar{k}(t))$ as $t \to \infty$, is found to be

$$\bar{\tau}_\infty = \frac{\sigma(\rho - 2\beta) + \beta}{\sigma(\rho - \beta)}.$$

For different values of the model parameters in this example it is possible to have positive or negative limiting capital income taxes. For instance, if $\sigma = 0.8$, $\rho = 0.25$, $\gamma = 7.1895$ and $\beta = 0.4$, then $\tau^*_\infty = 0.3333$. If we set $\sigma = 0.5$, $\rho = 0.5$, $\gamma = \sqrt{2}$ and $\beta = 2$, the equilibrium value of the tax is $\tau^*_\infty = -0.3333$. This is intuitive: with the above choice of parameters capital decreases over time. As capital becomes smaller, it becomes optimal from the welfare perspective to subsidize the capitalist to sustain some level of production.

3.2 Output Tax

Below we examine the problem of finding an optimal policy rule to tax output. A basic version of this model was analyzed by Xie [10] to demonstrate that sometimes open-loop Stackelberg solutions can be time consistent. Karp and Ho Lee [6] considered a slightly more general model with an arbitrary function of capital as a tax base and multiplicative tax policy rule. Here we propose a further generalization of the output tax model in the following sense: instead of solving the optimization problem of the government with respect to the tax rate by treating the tax base as given, we seek a function of capital which is optimal for the government and at the same time agrees with the problem of the consumer. Formally, consumer solves

$$\int_0^\infty e^{-\rho t} U(c(t)) \mathrm{d}t \to \max$$
$$\dot{k}(t) = F(k(t)) - b^*(k(t)) - c(t)$$
$$c(t) \geq 0,\ k(t) \geq 0,\ k(0) = k_0 > 0,$$

where $F(k)$ is the production function and b^* denotes the solution of the government's problem. Government decides about the rule by which to tax the consumer. The problem, which government solves is

$$\int_0^{\infty} \mathrm{e}^{-\rho t}[U(c^*(k(t))) + V(b(t))]\mathrm{d}t \to \max$$

with respect to b and subject to

$$\dot{k}(t) = F(k(t)) - b(t) - c^*(k(t))$$
$$k(t) \geq 0,\ k(0) = k_0 \geq 0.$$

Proposition 3. *Assume that the following condition holds true for some constant α:*

$$(F(k) - \bar{b}(k))U'(\bar{c}(k)) = \alpha,$$

where $\bar{c}(k)$ is given by

$$\bar{c}(k) = \rho \frac{F(k) - \bar{b}(k)}{F'(k) - \bar{b}'(k)}$$

Let $f(k) = F(k) - \bar{b}(k)$ satisfy the conditions of Proposition 1. Define

$$H^{g*}(k, \lambda) = \max_b [U(\bar{c}(k)) + V(b) + \lambda(f(k) - \bar{c}(k))].$$

We assume that the trajectory $k(\cdot)$ corresponding to $\bar{b}(k(\cdot))$ and $\bar{c}(k(\cdot))$ is well defined on the interval $[0, +\infty)$, the performance criteria for both players are also well defined and the following hold true:

(a) $H^{g*}(\cdot, \lambda)$ *is concave for each* $\lambda \geq 0$
(b) $H^{g*}(k(t), \lambda(t)) = H^g(k(t), \bar{b}(k(t)), \lambda(t))$ *where* $\lambda(t) := V'(\bar{b}(k(t)), k(t)))$ ≥ 0 *for each* $t \geq 0$
(c)

$$V''(\bar{b}(k)) = \frac{V'(\bar{b}(k))[\rho - F'(k) + \bar{c}'(k)] - U'(\bar{c}(k))\bar{c}'(k)}{\bar{b}'(k)(F(k) - \bar{b}(k) - \bar{c}(k))}$$

(d) $\lim_{t\to\infty} \mathrm{e}^{-\rho t} V'(\bar{b}(k(t)))k(t) = 0$

Then the Nash equilibrium of the differential game is given by $(\bar{b}(k), \bar{c}(k))$.

The proof is similar to that of Proposition 2 and we omit it.

Example 2. Let $U(c) = \dfrac{c^{1-\sigma}}{1-\sigma}$, $V(b) = \dfrac{b^{1-\sigma}}{1-\sigma}$, and $F(k) = \alpha \left(\dfrac{\rho}{\sigma}\right)^{\sigma} k^{\sigma} + \beta k$. Following the steps in Example 1, it is straightforward to verify that the controls $\bar{c}(k) = \dfrac{\rho}{\sigma} k$ and $\bar{b}(k) = \beta k$ solve the differential game whenever the constant β satisfies the following equation:

$$\beta + \left(\frac{\rho}{\sigma}\right)^{1-\sigma} \beta^{\sigma} = \frac{\rho}{\sigma}.$$

3.3 The Representative Agent Model

The model considered in this section is due to Chamley [1]. He showed that if the utility function is separable in consumption and labor, the optimal tax on capital income is zero in the long run. This is essentially the same result as the one obtained by Judd, despite the differences in the model formulation.

Frankel [4] studied the same problem and reconfirmed Chamley's findings in the case of separable utility. Here, we show that even for separable utility functions it is possible to have nonzero optimal capital tax if the controls of the players are chosen from the feedback strategy space. Before that, we briefly comment on the existing solutions of the representative agent model. For the purpose, it is convenient to state the problems of the government and the agent in open-loop form. We borrow the model formulation from [4]. Consumer solves:

$$\int_0^\infty e^{-\rho t} U(c(t), l(t)) dt \to \max \tag{15}$$

$$\dot{a}(t) = r^*(t)a(t) + w^*(t)l(t) - c(t) \tag{16}$$

$$a(0) = a_0$$

$$\lim_{t \to \infty} e^{-\int_0^t r^*(s) ds} a(t) \geq 0 \tag{17}$$

with respect to consumption c and labor l. Equation (16) describes the evolution of private assets a, which are composed of capital k and government bonds b. By w^* and r^* we denote after-tax returns on labor and capital, respectively.

Government solves the following problem:

$$\int_0^\infty e^{-\rho t} U(c^*(t), l^*(t)) dt \to \max \tag{18}$$

with respect to r and w and subject to

$$\dot{a}(t) = r(t)a(t) + w(t)l^*(t) - c^*(t) \tag{19}$$

$$\dot{k}(t) = F(k(t), l^*(t)) - c^*(t) - g(t)$$

$$a(0) = a_0,\ k(0) = k_0 > 0, k(t) \geq 0, \tag{20}$$

where $F(k,l)$ is the production technology. Whenever the solution of the game is sought as an open-loop Stackelberg equilibrium, the Hamiltonian functions for the OCPs (15)–(17) and (18)–(20) can be written as

$$H(c, l, a, \pi) = U(c, l) + \pi(ar^* + w^*l - c)$$

and

$$\begin{aligned} H(r, w, a, k, \pi, \lambda, \mu, \xi) = {} & U(c^*, l^*) + \lambda(ar + wl^* - c^*) \\ & + \mu(f(k, l^*) - c - g) + \xi\pi(\rho - r), \end{aligned}$$

respectively. In the process of derivation of their results, both Chamley and Frankel used informal arguments about the sign of the co-state variable λ at the initial moment. In particular, they claimed that $\lambda_0 < 0$, which was motivated with economic considerations. Furthermore, Chamley's proof relies essentially on the assumption that $\xi_0 = 0$ which, as Xie [10] later proved, is not necessary for optimality. Both these assumptions may fail to hold true as it can be seen in the example where the utility function is $u(c,l) = \ln c - \frac{l^2}{2}$ and the production function is $F(k,l) = \theta_1 k + \theta_2 l$ for some $\theta_1 > 0$ and $\theta_2 > 0$. For this special case, the optimal tax rate on capital is still zero but the sign of $\lambda(t)$, which coincides with the sign of λ_0, depends on the initial data in the following way:

$$\lambda_0 = \frac{\rho\pi_0^2(\theta_1(a_0 - k_0) + g_0)}{2\theta_1(\rho\pi_0 a_0 - 1)}.$$

Since in order for the integral (15) to be convergent, it is necessary that $\rho\pi_0 a_0 - 1 < 0$, the sign of λ_0 is determined by the sign of $\theta_1(a_0 - k_0) + g_0$. If the government starts with sufficient savings the latter expression could be negative and so λ_0 will be positive. Also, since in this example $\xi_0\pi_0 = a_0\lambda_0$, ξ_0 is different from zero.

To find the feedback Nash equilibrium of the representative agent model, we can adapt the approach used earlier in the paper to obtain a solution to the consumer's problem. Since there are two controls in this problem (c and l) and only one state variable, we need an additional assumption to have unique optimal controls. One possibility is to assume that $\frac{w(a)}{ar(a)} = \beta$, where β is a positive constant. We also replace the transversality condition (17) with

$$\lim_{t\to\infty} \mathrm{e}^{-\rho t} a(t) = 0. \tag{21}$$

The reason is that condition (17) is appropriate only for the case of open-loop controls. Furthermore, it must hold with equality in order the problem of the consumer to have a solution. The transversality condition (21) in turn is a sufficient condition for the considered problem whenever the trajectories are bounded.

Proposition 4. *Assume that the following conditions hold true:*

(B1) The function $U(\cdot,\cdot)$ is differentiable with $U'_c(c,l) > 0$ and $U'_l(c,l) < 0$ for each $l \geq 0$ and each $c \geq 0$

B2) The function $f(a) := ar(a)$ is differentiable and $\bar{c}(a) := \frac{\rho f(a)}{f'(a)} > 0$ for each $a \neq 0$

(B3) $H^*(a,\pi) = \max_{c,l}(U(c,l) + \pi(f(a) + \bar{w}(a)l - c))$ is concave in a for each $\pi > 0$

(B4) The following conditions hold true:

$$U_c'(\bar{c}(a), \bar{l}(a)) f(a) = \alpha,$$
$$\frac{w(a)}{f(a)} = \beta,$$

for some positive constants α *and* β*, where the function* $\bar{l}(a)$ *is determined as the solution of the equation* $U_l'(\bar{c}(a), \bar{l}(a)) = -\alpha\beta$

Then the optimal consumption and the optimal labour input are given by $\bar{c}(a)$ *and* $\bar{l}(a)$*, respectively, whenever the corresponding solution* $a(\cdot)$ *of (19) is defined and bounded on the interval* $[0, +\infty)$.

The proof is essentially the same as that of Proposition 1 and we omit it.

Next we present an example which shows that for a separable utility function it is possible to obtain a capital tax different from zero if the dependence of controls on the state on a is explicitly taken into account.

Example 3. For this example assume that we are given the following functions and initial data. $U(c, l) = \dfrac{c^{1-\sigma}}{1-\sigma} - \dfrac{l^2}{2}$, where $\sigma \in (0, 1)$. $F(k, l) = 2\left(\dfrac{\rho}{\sigma}\right)^{\sigma} k^{\sigma} l^{1-\sigma} + \gamma l$, $a(0) = k(0) = k_0 \geq 0$, $b(0) = 0$, $l \in (0, 1]$ and $g = \gamma$. It can be shown that the controls $\bar{c}(a) = \dfrac{\rho}{\sigma} a$, $\bar{l} = 1$, $\bar{r}(a) = \left(\dfrac{\rho}{\sigma}\right)^{\sigma} a^{\sigma-1}$, $\bar{w}(a) = \left(\dfrac{\rho}{\sigma}\right)^{\sigma} a^{\sigma}$ and the corresponding trajectories

$$a(t)^{1-\sigma} = k(t)^{1-\sigma} = 2\left(\frac{\rho}{\sigma}\right)^{\sigma-1} - \left(2\left(\frac{\rho}{\sigma}\right)^{\sigma-1} - k_0^{1-\sigma}\right) \mathrm{e}^{\frac{\rho}{\sigma}(\sigma-1)t}$$

solve the differential game.[1]

For this particular example government expenditure equals gross income minus net income (income after taxes) which is exactly the collected tax revenue. This in turn means that the government does not have an incentive to accumulate debt or assets and to redistribute income over time. It is interesting to examine the evolution of the capital income tax. Before that, it is convenient to state explicitly the before-tax capital income $\tilde{r}(k)$ using the standard equilibrium condition that $\tilde{r}(k) = F_k'(k, l)$. The gross return on capital is $\tilde{r}(k) = 2\sigma\left(\dfrac{\rho}{\sigma}\right)^{\sigma} k^{\sigma-1}$ and the net return is $\bar{r}(a) = \left(\dfrac{\rho}{\sigma}\right)^{\sigma} k^{\sigma-1} = (1-\tau_k)\tilde{r}(k)$. Thus, we obtain that

$$\tau_k = \frac{2\sigma - 1}{2\sigma}.$$

[1] Note that the controls $\bar{c}$ and $\bar{l}$ are easily derived applying Proposition 4 with $\alpha = \beta = 1$. For the OCP of the government, the sufficient conditions are verified for $\lambda = 0$, where λ is the co-state associated with the private assets equation.

Interestingly, in this example the optimal tax depends only on the intertemporal elasticity of substitution of consumption. It will be negative for $\sigma < 1/2$, values for which the maximized Hamiltonian of the consumer's problem is concave.

4 Conclusion

We find that whenever feedback solutions to the dynamic optimal taxation models are considered, the classical result of zero limiting capital income tax ceases to hold. Directions for future work may include relaxing some of the assumptions made in the paper to broaden the class of problems that can be analyzed in this framework.

Acknowledgements We would like to express our gratitude to the editors and to unknown referees for many constructive comments and remarks which helped us improve the quality of the paper.

References

1. Chamley, C.: Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives. Econometrica **54**(3), 607–622 (1986)
2. Cohen, D., Michel, P.: How Should Control Theory Be Used to Calculate a Time-Consistent Government Policy? Review of Economic Studies **55**(2), 263–274 (1988)
3. Dockner, E., Jorgensen, S., Van Long, N., Sorger, G.: Differential Games in Economics and Management Science. Cambridge University Press, Cambridge (2000)
4. Frankel, D.: Transitional Dynamics of Optimal Capital Taxation. Macroeconomic Dynamics **2**, 492–503 (1998)
5. Judd, D.: Redistributive Taxation in a Simple Perfect Foresight Model. Journal of Public Economics **28**, 59–83 (1985)
6. Karp, L., Ho Lee, In: Time-Consistent Policies. Journal of Economic Theory **112**(2), 353–364 (2003)
7. Lansing, K.: Optimal Redistributive Capital Taxation in a Neoclassical Growth Model. Journal of Public Economics **73**, 423–453 (1999)
8. Rubio, S.J.: On Coincidence of Feedback Nash Equilibria and Stackelberg Equilibria in Economic Applications of Differential Games. Journal of Optimization Theory and Applications **128**(1), 203–221 (2006)
9. Seierstad, A., Sydsaeter, K.: Sufficient Conditions in Optimal Control Theory. International Economic Review **18**(2), 367–391 (1977)
10. Xie, D.: On Time Inconsistency: A Technical Issue in Stackelberg Differential Games. Journal of Economic Theory **76**(2), 412–430 (1997)

On a Discounted Inventory Game

Heinz-Uwe Küenle

Abstract We investigate a problem where a company requires a comodity for its production. The demand is random. The company has one main supplier. It can also buy this commodity from the free market, but for a higher price. The supplier produces this commodity and delivers first to the company, but it has also an additional random demand. The problem is considered as a stochastic game with perfect information. We treat pairs of production strategies of a certain structure and show that there is such a pair which is a Nash equilibrium.

1 Introduction

The problem investigated in this paper belongs to the family of stochastic dynamic inventory problems. Such problems were introduced by Arrow, Harris, and Marschak [2] and they can be considered as Markov decision processes. Scarf [16], Iglehart [8], and Johnson [10] showed the optimality of so-called (s, S)-strategies. Inventory problems with incompletely known demand distribution lead to minimax decision problems (see [1, 9, 11]). These minimax problems can be considered as zero-sum Markov games with perfect information (see [6, 7, 12–14]), and optimal (s, S)-strategies (or (σ, S)-strategies in the multi-product case) can be shown to exist.

There are a lot of applications of nonzero-sum games in supply chain management [4, 15], but not too many of these applications concern stochastic games (see [4], Sect. 3.2).

Here, we treat a problem with two companies (players), which have different but not contrary aims: A company (Company 1) produces a commodity (Commodity 1) and requires for its production another commodity (Commodity 2). The demand for Commodity 1 is random. The commodities are stored in warehouses. Company 1

H.-U. Küenle (✉)
Brandenburg University of Technology Cottbus, Department of Mathematics, PF 10 13 44, D-03013 Cottbus, Germany
e-mail: kuenle@math.tu-cottbus.de

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_26,

has one main supplier for Commodity 2, namely Company 2. It can buy this product also from the free market, but for a higher price. Furthermore, both companies have convex holding and shortage costs. Company 2 produces Commodity 2 and delivers first to Company 1, but it has also an additional random demand. Unfilled demand is backlogged. The ordered quantities are instantaneously delivered or produced. We show that a certain pair of ordering/production strategies is a Nash equilibrium.

A similar but easier problem was treated by Schurath [17]. In his approach, the second player knows the quantity of the order of the first player before the production starts. Yang [18] has considered the above-mentioned problem where the cost function of Company 2 has a more complicated structure. Unfortunately, her proof is not complete, since some required convexity properties of the value function of the second player are not shown.

This chapter is organized as follows: In Sect. 2 stochastic games with perfect information are considered. The production-inventory model is introduced in Sect. 3. In Sect. 4, we prove that a certain pair of stationary strategies is a Nash equilibrium. A numerical example is given in Sect. 5.

2 Stochastic Games with Perfect Information

Stochastic games with perfect information considered in this paper are defined by the following objects:

(i) S is the *state space* and A and B are the *action spaces* for players 1 and 2. All three sets are assumed to be Borel sets in Polish spaces.
(ii) The *constraint set of the first player* K_A is assumed to be a Borel subset of $S \times A$. For each state $x \in S$, the set $A(s) := \{a \in A : (s,a) \in K_A\}$ represents the *set of admissible actions* for Player 1 in state s, and it is assumed that $A(\cdot)$ has a Borel measurable selector.
(iii) The *constraint set of the second player* K is assumed to be a Borel subset of $K_A \times B$. For each state $(s,a) \in K_A$, the set $B(s,a) := \{b \in B : (s,a,b) \in K\}$ represents the *set of admissible actions* for Player 2 in state S if Player 1 has chosen $a \in A(s)$. It is assumed that $B(\cdot)$ has a Borel measurable selector.
(iv) p is a Borel measurable transition probability from K to S, the *law of motion*.
(v) $k^{(i)} : K \to \mathbb{R}$ is a Borel measurable function from K to $\mathbb{R}$, the *cost function for Player* i $(i = 1, 2)$.
(vi) $\alpha \in (0, 1)$ is the *discount factor*.

We can describe the situation of such a stochastic game as follows: At the discrete time point $n \in \mathbb{N}_0$ the game is in state $s \in S$. First Player 1 decides for his action $a \in A(s)$ and then Player 2 for $b \in B(s,a)$. Player 1 has to pay costs $k^{(1)}(s,a,b)$ while Player 2 has to pay an amount of $k^{(2)}(s,a,b)$. At the beginning of the next period $n+1$, the game is in a state s' which is chosen according to the probability distribution $p(\cdot | s,a,b)$. The goal of both players is to minimize their expected total discounted costs.

Let $H_n := K^n \times S$ for $n \geq 1$, $H_0 = S$. A transition probability π_n from H_n to A with $\pi_n(A(s_n)|\, s_0, a_0, b_0, \ldots, s_n) = 1$ for all $(s_0, a_0, b_0, \ldots, s_n) \in H_n$ is called a *decision rule of the first player at time* n.

Let $H'_n := K^n \times K_A$ for $n \geq 1$, $H'_0 = K_A$. A transition probability ρ_n from H'_n to B with $\rho_n(B(s_n)|\, s_0, a_0, b_0, \ldots, s_n, a_n) = 1$ for all $(s_0, a_0, b_0, \ldots, s_n, a_n) \in H'_n$ is called a *decision rule of the second player at time* n.

A decision rule π_n of the first player is called *Markov deterministic* iff a function $f : S \to A$ with $f(s) \in A(s)$ for all $s \in S$ exists such that $\pi_n(\{f(s_n)\}|\, s_0, a_0, b_0, \ldots, s_n) = 1$ for all $(s_0, a_0, b_0, \ldots, s_n) \in H_n$. A decision rule ρ_n of the second player is called *Markov deterministic* iff a function $g : K_A \to B$ with $g(s,a) \in B(s,a)$ for all $(s,a) \in K_A$ exists such that $\rho_n(\{g(s_n, a_n)\}|\, s_0, a_0, b_0, \ldots, s_n, a_n) = 1$ for all $(s_0, a_0, b_0, \ldots, s_n, a_n) \in H'_n$.

(Notation: We identify f with π_n and g with ρ_n .)

E and F denote the sets of all Markov deterministic decision rules. A sequence $\Pi = (\pi_n)$ or $P = (\rho_n)$ of decision rules of the first or second player is called a *strategy* of that player.

A strategy $\Pi = (\pi_n)$ of Player 1 with $\pi_n = f$ for all $n \in \mathbb{N}_0$ is called *stationary deterministic* and denoted by f^∞. A stationary deterministic strategy g^∞ of the second player is analogously defined.

Assumption 1. *There are a measurable function* $V : S \to [1, \infty)$*, and a constant* $\lambda \in (0,1)$ *such that*

(a) $\displaystyle\sup_{(s,a,b)\in K} \frac{|k^{(i)}(s,a,b)|}{V(s)} < \infty$ *for* $i = 1, 2$

(b) $\alpha \int_S V(t) p(\mathrm{d}t|s,a,b) \leq \lambda V(s)$ *for each* $(s,a,b) \in K$

We assume in this section that Assumption 1. is fulfilled.

Let us consider the space $\Omega := (S \times A \times B)^\infty$, endowed with the product σ-algebra $\mathscr{F}$. For every strategy pair (Π, P) and each initial state $s = s_0 \in S$, there exist a probability measure $P_{s,\Pi,P}$ and a stochastic process (S_n, A_n, B_n) on $(\Omega, \mathscr{F})$, which are defined in a canonical way. Here, the random variables S_n, A_n, and B_n represent the state and the actions of players 1 and 2 in stage n. Let $E_{s,\Pi,P}$ be the expectation operator with respect to $P_{s,\Pi,P}$. Then, for a strategy pair (Π, P) and initial state $s \in S$, the *total expected discounted reward* for Player i is

$$V^{(i)}_{\Pi,P}(s) = E_{s,\Pi,P} \sum_{m=0}^{\infty} \alpha^m k^{(i)}(S_m, A_m, B_m),$$

where Assumption 1. implies that the expectation exists. Let $S' \subseteq S$. A strategy pair (Π^*, P^*) is called a *Nash equilibrium pair on* S' if

$$V^{(1)}_{\Pi^*,P^*}(s) \leq V^{(1)}_{\Pi,P^*}(s) \tag{1}$$

and

$$V^{(2)}_{\Pi^*,P^*}(s) \le V^{(2)}_{\Pi^*,P}(s) \tag{2}$$

for all $s \in S'$, Π, P. In the zero-sum case, every discounted stochastic game with perfect information and finite state and action spaces has a Nash equilibrium pair (that means an optimal strategy pair). Unfortunately, this property does not hold in a general case. An appropriate example can be found in [5]. On the other hand, making use of standard methods one is able to prove that a given stationary strategy pair is a Nash equilibrium. Namely, (1) and (2) are equivalent to

$$V^{(1)}_{\Pi^*,P^*}(s) = \min_{\Pi} V^{(1)}_{\Pi,P^*}(s) \tag{3}$$

and

$$V^{(2)}_{\Pi^*,P^*}(s) = \min_{P} V^{(2)}_{\Pi^*,P}(s) \tag{4}$$

for all $s \in S'$. If $\Pi^* = f^{*\infty}$ and $P^* = g^{*\infty}$ are stationary deterministic strategies then proving of (3) and (4) reduce to problems of Markov decision theory.

Let $\mathfrak{V}$ be the set of all lower semianalytic functions u with

$$\|u\|_V := \sup_{s\in S} \frac{|u(s)|}{V(s)} < \infty.$$

We consider on $\mathfrak{V}$ a metric defined by the V-norm $\|\cdot\|_V$. Then $\mathfrak{V}$ is a complete metric space. Let f be an arbitrary Markov deterministic decision rule of the first player and g an arbitrary Markov deterministic decision rule of the second player. We define on $\mathfrak{V}$ the following operators:

$$T^{(i)}_{f,g}u(s) := k^{(i)}(s, f(s), g(s, f(s))) + \alpha \int_S u(t)p(\mathrm{d}t \,|\, s, f(s), g(s, f(s)))$$

$$T^{(1)}_{g}u(s) := \inf_{a\in A(s)} \left\{ k^{(1)}(s, a, g(s, a)) + \alpha \int_S u(t)p(\mathrm{d}t \,|\, s, a, g(s, a)) \right\}$$

$$T^{(2)}_{f}u(s) := \inf_{b\in B(s,f(s))} \left\{ k^{(2)}(s, f(s), b) + \alpha \int_S u(t)p(\mathrm{d}t \,|\, s, f(s), b) \right\}$$

for all $s \in S$, $u \in \mathfrak{V}$, $i = 1, 2$.

Lemma 1.

(i) *The operators $T^{(i)}_{f,g}$, $T^{(1)}_{g}$, and $T^{(2)}_{f}$ are isotonic and contractive on $\mathfrak{V}$.*
(ii) *Let $u_0 \in \mathfrak{V}$, $u_n := T^{(1)}_{g}u_{n-1}$. Then $(u_n)_{n\in\mathbb{N}_0}$ converges in V-norm to the unique fixed point $v^{(1)}_{g}$ of $T^{(1)}_{g}$ in $\mathfrak{V}$.*
(iii) *Let $u_0 \in \mathfrak{V}$, $u_n := T^{(2)}_{f}u_{n-1}$. Then $(u_n)_{n\in\mathbb{N}_0}$ converges in V-norm to the unique fixed point $v^{(2)}_{f}$ of $T^{(2)}_{f}$ in $\mathfrak{V}$.*

Theorem 1. *Let $(f^{*\infty}, g^{*\infty})$ be a stationary strategy pair. $(f^{*\infty}, g^{*\infty})$ is a Nash equilibrium pair on $S' \subseteq S$ if $p(S'|\, s, f^*(s), g^*(s, f^*(s))) = 1$ for all $s \in S'$ and the following equations are satisfied for all $s \in S'$:*

$$v^{(1)}_{g^*}(s) = T^{(1)}_{f^*,g^*} v^{(1)}_{g^*}(s) \tag{5}$$

and

$$v^{(2)}_{f^*}(s) = T^{(2)}_{f^*,g^*} v^{(2)}_{f^*}(s). \tag{6}$$

Proof. We remark that the fixed point equation $v^{(1)}_{g^*} = T^{(1)}_{g^*} v^{(1)}_{g^*}$ is Bellman's optimality equation in a corresponding Markov decision model. It follows by standard methods (see [3], for instance) that $v^{(1)}_{g^*}(s) \le V^{(1)}_{\Pi,g^{*\infty}}(s)$ for all Π and $s \in S'$. On the other hand, from (5) it follows that $v^{(1)}_{g^*}(s) = V^{(1)}_{f^{*\infty},g^{*\infty}}(s)$ for all $s \in S'$. Therefore, $V^{(1)}_{f^{*\infty},g^{*\infty}}(s) \le V^{(1)}_{\Pi,g^{*\infty}}(s)$ for all Π and $s \in S'$. Analogously, (6) implies that $V^{(2)}_{f^{*\infty},g^{*\infty}}(s) \le V^{(2)}_{f^{*\infty},P}(s)$ for all P and $s \in S'$. □

3 The Production-Inventory Model

We treat the following periodic review production-inventory problem: A company (Company 1) requires a commodity for its production. The demand in period $[n, n+1)$ is random and denoted by $\hat{\xi}_n$. The commodity is stored in a warehouse. Company 1 has one main supplier for this commodity (Company 2). It can buy this product also from the free market, but for a higher price. Company 2 produces this commodity and delivers first to Company 1, but it has also an additional random demand $\hat{\eta}_n$ in period $[n, n+1)$. We assume that $\hat{\xi}_n$ and $\hat{\eta}_n$ are independent, and also that the demands in several periods are independent of one another and have finite expectations. Unfilled demand is backlogged (as negative stock). We assume that the ordered quantities will be instantaneously delivered or produced. Let x_n (y_n) be the stock in the inventory of Company 1 (2) at time n. If Company 1 orders a quantity $\tilde{a}_n$, then it means that it buys $\min\{\tilde{a}_n, y_n\}$ from Company 2 and the rest from the free market. We put $a_n := \tilde{a}_n + x_n$. Let $\tilde{b}_n$ be the quantity that will be produced by Company 2 in period $[n, n+1)$. Then we put $b_n := \tilde{b}_n + y_n - \min\{\tilde{a}_n, y_n\}$. We get

$$x_{n+1} = a_n - \hat{\xi}_n,$$
$$y_{n+1} = b_n - \hat{\eta}_n.$$

Let $L^{(1)}(a_n)$ be the (expected) holding and shortage cost of Company 1 in period $[n, n+1)$ where the stock after the order is a_n. We suppose that $L^{(1)} : \mathbb{R} \to \mathbb{R}$ is convex on $\mathbb{R}$ with $\lim_{|u|\to\infty} L^{(1)}(u) = \infty$. Often it is assumed that the costs are $h^{(1)}$ for a unit of stock and $g^{(1)}$ for a unit of backlogged demand at the end of the period. This leads to

$$L^{(1)}(u) = \int\limits_{R_+} \left(\mathbb{I}_{[0,\infty)}(u-t)h^{(1)}(u-t) + \mathbb{I}_{[-\infty,0)}(u-t)g^{(1)}(t-u) \right) \psi^{(1)}(t)\lambda(\mathrm{d}t), \tag{7}$$

where ψ^i denotes the density of the random demand of the inventory of Company i. Furthermore, let c_1 be the price for one unit of the commodity, if Company 1 buys it from Company 2, and c_2 the price for one unit, if Company 1 buys it from the free market ($c_2 > c_1$). We assume $\lim_{x \to -\infty}((c_2 - \alpha c_1)x + L^{(1)}(x)) = \infty$. The last property means that for $x \to -\infty$ the shortage cost predominate over the ordering cost.

Concerning Company 2, we assume that in period $[n, n+1)$ cost $L^{(2)}(z_n)$ occurs where $z_n := y_n - (a_n - x_n)$. $|z_n|$ is the remaining stock in the inventory of Company 2 after subtraction of the order of Company 1 (for $z_n \geq 0$) or the part of the order of Company 1, which cannot be satisfied by Company 2 (for $z_n < 0$). We suppose that $L^{(2)} : \mathbb{R} \to \mathbb{R}$ is convex on $\mathbb{R}$ with $\lim\limits_{|z| \to \infty} L^{(2)}(z) = \infty$, $L^{(2)} \geq 0$, and $L^{(2)}(0) = 0$. An example for $L^{(2)}$ is

$$L^{(2)}(z) = g^{(2)}[z]^- + h^{(2)}[z]^+ \text{ for } z \in (-\infty, \kappa^{(2)}] \quad (h^{(2)}, g^{(2)} > 0) \tag{8}$$

with $[z]^+ := \max\{z, 0\}$, $[z]^- := \max\{-z, 0\}$.

We will consider the above described inventory problem as a noncooperative stochastic game. Clearly, the players are the companies. Furthermore:

- Let $S := X \times Y$ be the state space of the stochastic game where $X := (-\infty, \kappa^{(1)}]$ and $Y := (-\infty, \kappa^{(2)}]$ with $\kappa^{(1)} > 0, \kappa^{(2)} > 0$. If $x \in X$, $x \geq 0$, then x denotes the stock in the inventory of Company 1. If $x < 0$, then $-x$ means the backlogged demand. Analogously, $y \in Y$ is the stock or the negative backlogged demand of the inventory of Company 2. $\kappa^{(1)}$ and $\kappa^{(2)}$ are the capacities of the inventories.
- Let $A(x, y) = [x^+, \kappa^{(1)}]$ for $(x, y) \in S$. (Remember that we have assumed that the ordered quantities will be delivered instantaneously.)
- Let $B(x, y, a) = \left[y - \min\{a - x, y\}, \kappa^{(2)} \right] = \left[[x + y - a]^+, \kappa^{(2)} \right]$ for $(x, y) \in S$, $a \in A(x, y)$. Then $b \in B(x, y, a)$ means the stock of Company 2 after an order, if the stock before this order was y and the quantity $\min\{a - x, y\}$ was delivered to Company 1. That means that $b - y + \min\{a - x, y\}$ is the quantity ordered by Company 2.
- p is given by $p(C \mid x, y, a, b) := \int_0^\infty \int_0^\infty \mathbb{I}_C(a - \xi, b - \eta)\psi^{(1)}(\xi)\psi^{(2)}(\eta)\mathrm{d}\xi\mathrm{d}\eta$ for every open set $C \in \mathbb{R}$.
- $k^{(1)}(x, y, a, b) := c_1 \cdot \min\{a - x, y^+\} + c_2 \cdot [a - (x + y^+)]^+ + L^{(1)}(a)$
  $k^{(2)}(x, y, a, b) := L^{(2)}(x + y - a)$

We remark that it is easy to find a linear function V for which Assumption 1. is satisfied, if (8) holds.

4 Equilibrium Pairs in the Production-Inventory Model

In this section, we introduce a pair of stationary strategies and show that this pair is a Nash equilibrium.

Let

$$G_1(a) := (c_2 - \alpha c_1)a + L^{(1)}(a)$$

and

$$G_2(a) := (1-\alpha)c_1 a + L^{(1)}(a)$$

for all $a \in (-\infty, \kappa^{(1)}]$.

Let furthermore $a^* \in (-\infty, \kappa^{(1)}]$ with

$$G_1(a^*) = \min_{a\in(-\infty,\kappa^{(1)}]} G_1(a)$$

and $a^{**} \in (-\infty, \kappa^{(1)}]$ with

$$G_2(a^{**}) = \min_{a\in(-\infty,\kappa^{(1)}]} G_2(a).$$

Since $c_1 < c_2$ and $G_1(a) = G_2(a) + (c_2 - c_1)a$, we get $G_1(a^*) \le G_1(a^{**}) = G_2(a^{**}) + (c_2 - c_1)a^{**} \le G_2(a^*) + (c_2 - c_1)a^{**} = G_1(a^*) + (c_2 - c_1)(a^{**} - a^*)$. Hence, $a^* \le a^{**}$. We assume that $a^* \ge 0$.

We denote by f^* a decision rule of the first player with

$$f^*(x,y) = \begin{cases} x & \text{if } a^{**} < x \\ a^{**} & \text{if } x \le a^{**} < x+y \\ x+y & \text{if } a^* \le x+y < a^{**} \\ a^* & \text{if } x+y < a^* \end{cases}. \tag{9}$$

Let $\tilde{L}(x,y) := L^{(2)}(x + y - f^*(x,y))$. For $x \le a^{**}$ it holds

$$\tilde{L}(x,y) = \begin{cases} L^{(2)}(x+y-a^{**}) & \text{if } a^{**} \le x+y \\ L^{(2)}(0) & \text{if } a^* \le x+y < a^{**} \\ L^{(2)}(x+y-a^*) & \text{if } x+y < a^* \end{cases}.$$

Obviously, $\tilde{L}(x,y)$ depends only on the sum $z := x+y$ (for $x \le a^{**}$) and is convex in z. Hence,

$$K(x+y) := \int_0^\infty \int_0^\infty \tilde{L}(x-\xi, y-\eta)\psi_1(\xi)\psi_2(\eta)\mathrm{d}\xi\mathrm{d}\eta$$

has the same properties. Furthermore, $\lim\limits_{|z|\to\infty} K(z) = \infty$.

Let $b^* \in \mathbb{R}$ with

$$K(b^*) = \min_{b\in\mathbb{R}} K(b).$$

Since $K(z)$ is not increasing for $z \le a^{**}$, it holds that $b^* \ge a^{**} \ge 0$. Let g^* be a decision rule of the second player with

$$g^*(x, y, a) = b^* - a \quad \text{for} \quad b^* - a \in B(x, y, a). \tag{10}$$

That means

$$g^*(x, y, a) = b^* - a \quad \text{for} \quad \max\{x + y, a\} \le b^*. \tag{11}$$

For $x + y > b^*$ or $a > b^*$, let $g^*(x, y, a) \in B(x, y, a)$ be arbitrary.

Our aim is to show that $(f^{*\infty}, g^{*\infty})$ is an equilibrium pair. Unfortunately, we will succeed in proving the equilibrium property only for initial states from a certain subset S' of the state space S, where S' is a closed set of the Markov chain (S_n) under $(f^{*\infty}, g^{*\infty})$. Such a phenomenon is also known for other multiitem inventory problems (see [12], for instance). We set

$$S' := \{(x, y) \in S : x \le a^{**}, 0 \le y \le b^*, x + y \le b^*\}.$$

For proving that S' is closed, we need the following additional assumptions.

Assumption 2.

(a) $\kappa^{(2)} \ge b^* + a^{**} \ge \kappa^{(1)}$.

(b) $a^{**} < b^*$ *and* $\psi_2(\eta) = 0$ *for* $\eta > b^* - a^{**}$.

Now we will show that the state cannot leave the set S', if the strategy pair $(f^{*\infty}, g^{*\infty})$ is played.

Lemma 2. *Let Assumption 2. be satisfied. Then* $p(S'| x, y, f^*(x, y), g^*(x, y, f^*(x, y))) = 1$ *for all* $(x, y) \in S'$.

Proof. If $(x_n, y_n) \in S'$ and the decision rules f^* and g^* are used, then we have first $a_n = f^*(x_n, y_n) \le a^{**}$. Therefore, $x_{n+1} = a_n - \hat{\xi}_n \le a^{**}$ (a.e.). Moreover, it holds that $b^* - a_n \le \kappa^{(2)}$, $b^* - a_n \ge 0$, and $b^* - a_n \ge x_n + y_n - a_n$. Hence, $b^* - a_n \in B(x_n, y_n, a_n)$. It follows that $b_n = g^*(x_n, y_n, a_n) = b^* - a_n$. We get $y_{n+1} = b_n - \hat{\eta}_n = b^* - a_n - \hat{\eta}_n \le b^*$ and $x_{n+1} + y_{n+1} = a_n + b_n - \hat{\xi}_n - \hat{\eta}_n = b^* - \hat{\xi}_n - \hat{\eta}_n \le b^*$ (a.e.). Furthermore, $y_{n+1} = b^* - a_n - \hat{\eta}_n \ge b^* - a^{**} - \hat{\eta}_n \ge 0$ (a.e.). Therefore, $(x_{n+1}, y_{n+1}) \in S'$ (a.e.). This implies the statement. □

Now we can formulate the main theorem.

Theorem 2. *Let Assumption 2. be satisfied. Then a strategy pair* $(f^{*\infty}, g^{*\infty})$ *with (9) and (10) is a Nash equilibrium pair on* S'.

Proof. First, we will consider the best answer of the first player to the strategy $g^{*\infty}$ of the second player.

Let $\bar{\psi}$ be a density of $\zeta_n := \xi_n + \eta_n$. We put

$$\bar{W} := (1-\alpha)^{-1}\Bigg(G_2(a^{**})\int_0^{b^*-a^{**}} \bar{\psi}(\zeta)\mathrm{d}\zeta + \int_{b^*-a^{**}}^{b^*-a^*} G_2(b^*-\zeta)\bar{\psi}(\zeta)\mathrm{d}\zeta$$
$$+ \int_{b^*-a^*}^{\infty} (G_1(a^*) - (c_2-c_1)(b^*-\zeta))\bar{\psi}(\zeta)\mathrm{d}\zeta + c_1\int_0^{\infty} \xi\psi_1(\xi)\mathrm{d}\xi\Bigg).$$

Let

$$u_0(x,y) = \alpha\bar{W} - c_1 x + \begin{cases} G_2(a^{**}) & \text{for } x \le a^{**} \le x+y \\ G_2(x+y) & \text{for } a^* \le x+y < a^{**} \\ G_1(a^*) - (c_2-c_1)(x+y) & \text{for } x+y < a^* \end{cases}$$

Then it holds for $(x,y) \in S'$

$$u_0(a-\xi, g^*(x,y,a)-\eta) = \alpha\bar{W} - c_1(a-\xi) + \begin{cases} G_2(a^{**}) & \text{for } \zeta \le b^*-a^{**} \\ G_2(b^*-\zeta) & \text{for } b^*-a^{**} < \zeta \le b^*-a^* \\ G_1(a^*) - (c_2-c_1)(b^*-\zeta) & \text{for } b^*-a^* < \zeta \end{cases}$$

Hence,

$$\int_0^{\infty}\int_0^{\infty} u_0(a-\xi, g(x,y,a)-\eta)\psi_1(\xi)\psi_2(\eta)\mathrm{d}\xi\mathrm{d}\eta$$
$$= \alpha\bar{W} - c_1 a + (1-\alpha)\bar{W} = \bar{W} - c_1 a.$$

It holds for $u \in \mathfrak{V}$

$$T_{g^*}^{(1)}u(x,y) = \inf_{a\in[x^+,\kappa^{(1)}]}\Bigg\{k^{(1)}(x,y,a,g^*(x,y,a))$$
$$+\alpha\int_0^{\infty}\int_0^{\infty} u(a-\xi, g^*(x,y,a)-\eta)\psi_1(\xi)\psi_2(\eta)\mathrm{d}\xi\mathrm{d}\eta\Bigg\}$$

for all $(x,y) \in S$.

Let $u_n(x,y) := u_0(x,y) + w_n(x)$ for $(x,y) \in S$, where $w_n \ge 0$ is a convex function with $w_n(x) = 0$ for $x \le a^{**}$, $w_0 \equiv 0$. Let $\bar{w}_n(a) := \int_0^{\infty} w_n(a-\xi)\psi_1(\xi)\mathrm{d}\xi$. Clearly, $\bar{w}_n(a) = 0$ for $a \le a^{**}$ and $\bar{w}_n$ is convex. We put

$$w_{n+1}(a) := G_2(a) - G_2(a^{**}) + \alpha\bar{w}_n(a)$$

for $a > a^{**}$ and $w_{n+1}(a) = 0$ for $a \le a^{**}$. Then w_{n+1} is also nonnegative and convex. Because $w_1 \ge w_0 \equiv 0$, it follows also that the sequence (w_n) is nondecreasing.

We will consider $u_{n+1} = T^{(1)}_{g^*} u_n$. Let first assume $x \ge 0$. Then it holds

$$\begin{aligned}
T^{(1)}_{g^*} u_n(x,y) &= \min_{a\in[x,\kappa^{(1)}]} \big\{k^{(1)}(x,y,a,g^*(x,y,a)) + \alpha(\bar{W} - c_1 a + \bar{w}_n(a))\big\} \\
&= \min\Bigg\{ \min_{a\in[x+y,\kappa^{(1)}]} \big\{c_1 y + c_2\cdot(a-(x+y)) + L^{(1)}(a) - \alpha c_1 a + \alpha\bar{w}_n(a)\big\}, \\
&\qquad \min_{a\in[x,x+y]} \big\{c_1(a-x) + L^{(1)}(a) - \alpha c_1 a + \alpha \bar{w}_n(a)\big\}\Bigg\} + \alpha\bar{W} \\
&= \min\Bigg\{ c_1 y - c_2\cdot(x+y) + \min_{a\in[x+y,\kappa^{(1)}]}\{G_1(a) + \alpha\bar{w}_n(a)\}, \\
&\qquad -c_1 x + \min_{a\in[x,x+y]}\{G_2(a) + \alpha\bar{w}_n(a)\}\Bigg\} + \alpha\bar{W} \\
&= \min\Bigg\{ (c_1 - c_2)(x+y) + \min_{a\in[x+y,\kappa^{(1)}]}\{G_1(a) + \alpha\bar{w}_n(a)\}, \\
&\qquad \min_{a\in[x,x+y]}\{G_2(a) + \alpha\bar{w}_n(a)\}\Bigg\} + \alpha\bar{W} - c_1 x
\end{aligned}$$

for all $(x,y) \in S$, $x \ge 0$.

Let $x + y \le a^*$. Then

$$\begin{aligned}
T^{(1)}_{g^*} u_n(x,y) &= \alpha\bar{W} - c_1 x + \min\{(c_1-c_2)(x+y) + G_1(a^*), G_2(x+y)\} \\
&= \alpha\bar{W} - c_1 x + \min\{(c_1-c_2)(x+y) + G_1(a^*), \\
&\qquad\qquad (c_1-c_2)(x+y) + G_1(x+y)\} \\
&= \alpha\bar{W} - c_1 x + (c_1-c_2)(x+y) + G_1(a^*) \\
&= u_0(x,y).
\end{aligned}$$

Let $a^* \le x + y \le a^{**}$. Then

$$\begin{aligned}
T^{(1)}_{g^*} u_n(x,y) &= \alpha\bar{W} - c_1 x + \min\{(c_1-c_2)(x+y) + G_1(x+y), G_2(x+y)\} \\
&= \alpha\bar{W} - c_1 x + \min\{G_2(x+y), G_2(x+y)\} \\
&= \alpha\bar{W} - c_1 x + G_2(x+y) \\
&= u_0(x,y).
\end{aligned}$$

Let $x \le a^{**} \le x + y$. Then

$$\begin{aligned}
&T^{(1)}_{g^*} u_n(x,y) \\
&\quad= \alpha\bar{W} - c_1 x + \min\{(c_1-c_2)(x+y) + G_1(x+y) + \alpha\bar{w}_n(x+y), G_2(a^{**})\} \\
&\quad= \alpha\bar{W} - c_1 x + \min\{G_2(x+y) + \alpha\bar{w}_n(x+y), G_2(a^{**})\}
\end{aligned}$$

$$= \alpha\bar{W} - c_1 x + G_2(a^{**})$$
$$= u_0(x, y).$$

Let $a^{**} \le x$. Then

$$\begin{aligned}
&T^{(1)}_{g^*} u_n(x, y)\\
&\quad = \alpha\bar{W} - c_1 x + \min\{(c_1 - c_2)(x + y) + G_1(x + y)\\
&\qquad + \alpha\bar{w}_n(x + y), G_2(x) + \alpha\bar{w}_n(x)\}\\
&\quad = \alpha\bar{W} - c_1 x + \min\{G_2(x + y) + \alpha\bar{w}_n(x + y), G_2(x) + \alpha\bar{w}_n(x)\}\\
&\quad = \alpha\bar{W} - c_1 x + G_2(x) + \alpha\bar{w}_n(x)\\
&\quad = u_0(x, y) + \alpha\bar{w}_n(x) + G_2(x) - G_2(a^{**})\\
&\quad = u_0(x, y) + w_{n+1}(x) = u_{n+1}(x, y).
\end{aligned}$$

We will now consider the case $x < 0$. For $x + y \le 0$, we have

$$\begin{aligned}
T^{(1)}_{g^*} u_n(x, y) &= \min_{a\in[0,\kappa^{(1)}]} \left\{k^{(1)}(x, y, a, g^*(x, y, a)) + \alpha(\bar{W} - c_1 a + \bar{w}_n(a))\right\}\\
&= \min_{a\in[0,\kappa^{(1)}]} \left\{c_1 y + c_2 \cdot (a - (x + y)) + L^{(1)}(a) - \alpha c_1 a + \alpha\bar{w}_n(a)\right\}\\
&= \alpha\bar{W} - c_1 x + (c_1 - c_2)(x + y) + G_1(a^*)\\
&= u_0(x, y).
\end{aligned}$$

For $x + y > 0$, the calculations are analogous to the case when $x \ge 0$.

Hence,

$$T^{(1)}_{g^*} u_n(x, y) = u_0(x, y) + w_{n+1}(x) = u_{n+1}(x, y)$$

for all $(x, y) \in S$. Furthermore, we observe that

$$T^{(1)}_{g^*} u_n(x, y) = T^{(1)}_{f^*g^*} u_n(x, y)$$

for all $(x, y) \in S$.

The sequence (u_n) converges to the fixed point u^* of the operator $T^{(1)}_{g^*}$. Hence, $u^* = u_0 + w^*$, where $w^* = \lim_{n\to\infty} w_n$. $w^* \ge 0$ is a convex function with $w^*(x) = 0$ for $x \le a^{**}$. An analogous calculation as above shows that

$$u^* = T^{(1)}_{g^*} u^* = T^{(1)}_{f^*g^*} u^*.$$

Hence, (5) is satisfied.

Now we ask for the best answer of the second player to the strategy $f^{*\infty}$ of the first player.

We consider the optimality equation

$$v = T^{(2)}_{f^*} v \tag{12}$$

with

$$T^{(2)}_{f^*}v(x,y) = \inf_{b\in B(x,y,f^*(x,y))}\Big\{k^{(2)}(x,y,f^*(x,y),b) + \alpha\int_0^\infty\int_0^\infty v(f^*(x,y)-\xi, b-\eta)\psi_1(\xi)\psi_2(\eta)\mathrm{d}\xi\mathrm{d}\eta\Big\}.$$

For $x \le a^{**}$, the optimality equation (12) is equivalent to

$$v(x,y) = L^{(2)}(x+y-f^*(x,y)) + \alpha\inf_{b\in[[y-f^*(x,y)]^+,\kappa^{(2)}]}\Big\{\int_0^\infty\int_0^\infty v(f^*(x,y)-\xi, b-\eta)\psi_1(\xi)\psi_2(\eta)\mathrm{d}\xi\mathrm{d}\eta.\Big\} \tag{13}$$

We put $\tilde{v}(x,y) := v(x,y) - L^{(2)}(x+y-f^*(x,y))$. Then (13) is equivalent to

$$\tilde{v}(x,y) = \alpha\inf_{b\in[[y-f^*(x,y)]^+,\kappa^{(2)}]}\Big\{K(f^*(x,y)+b) + \int_0^\infty\int_0^\infty \tilde{v}(f^*(x,y)-\xi, b-\eta)\psi_1(\xi)\psi_2(\eta)\mathrm{d}\xi\mathrm{d}\eta.\Big\} \tag{14}$$

We remember that $b^* \ge a^{**} \ge f^*(x,y)$ and $y \le b^*$ for $(x,y) \in S'$. If $y \ge f^*(x,y)$, then $[y-f^*(x,y)]^+ = y - f^*(x,y) \le b^* - f^*(x,y) \le \kappa^{(2)}$. If $y < f^*(x,y)$, then $[y-f^*(x,y)]^+ = 0 \le b^* - f^*(x,y) \le \kappa^{(2)}$. Hence,

$$\inf_{b\in\left[[y-f^*(x,y)]^+,\kappa^{(2)}\right]} K(f^*(x,y)+b) = K(b^*).$$

Let $\tilde{v}^*(x,y) := \frac{\alpha}{1-\alpha}K(b^*)$ for all $(x,y)\in S'$. Then

$$\begin{aligned}
&\alpha\inf_{b\in\left[[y-f^*(x,y)]^+,\kappa^{(2)}\right]}\Big\{K(f^*(x,y)+b) + \int_0^\infty\int_0^\infty \tilde{v}^*(f^*(x,y)-\xi, b-\eta)\psi_1(\xi)\psi_2(\eta)\mathrm{d}\xi\mathrm{d}\eta\Big\}\\
&= \alpha\inf_{b\in\left[[y-f^*(x,y)]^+,\kappa^{(2)}\right]}\Big\{K(f^*(x,y)+b) + \frac{\alpha}{1-\alpha}K(b^*)\Big\}\\
&= \alpha K(b^*) + \frac{\alpha^2}{1-\alpha}K(b^*) = \tilde{v}^*(x,y).
\end{aligned}$$

Hence, $\tilde{v}^*$ is the solution of (14) and v^* with $v^*(x,y) = L(x+y-f^*(x,y)) + \tilde{v}^*(x,y)$ is the solution of (13) for $(x,y)\in S'$. This implies that (6) is also satisfied.

The statement follows now from Theorem 1. □

5 Example

We assume that $L^{(1)}$ has form (7) and $L^{(2)}$ has form (8) with $h^{(1)} = 1$, $g^{(1)} = 11$, $h^{(2)} = 1$, $g^{(2)} = 10$. Let furthermore $c_1 = 0$, $c_2 = 5$, $\psi^{(1)}(\xi) = \mathbb{I}_{[0,\infty)}(\xi)\lambda e^{-\lambda\xi}$ with $\lambda = 0.1$, $\psi^{(2)}(\eta) = \frac{1}{8}\mathbb{I}_{[0,8]}(\eta)$, $\kappa^{(1)} = 30$, and $\kappa^{(2)} = 70$.

Then it holds that $G_1(a) = c_2 \cdot a + L^{(1)}(a)$, $G_2(a) = L^{(1)}(a)$ for $a \in \mathbb{R}$. We have

$$L^{(1)}(u) = \mathbb{I}_{[0,\infty]}(u)(g^{(1)} + h^{(1)}) \int_0^u (u-t)\psi^{(1)}(t)\mathrm{d}t + g^{(1)} \cdot \left(\frac{1}{\lambda} - u\right).$$

It follows for $u > 0$

$$L^{(1)}(u) = (g^{(1)} + h^{(1)})\frac{1}{\lambda}\mathrm{e}^{-\lambda u} + h^{(1)} \cdot \left(u - \frac{1}{\lambda}\right).$$

Hence,

$$\frac{\mathrm{d}}{\mathrm{d}u}L^{(1)}(u) = -(g^{(1)} + h^{(1)})\mathrm{e}^{-\lambda u} + h^{(1)}.$$

It follows that $L^{(1)}$ has the minimum in

$$a^{**} = \frac{1}{\lambda}\ln\left(1 + \frac{g^{(1)}}{h^{(1)}}\right) = 24.849.$$

Similarly, we get

$$G_1(u) = g^{(1)}\frac{1}{\lambda} - (g^{(1)} - c_2)u$$

for $u \leq 0$ and

$$G_1(u) = (g^{(1)} + h^{(1)})\frac{1}{\lambda}\mathrm{e}^{-\lambda u} - h^{(1)}\frac{1}{\lambda} + (h^{(1)} + c_2)u$$

for $u > 0$. Hence,

$$\frac{\mathrm{d}}{\mathrm{d}u}G_1(u) = -(g^{(1)} + h^{(1)})\mathrm{e}^{-\lambda u} + h^{(1)} + c_2.$$

It follows that G_1 has the minimum in

$$a^* = \frac{1}{\lambda}\ln\frac{g^{(1)} + h^{(1)}}{c_2 + h^{(1)}} = 10\ln 2 = 6.931.$$

From (8), it follows

$$\tilde{L}(z - \xi - \eta) = \begin{cases} h^{(2)} \cdot (z - a^{**} - \xi - \eta) & \text{for} \quad \xi \leq z - a^{**} - \eta \\ 0 & \text{for} \quad z - a^{**} - \eta < \xi \leq z - a^{*} - \eta \\ -g^{(2)} \cdot (z - a^{*} - \xi - \eta) & \text{for} \quad z - a^{*} - \eta < \xi \end{cases}.$$

Let

$$L_\eta(z) := \int_0^\infty \tilde{L}(z-\xi-\eta)\psi^{(1)}(\xi)\mathrm{d}\xi.$$

Then we get for $z \geq a^{**} + \eta$

$$\begin{aligned} L_\eta(z) &= h^{(2)} \int_0^{z-a^{**}-\eta} (z-a^{**}-\eta-\xi)\lambda \mathrm{e}^{-\lambda\xi}\mathrm{d}\xi \\ &\quad - g^{(2)} \int_{z-a^*-\eta}^{\infty} (z-a^*-\eta-\xi)\lambda \mathrm{e}^{-\lambda\xi}\mathrm{d}\xi \\ &= h^{(2)}\left(z-a^{**}-\eta-\frac{1}{\lambda}\right) + \frac{1}{\lambda}(h^{(2)}\mathrm{e}^{\lambda a^{**}} + g^{(2)}\mathrm{e}^{\lambda a^*})\mathrm{e}^{-\lambda(z-\eta)}. \end{aligned}$$

For $a^* + \eta \leq z < a^{**} + \eta$, it holds

$$L_\eta(z) = g^{(2)}\frac{1}{\lambda}\mathrm{e}^{\lambda(a^*-z+\eta)}.$$

For $z < a^* + \eta$, it follows

$$L_\eta(z) = g^{(2)} \cdot \left(a^* - z + \eta + \frac{1}{\lambda}\right).$$

L_η is convex for every fixed η, since $\tilde{L}$ is convex. Furthermore, $L_\eta(z)$ is decreasing for $z < a^{**} + \eta$. Therefore, we have to look for the minimum of $L_\eta(z)$ in $[a^{**} + \eta, \infty)$. It holds for $z \geq a^{**} + \eta$

$$\frac{\mathrm{d}}{\mathrm{d}z}L_\eta(z) = h^{(2)} - (h^{(2)}\mathrm{e}^{\lambda a^{**}} + g^{(2)}\mathrm{e}^{\lambda a^*})\mathrm{e}^{-\lambda(z-\eta)}.$$

It follows that L_η has its minimum in

$$z_\eta = \eta + \frac{1}{\lambda}\ln\left(\mathrm{e}^{\lambda a^{**}} + \frac{g^{(2)}}{h^{(2)}}\mathrm{e}^{\lambda a^*}\right).$$

Since $\eta \geq 0$, it follows for every η that $L_\eta(z)$ is (strong) decreasing in z for $z \leq \underline{b}$, where $\underline{b} := \frac{1}{\lambda}\ln\left(\mathrm{e}^{\lambda a^{**}} + \frac{g^{(2)}}{h^{(2)}}\mathrm{e}^{\lambda a^*}\right)$. It follows that $b^* \geq \underline{b} = 34.657$. Therefore, for the calculation of b^* it is sufficient to consider $K(z)$ for $z \geq \underline{b}$. For $z \geq \underline{b}$ we have $z - a^{**} - \eta \geq 34.65 - 24.85 - 8 > 0$. Hence,

$$\begin{aligned} K(z) &= h^{(2)}z - h^{(2)} \cdot \left(a^{**} - 4 - \frac{1}{\lambda}\right) \\ &\quad + \mathrm{e}^{-\lambda z} \cdot \frac{1}{\lambda}(h^{(2)}\mathrm{e}^{\lambda a^{**}} + g^{(2)}\mathrm{e}^{\lambda a^*}) \cdot \frac{1}{8}\int_0^8 \mathrm{e}^{\lambda\eta}\mathrm{d}\eta \end{aligned}$$

$$= h^{(2)}z - h^{(2)} \cdot \left(a^{**} - 4 - \frac{1}{\lambda}\right)$$
$$+\mathrm{e}^{-\lambda z} \cdot \frac{1}{\lambda}(h^{(2)}\mathrm{e}^{\lambda a^{**}} + g^{(2)}\mathrm{e}^{\lambda a^{*}}) \cdot \frac{1}{8\lambda}(e^{0.8} - 1)$$

From

$$K'(z) = h^{(2)} - \mathrm{e}^{-\lambda z} \cdot (h^{(2)}\mathrm{e}^{\lambda a^{**}} + g^{(2)}\mathrm{e}^{\lambda a^{*}}) \cdot \frac{1}{8\lambda}(e^{0.8} - 1)$$

we get

$$b^* = 38.922.$$

Hence, Assumption 2. is fulfilled. According to (9) and (11), we get that $(f^{*\infty}, g^{*\infty})$ with

$$f^*(x,y) = \begin{cases} x & \text{if } 24.849 < x \\ 24.849 & \text{if } x \le 24.849 < x + y \\ x + y & \text{if } 6.931 \le x + y < 24.849 \\ 6.931 & \text{if } x + y < 6.931 \end{cases}.$$

and

$$g^*(x,y,a) = 38.922 - a \quad \text{for} \quad \max\{x + y, a\} \le 38.922.$$

is a Nash equilibrium pair.

Acknowledgements The author thanks the anonymous referees for many valuable hints and comments.

References

1. Agin, N.: A min-max inventory model. Management Science **12**, 517–529 (1966)
2. Arrow, K.J., Harris, T., Marschak, J.: Optimal inventory policy. Econometrica **19**, 250–272 (1951)
3. Bertsekas, D.P., Shreve, S.E.: Stochastic optimal control: the discrete time case. Academic, New York (1978)
4. Cachon, G.P., Netessine, S.: Game theory in supply chain analysis. In: Simchi-Levi, Wu, Shen (eds.) Handbook of Quantitative Supply Chain Analysis: Modeling in the E-Business Era, pp. 13–66. Kluwer, Boston (2004)
5. Federgruen, A.: On N-person stochastic games with denumerable state space. Advances in Applied Probability **10**, 452–471 (1978)
6. Girlich, H.-J.: On the optimal inventory equation in the minimax case. In: Inventory in theory and practice (Budapest, 1984), *Stud. Prod. Engrg. Econom.*, vol. 6, pp. 505–512. Elsevier, Amsterdam (1986)
7. Girlich, H.-J., Küenle, H.-U.: On dynamic min-max decision models. In: Transactions of the ninth Prague conference on information theory, statistical decision functions, random processes, Vol. A (Prague, 1982), pp. 257–262. Reidel, Dordrecht (1983)
8. Iglehart, D.L.: Optimality of (s, S) policies in the infinite horizon dynamic inventory problem. Management Science **9**, 259–267 (1963)

9. Jagannathan, R.: A minimax ordering policy for the infinite stage dynamic inventory problem. Management Science **24**, 1138–1149 (1978)
10. Johnson, E.L.: On (s, S) policies. Management Science **15**, 80–101 (1968)
11. Kasugai, H., Kasegai, T.: Characteristics of dynamic maximin ordering policy. Journal of the Operations Research Society of Japan **3**, 11–26 (1960/1961)
12. Küenle, H.-U.: Stochastische Spiele und Entscheidungsmodelle. Teubner-Verlag, Leipzig (1986)
13. Küenle, H.-U.: On the optimality of (s, S)-strategies in a minimax inventory model with average cost criterion. Optimization **22**(1), 123–138 (1991)
14. Kurano, M.: Minimax strategies for average cost stochastic games with an application to inventory models. Journal of the Operations Research Society of Japan **30**(2), 232–247 (1987)
15. Leng, M., Parlar, M.: Game theoretic applications in supply chain management: a review. INFOR Information Systems and Operations Research **43**(3), 187–220 (2005)
16. Scarf, H.: The optimality of (S, s) policies in the dynamic inventory problem. In: Arrow, Karlin, Suppes (eds.) Mathematical methods in the social sciences, 1959, pp. 196–202. Stanford University Press, Stanford (1960)
17. Schurath, R.: On a two-player production-inventory problem. Tech. Rep. M-07/2000, Brandenburg University of Technology Cottbus, Department of Mathematics (2000)
18. Yang, S.: Ein Produktions-Lagerhaltungsproblem als nicht-kooperatives stochastisches Spiel. Master's thesis, Brandenburg University of Technology Cottbus, Faculty 1, (2003)

Time-Consistent Emission Reduction in a Dynamic Leader-Follower Game

Yulia Pavlova

Abstract In this paper, we search for multistage realization of international environmental agreements (IEAs). To analyze countries incentives and results of their interactions, we mathematically represent players' strategic preferences and apply game-theoretic approach to make predictions about their outcomes. Initial decision on emission reduction is determined by the Stackelberg equilibrium concept. We generalize Barrett's static 'emission' model to a dynamic framework and answer the question 'how fast should the emission reduction be?' It appears that sharper abatement is desirable in the early terms, which is similar to the conclusion of the *Stern* review. As discounting of the future payoffs becomes larger, more immediate reductions should be undertaken by the agreement parties. We show that without incentives from external organizations or governments, such depollution path can lead to a decline of the membership size.

1 Introduction

The increasing urgency of environmental problems has attracted the attention of scientists, politicians and, society at large. Environmental issues include the depletion of the ozone layer, the loss of biodiversity, and the effects of climate change. To protect the environment and ensure the stability of the ecosystem, a variety of international environmental agreements (IEAs) have been developed. These documents prescribe, among other things, pollution limits, the growth of industrial efficiency, and careful usage of resources. Although the parties to such agreements are often well intentioned, IEA negotiation, enforcement, and achievement face significant obstacles. At the root of such problems lies the realistic principle that countries act only in their own interest. It is obvious that the agreement requirements can be costly in an economic and social sense, and thus each country wants to avoid paying for

Y. Pavlova (✉)
Economic Research, MTT Agrifood Research Finland, Luutnantintie 13, Helsinki, 00410 Finland
e-mail: yulia.pavlova@mtt.fi

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_27,

environmental protection, even while recognizing that if every country does so the overall result will not be satisfactory.

To analyze countries' incentives and the results of their interactions, we mathematically represent players' preferences and apply a game-theoretic approach to make predictions about players' decisions and outcomes. In the last decades, an extensive body of literature on IEAs has appeared (see [6, 10] for excellent surveys of this literature).

The question of negotiation on agreement participation and the choice of commitments is well represented in the vast part of the literature (starting with [2, 5, 11]) and based on a static coalition formation game. The idea is to see when an individual country is indifferent between joining or leaving the coalition and to use this as the stability concept, also known as self-enforcement or internal/external stability [1]. Results in [2, 5, 11] suggest that the size of a stable coalition without additional enforcing mechanisms is typically very small (2, 3 or 4 players) [8], and that the size of a stable agreement depends on how the net benefits from abatement are distributed [7, 13]. These and other studies, e.g. [4, 9, 10, 23], highlight the role of the design of the treaty in making a large number of countries sign an IEA and have incentives to comply with its terms (i.e. through a combination of mechanisms of positive and negative incentives). Additionally, Diamantoudi and Srtzetakis [8] demonstrate some contradictions in the results between the 'emission' game (an individual strategy is an amount of pollution) and the 'abatement' game (a strategic choice is an amount of pollution reduction) and reconciles it by proposing a feasibility restraint on model parameters.

Another option to examine the problem of agreement enforcement is to include a realistic assumption of a time dimension, as most agreements usually state that a certain emission reduction has to be realized at a future point in time. Operating with the pollution stock (or the pollution flow) and abatement dynamics allows for consideration of a model with a variable at each period membership [19, 22]. The results of the dynamic models generalize the 'pessimistic' conclusion of the static ones: as the pollution stock approaches a steady state from above, the steady-state membership declines to size 2, while compliance with the agreement targets is subjected to free-riding. The presented paper contributes to the analysis of pollution flow and agreement membership dynamics and is close in its settings to [19, 22]. In contrast with the mentioned literature, which avoids specific remarks on the depollution process, our interest is directed to the question of how rapid the emission reduction should be.

The described problem is more generously discussed in economic studies utilizing cost-benefit modeling rather than the game-theoretic approach. While most of them call for the imposing of restraints on greenhouse gas emissions, the difficult question of how much and how fast is a debated issue. There are two main streams among economic models of climate change: the first one states that efficient or 'optimal' economic policies to slow climate change should involve modest rates of emission reduction in the near term, followed by sharp reductions in the medium and long term [3, 12, 14]; the second stream advocates sharp immediate reductions, pointing out that early depollution activities outweigh the costs due to unknown discounting of climate change and uncertain future catastrophes [20, 24]. This paper seeks a solution to the problem from a game-theoretic perspective. We construct a time-consistent path of emission reduction utilizing Bellman's principle of optimality, which guarantees that the initially chosen solution remains optimal at later stages of the process. It appears that the depollution process should start with substantial reductions in the initial stages followed by much lower reductions at later stages. Mathematically supporting the idea expressed in the *Stern* review [20], we nevertheless argue that such an approach is reasonable if the design of the treaty exogenously enforces its stability and deters possible free-riding.

This is organized as follows. In Sect. 2, we interpret a multilateral collaboration among countries as a coalition formation game [2, 5, 7, 8, 13]. We specify players' feasible strategies and correspondent payoffs as the difference between the benefits of total emission reduction and individual costs. Assuming that all the countries that have not joined the agreement make their decision on emission reduction after the analogous decision of the agreement signatories is announced allows us to determine initial abatement applying the Stackelberg equilibrium concept. The agreement membership is described as a stable coalition structure, determined by the principle of self-enforcement (conditions of internal/external stability) [1].

In Sect. 3, we examine depollution options and suggest a time-consistent scheme of stepwise emission reduction, which is analytically formulated in Theorem 1. The time-consistency term [17] requires that the remaining part of emission reduction, left after intermediate time interval, would be optimal (the Stackelberg equilibrium) in the dynamic game with pollution flow as a state variable [15, 16]. Such a depollution path is based on Bellmann's principle of optimality, and its construction mechanism is similar to approaches for allocation over time of players' payoffs as discussed in [18,21]. We show that even at low discounting, sharper emission reduction is desirable in the early terms, and as future payoffs become more discounted, sharper reductions should be undertaken by the agreement parties at early stages.

In Sect. 4, we look at how the initially stable coalition reacts to changes in the pollution flow state variable. We generalize the D'Aspremont et al. (1983) principle of internal–external stability [1] to the dynamic framework and carry out simulations for arbitrary values of the model to check whether coalition stability is violated. It appears that as the discount factor becomes smaller, free-riding incentives grow, and similarly to [19, 22], in the considered model the coalition size declines to 2 players. Thus, mathematically supporting the opinion expressed in the *Stern* review, we nevertheless should be aware that large abatement at an early stage leads to possible free-riding from the agreement. Additionally, we suggest that if a sharp depollution path is chosen, the treaty should be designed so as to maintain the dynamic stability of the coalition and deter free-riding.

2 The Static Model

Let $\mathcal{N}$ be a set of N identical players, countries of the world, each of which emits a pollutant that damages a shared environmental resource. Let S, $\varnothing \neq S \subseteq \mathcal{N}$ be a coalition of players that jointly intend to reduce their emissions. Players simultaneously and voluntarily decide to join the coalition S or act independently. Denote n ($n \leq N$) as the number of players that joined the agreement. They choose their abating strategies to maximize the net benefit to the coalition. The remaining $N-n$ players (free-riders) adjust their abatement levels noncooperatively, maximizing individual net benefit.

We will use the following notation:

- $F = \mathcal{N} \setminus S$ is the set of free-riders
- q_i^S and q_j^F are the individual abatement commitments chosen by players from coalition S and from the group of free-riders F, respectively
- $\mathbf{q}^S$ is a vector of all signatories' strategies and $\mathbf{q}^F$ is a vector of all free-riders' strategies
- $Q_S = \sum_{i \in S} q_i^S$ and $Q_F = \sum_{j \in F} q_j^F$ are the abatements that all signatories of the coalition S and all free-riders from the set F commit to reduce
- $Q = Q_S + Q_F$ is total abatement by all players under the IEA
- $\lambda = b/c$

We assume that the net benefit $\pi_i(\mathbf{q}^S, \mathbf{q}^F)$ of each player i depends on its own abatement commitments q_i and on the emission reduction Q undertaken by all players [5]. In particular,

$$\pi_i = B(Q) - C_i(q_i),$$

where

$$B(Q) = \frac{b}{N}(aQ - Q^2/2)$$

is the benefit function. The positive parameters a and b describe the initial level of global pollution and the slope of the marginal benefit function, respectively. The country's abatement cost function $C_i(q_i)$ is assumed to be given by

$$C_i(q_i) = \frac{1}{2}cq_i^2,$$

where c is a positive parameter equal to the slope of each country's marginal abatement cost curve.

We interpret agreement formation as a static Stackelberg game $\Gamma_0(S) = \langle \mathcal{N}, \{q_i^S, q_j^F\}, \{\pi_i^S, \pi_j^F\}, i \in S, j \in F \rangle$, where coalition S makes the first move (the leader), and the free-riders in F adopt the position of followers and rationally react to the strategy of the leader. We say that the strategies are feasible if $Q \leq a$. Supposing that a nonempty coalition S has formed, let us consider the problem of the existence of the Stackelberg equilibrium in the game $\Gamma_0(S)$.

Lemma 1. *In the two-level game $\Gamma_0(S)$, the Stackelberg equilibrium is unique and is given by the following strategies:*

$$q_i^S = a\gamma, \quad i \in S, \tag{1}$$

$$q_j^F = ag(1 - n\gamma), \quad j \in F, \tag{2}$$

where

$$g = \frac{\lambda(N-n)}{N+\lambda(N-n)}, \qquad \gamma = \frac{\lambda n(1-g)^2}{N+\lambda n^2(1-g)^2}.$$

Proof. We first construct the Stackelberg equilibrium. Assuming that coalition S has chosen some feasible strategies q_i^S, free-riders adjust their optimal abatement efforts q_j^F by maximizing the individual net benefit

$$\max_{q_j^F} \pi_j^F, \quad j \in F.$$

The first-order condition $\partial \pi_j^F / \partial q_j^F = 0$, $j \in F$, yields the solution

$$q_j^F = \frac{\lambda(a - Q_S)}{N + \lambda(N-n)}. \tag{3}$$

Since $\partial^2 \pi_j^F / \partial^2 q_j^F = -\frac{b}{N} - c < 0$, the solution given in (3) is maximum and determines the optimal strategies of the followers (free-riders' individual abatements). The total abatement of the free-riders is

$$Q_F = g(a - Q_S). \tag{4}$$

Expressions (3) and (4) can be interpreted as a rational (Nash equilibrium) reply of the followers to any of the leader's strategy. Taking the reaction functions (3) and (4) into account, the leader chooses its abatement Q_S by maximizing its aggregate net benefit

$$\max_{q_i^S} \sum_{i \in S} \pi_i^S. \tag{5}$$

The solution for the optimization problem (5) is

$$q_i^S = \frac{a\lambda n(1-g)^2}{N + (1-g)^2 \lambda n^2}, \quad i \in S. \tag{6}$$

To verify this, it is sufficient to check that

$$\frac{\partial \sum_{i \in S} \pi_i^S}{\partial q_i^S} = 0,$$

$$\frac{\partial^2 \sum_{i \in S} \pi_i^S}{\partial^2 q_i^S} = -\frac{b}{N} - c < 0$$

holds for the solution given by (6).

The aggregate coalitional abatement is given by

$$Q_S = \frac{a(1-g)^2 n^2 \lambda}{N + (1-g)^2 n^2 \lambda}.$$

The solution $(\mathbf{q}^S, \mathbf{q}^F)$ defined by (3) and (6) is feasible because

$$Q = Q_S + Q_F = Q_S + g(a - Q_S) = a\left[1 - \frac{(1-g)N}{N + (1-g)^2 n^2 \lambda}\right].$$

Here, we have

$$1 - g = \frac{1}{N + \lambda(N - n)} \in (0, 1]$$

and

$$\frac{N}{N + (1-g)^2 n^2 \lambda} \in (0, 1].$$

As a result, $Q = Q_S + Q_F \le a$. This completes the proof. □

The values $(\mathbf{q}^S, \mathbf{q}^F)$ are positive and finite because the parameters a, b, and c are assumed to be positive. It is useful to mention that additional restraints on the model parameters should be taken into account. In particular, in [8] the initial pollution flow a is obtained as unconstrained emission $\bar{E}$ in the game-theoretic 'emission' model with a zero damage function, and the chosen level of abatement q_i for each country is set below the country's uncontrolled emissions $\bar{e}_i$ and above zero. Such a natural assumption leads to constraints:[1] $0 \le \lambda = \frac{b}{c} \le \frac{4}{N-4}$.

We now assume that the stability of the coalition S is associated with a principle of self-enforcement, as introduced in [1].

Definition 1. A coalition S is self-enforcing in the game $\Gamma_0(S)$ if

$$\pi_i^S \ge \pi_i^{F \cup i}, \quad i \in S, \tag{7}$$

where $(\mathbf{q}^S, \mathbf{q}^F)$ is the Stackelberg equilibrium in the game $\Gamma_0(S)$ and $(\mathbf{q}^{S\setminus i}, \mathbf{q}^{F \cup i})$ is the Stackelberg equilibrium in the game $\Gamma_0(S \setminus i)$, and

$$\pi_j^{S \cup j} \le \pi_j^F, \quad j \in F, \tag{8}$$

where $(\mathbf{q}^S, \mathbf{q}^F)$ is the Stackelberg equilibrium in the game $\Gamma_0(S)$ and $(\mathbf{q}^{S \cup j}, \mathbf{q}^{F \setminus j})$ is the Stackelberg equilibrium in the game $\Gamma_0(S \cup j)$.

The inequality (7) guarantees the internal stability of the coalition, i.e. no member has a reason to leave the IEA. The external stability condition (8) guarantees that no nonmember has an incentive to join the coalition. In general, stability conditions

[1] It is useful to mention here that the 'abatement' model can appear to be not subordinate to the 'emission' game, as it is assumed in [8]. For instance, if the countries choose their initial pollution flow a according to other assessment methods for comparing technology costs with possible environmental damage, one should solve the conditional optimization problem to guarantee that abatement targets do not exceed present emissions. At the same time, since this paper is strictly theoretical, we shall further follow the proposed assumption and study the process of depollution and coalition dynamic stability given the outlined restraints on parameters c and b.

ensure that no player benefits from unilateral deviation. To identify the structure of the self-enforcing coalition S, we substitute the abatement strategies $(\mathbf{q}^S, \mathbf{q}^F)$, presented in (1) and (2), into the conditions of internal/external stability (7) and (8).

Within this framework, it is analytically shown that the stable coalition may consist of $2, 3$, or 4 players. One possible intuitive explanation can be that while the coalition size grows the positive externality for free-riders goes higher, and thus, when the coalition size exceeds a threshold of 4 members, signatories of such an agreement become worse-off than free-riders. Numerical simulations with heterogeneous players can be found in [7, 13]. It is shown that a solution of the system exists for a sufficiently large set of model parameters, and that a stable coalition solution is often not unique. This means that several coalitions can form. Additionally, we can conclude that redistribution of coalitional payoffs increases the number of possibly stable coalitions with heterogeneous membership, but the size of the membership is left unaffected.

3 Time-Consistent Depollution Scheme

Let us suppose that an agreement formed in the game $\Gamma_0(S)$ $(S \neq \emptyset)$ is in force during T time periods, $T > 1$ (given exogenously[2]). Assuming the agreement membership to be fixed, let us allocate the abatement commitments $(\mathbf{q}^S, \mathbf{q}^F)$, determined in (1)-(2), so that at each stage $t = 1, \dots, T$ each player reduces emissions by the amount $\Delta q_i^S(t)$ if $i \in S$, or $\Delta q_j^F(t)$ if $j \in F$.

Let us solve a discrete optimal control problem of duration T periods. At the beginning of each period $t, t = 1, \dots, T$, the environment (as a system) is specified by a state parameter $a(t)$ (current pollution flow). A control vector $\left(\Delta \mathbf{q}^S(t), \Delta \mathbf{q}^F(t)\right)$ changes the state of the system from $a(t-1)$ to $a(t)$ according to the relation

$$a(t) = a(t-1) - \Delta Q(t).$$

Here, $\Delta Q(t) = \Delta Q_S(t) + \Delta Q_F(t)$, $\Delta Q_S(t) = \sum_{i \in S} \Delta q_i^S(t)$, $\Delta Q_F(t) = \sum_{j \in F} \Delta q_j^F(t)$. Given the initial state $a(0) = a$, a sequence of controls $(\Delta \mathbf{q}^S(1), \Delta \mathbf{q}^F(1)), \dots, (\Delta \mathbf{q}^S(T), \Delta \mathbf{q}^F(T))$ and state vectors $a(1), \dots, a(T)$ are called feasible if for all $i \in S$, $j \in F$ and $t = 1, \dots, T$ they satisfy

$$\begin{aligned} &0 \leqslant \Delta q_i^S(t) \leqslant q_i^S, \quad 0 \leqslant \Delta q_j^F(t) \leqslant q_j^F, \\ &\sum_{t=1}^{T} \Delta q_i^S(t) = q_i^S, \quad \sum_{t=1}^{T} \Delta q_j^F(t) = q_j^F, \ 0 \leqslant a(t) \leqslant a. \end{aligned} \tag{9}$$

[2] Accomplishment of each agreement is typically a finite horizon process, and the fulfillment of its targets is checked at discrete times.

Among all feasible controls and trajectories, we seek a control and a corresponding trajectory that optimizes the objective function of a free-rider

$$\max \phi_j^F(\{a(t)\}_{t=0}^T, \{\Delta \mathbf{q}^S(t), \Delta \mathbf{q}^F(t)\}_{t=1}^T) = \sum_{t=1}^T \rho^{t-1} \pi_{jt}^F,$$

subject to $a(t) = a(t-1) - \Delta Q(t), t = 1, \ldots, T$. Parameter ρ is a discount factor.

A coalition S objective is

$$\max \phi^S(\{a(t)\}_{t=0}^T, \{\Delta \mathbf{q}^S(t), \Delta \mathbf{q}^F(t)\}_{t=1}^T) = \sum_{t=1}^T \rho^{t-1} \pi_t^S,$$

subject to $a(t) = a(t-1) - \Delta Q(t), t = 1, \ldots, T$.

Here,

$$\pi_t^S = \sum_{i \in S} \left(b(a(t)\Delta Q(t) - \frac{1}{2}\Delta Q^2(t)) - \frac{1}{2}c(\Delta q_i^S(t))^2 \right),$$

$$\pi_{jt}^F = b(a(t)\Delta Q(t) - \frac{1}{2}\Delta Q^2(t)) - \frac{1}{2}c(\Delta q_j^F(t))^2, \quad j \in F.$$

We introduce a finite-horizon dynamic game $\Gamma_t \left(S, \mathbf{q}^S(t), \mathbf{q}^F(t)\right)_{t=1}^T$, where $(\mathbf{q}^S(t), \mathbf{q}^F(t))$ is a vector of strategies such that $q_i^S(t) = \sum_{\tau=t}^T \Delta q_i^S(\tau)$, $i \in S$ and $q_j^F(t) = \sum_{\tau=t}^T \Delta q_j^F(\tau)$, $j \in F$, and payoffs of the coalition and the free-riders are $\pi^S(t) = \sum_{\tau=t}^T \rho^{\tau-1}\pi_\tau^S$ and $\pi_j^F(t) = \sum_{\tau=t}^T \rho^{\tau-1}\pi_{j\tau}^F$, respectively.

Definition 2. A scheme of stepwise emission reduction $\{\Delta q_i^S(t), \Delta q_j^F(t)\}_{t=1}^T$, $i \in S, j \in F$, is called time-consistent if it constitutes a Stackelberg equilibrium in the current game $\Gamma_t \left(S, \mathbf{q}^S(t), \mathbf{q}^F(t)\right)$.

Let us compute an optimal control by a recursive method, applying Bellman's principle of optimality and supposing that at any $t = 1, \ldots, T$ the players are acting as before: The coalition is a leader and the free-riders are followers [15, 16]. The design of a time-consistent scheme of stepwise emission reduction is based on a mechanism for the allocation over time of players' payoffs [18, 21].

Theorem 1. *Consider a coalition S and the corresponding Stackelberg equilibrium $\left(\mathbf{q}^S, \mathbf{q}^F\right)$ of the game $\Gamma_0(S)$. The scheme of stepwise abatement is time-consistent*

$$\Delta q_j^F(t) = \frac{\varphi}{N-n}a(t-1) - \frac{\varphi}{N-n}\Delta Q_S(t) + \frac{(1-\varphi)\rho}{1+\rho}\left(q_j^F - \sum_{\tau=1}^{t-1}\Delta q_j^F(\tau)\right),$$

$$\Delta q_i^S(t) = \frac{\psi}{n}a(t-1) - \frac{\psi}{n}\frac{\rho}{\rho+1}(Q_F - \sum_{\tau=1}^{t-1}\Delta Q_F(t))$$
$$+ \frac{(1-\psi)\rho}{1+\rho}\left(q_i^S - \sum_{\tau=1}^{t-1}\Delta q_i^S(t)\right),$$
$$t = 1,\ldots,T-1,$$

where

$$\psi = \frac{\lambda(1-\rho)(1-\varphi)^2 n^2}{\lambda(1-\rho)(1-\varphi)^2 n^2 + N(1+\rho)},$$
$$\varphi = \frac{\lambda(1-\rho)(N-n)}{\lambda(1-\rho)(N-n) + N(1+\rho)},$$

and

$$\Delta q_i^S(T) = q_i^S - \sum_{t=1}^{T-1}\Delta q_i^S(t), \quad \Delta q_j^F(T) = q_j^F - \sum_{t=1}^{T-1}\Delta q_j^F(t).$$

Proof. The functional equations of the free-riders at T, $i = 1,\ldots,K$, are

$$W_{jT}^F(\Delta\mathbf{q}^S(T), \Delta\mathbf{q}^F(T)) = \max_{\Delta q_j^F(T)} \pi_{jT}^F$$
$$= \max_{\Delta q_j^F(T)} b(a(T-1)\Delta Q(T) - \frac{1}{2}\Delta Q^2(T)) - \frac{1}{2}c(\Delta q_j^F(T))^2,$$

and the objective of the coalition S

$$W_T^S(\Delta\mathbf{q}^S(T), \Delta\mathbf{q}^F(T)) = \max_{\Delta\mathbf{q}^S(T)} \pi_T^S$$
$$= \max_{\Delta\mathbf{q}^S(T)} \sum_{i\in S} b(a(T-1)\Delta Q(T) - \frac{1}{2}\Delta Q^2(T)) - \frac{1}{2}c(\Delta q_i^S(T))^2,$$

Due to feasibility constraints (9), the control parameters $(\Delta\mathbf{q}^S(T), \Delta\mathbf{q}^F(T))$ are expressed through all the control parameters $(\Delta\mathbf{q}^S(t), \Delta\mathbf{q}^F(t))$ of the stages from 1 to $T-1$

$$\Delta q_i^S(T) = q_i^S - \sum_{t=1}^{T-1}\Delta q_i^S(t), \quad \Delta q_j^F(T) = q_j^F - \sum_{t=1}^{T-1}\Delta q_j^F(t).$$

Consider stage $t = T-1$. According to the dynamic program, the objective of the free-riders is

$$W^F_{j,T-1}(\Delta\mathbf{q}^S(T-1),\Delta\mathbf{q}^F(T-1)) = \max_{\Delta q^F_j(T-1)} \left(\pi^F_{j,T-1} + \rho W^F_{jT}(\Delta\mathbf{q}^S(T),\Delta\mathbf{q}^F(T))\right),$$

thus we solve

$$\max_{\Delta q^F_j(T-1)} \Bigg(b\left(a(T-2)\Delta Q(T-1) - \frac{1}{2}\Delta Q^2(T-1)\right) - \frac{1}{2}c\left(\Delta q^F_j(T-1)\right)^2$$
$$+\rho b\left((a(T-2)-\Delta Q(T-1))(Q-\sum_{t=1}^{T-1}\Delta Q(t)) - \frac{1}{2}(Q-\sum_{t=1}^{T-1}\Delta Q(t))^2\right)\Bigg)$$
$$-\frac{\rho}{2}c(q^F_i - \sum_{t=1}^{T-1}\Delta q^F_j(t))^2\Bigg).$$

The objective of the signatories is

$$W^S_{T-1}(\Delta\mathbf{q}^S(T-1),\Delta\mathbf{q}^F(T-1)) = \max_{\Delta\mathbf{q}^S(T-1)} \left(\sum_{i\in S}\pi^S_{T-1} + \rho W^S_T(\Delta\mathbf{q}^S(T),\Delta\mathbf{q}^F(T))\right),$$

and we need to solve

$$\max_{\Delta\mathbf{q}^S(T-1)} \sum_{i\in S}\Bigg(b\left(a(T-2)\Delta Q(T-1) - \frac{1}{2}\Delta Q^2(T-1)\right) - \frac{1}{2}c\left(\Delta q^S_i(T-1)\right)^2$$
$$+\rho b\left((a(T-2)-\Delta Q(T-1))\left(Q-\sum_{t=1}^{T-1}\Delta Q(t)\right) - \frac{1}{2}\left(Q-\sum_{t=1}^{T-1}\Delta Q(t)\right)^2\right)$$
$$-\frac{\rho}{2}c\ (q^S_i - \sum_{t=1}^{T-1}\Delta q^S_i(t))^2\Bigg).$$

The first-order condition for the maximization problem of the free-riders and the signatories at stage $T-1$ yields

$$\Delta q^F_j(T-1) = \frac{\varphi}{N-n}a(T-2) - \frac{\varphi}{N-n}\Delta Q_S(T-1) + \frac{(1-\varphi)\rho}{1+\rho}\left(q^F_j - \sum_{t=1}^{T-2}\Delta q^F_j(t)\right),$$

$$\Delta q^S_i(T-1) = \frac{\psi}{n}a(T-2) - \frac{\psi}{n}\frac{\rho}{\rho+1}\left(Q_F - \sum_{t=1}^{T-2}\Delta Q_F(t)\right) + \frac{(1-\psi)\rho}{1+\rho}\left(q^S_i - \sum_{t=1}^{T-2}\Delta q^S_i(t)\right).$$

Consider stage $t = T-2$. In a similar way as before, we determine the objective of the free-riders as

$$
\begin{aligned}
&W^F_{j,T-2}(\Delta\mathbf{q}^S(T-2),\Delta\mathbf{q}^F(T-2))\\
&\quad=\max_{\Delta q_j^F(T-2)}\left(\pi^F_{j,T-2}+\rho W^F_{j,T-1}(\Delta\mathbf{q}^S(T-1),\Delta\mathbf{q}^F(T-1))\right),
\end{aligned}
$$

which is analogous to

$$
\begin{aligned}
&\max_{\Delta q_j^F(T-2)}\Bigg(b\left(a(T-3)\Delta Q(T-2)-\frac{1}{2}\Delta Q^2(T-2)\right)-\frac{1}{2}c\left(\Delta q_j^F(T-2)\right)^2\\
&\quad+\rho b\left((a(T-3)-\Delta Q(T-2))\Delta Q(T-1)-\frac{1}{2}\Delta Q(T-1)^2\right)\\
&\quad-\rho\frac{1}{2}c\Delta q_j^F(T-1)^2\\
&\quad+\rho^2 b\Bigg((a(T-3)-\Delta Q(T-2)-\Delta Q(T-1))\left(Q-\sum_{t=1}^{T-1}\Delta Q(t)\right)\\
&\quad-\frac{1}{2}\left(Q-\sum_{t=1}^{T-1}\Delta Q(t)\right)^2\Bigg)-\frac{1}{2}\rho^2 c\left(q_j^F-\sum_{t=1}^{T-1}\Delta q_j^F(t)\right)^2\Bigg).
\end{aligned}
$$

The objective of the signatories is

$$
\begin{aligned}
&W^S_{T-2}(\Delta\mathbf{q}^S(T-2),\Delta\mathbf{q}^F(T-2))\\
&\quad=\max_{\Delta\mathbf{q}^S(T-2)}\left(\sum_{i\in S}\pi^S_{T-2}+\rho W^S_{T-1}(\Delta\mathbf{q}^S(T-1),\Delta\mathbf{q}^F(T-1))\right),
\end{aligned}
$$

and we need to solve

$$
\begin{aligned}
&\max_{\Delta q_i^S(T-2)}\sum_{i\in S}\Bigg(b\left(a(T-3)\Delta Q(T-2)-\frac{1}{2}\Delta Q^2(T-2)\right)-\frac{1}{2}c\left(\Delta q_i^S(T-2)\right)^2\\
&\quad+\rho b\left((a(T-3)-\Delta Q(T-2))\Delta Q(T-1)-\frac{1}{2}\Delta Q(T-1)^2\right)\\
&\quad-\rho\frac{1}{2}c\Delta q_i^S(T-1)^2\\
&\quad+\rho^2 b\Bigg((a(T-3)-\Delta Q(T-2)-\Delta Q(T-1))(Q-\sum_{t=1}^{T-1}\Delta Q(t))\\
&\quad-\frac{1}{2}\left(Q-\sum_{t=1}^{T-1}\Delta Q(t)\right)^2\Bigg)-\frac{1}{2}\rho^2 c\left(q_i^S-\sum_{t=1}^{T-1}\Delta q_i^S(t)\right)^2\Bigg).
\end{aligned}
$$

The first-order condition for the maximization problem of the free-riders and the signatories at stage $T-2$ yields

$$\Delta q_j^F(T-2) = \frac{\varphi}{N-n} a(T-3) - \frac{\varphi}{N-n} \Delta Q_S(T-2) + \frac{(1-\varphi)\rho}{1+\rho} \left(q_j^F - \sum_{t=1}^{T-3} \Delta q_j^F(t) \right),$$

$$\Delta q_i^S(T-2) = \frac{\psi}{n} a(T-3) - \frac{\psi}{n} \frac{\rho}{\rho+1} \left(Q_F - \sum_{t=1}^{T-3} \Delta Q_F(t) \right) + \frac{(1-\psi)\rho}{1+\rho} \left(q_i^S - \sum_{t=1}^{T-3} \Delta q_i^S(t) \right).$$

Carrying-on in such a manner, we determine that the solution of $W_{jt}^F(\Delta \mathbf{q}^S(t), \Delta \mathbf{q}^F(t))$ and $W_t^S(\Delta \mathbf{q}^S(t), \Delta \mathbf{q}^F(t))$, $t = 1, \ldots, T-1$, is (10). $\square$

According to such a scheme, a large emission reduction should be undertaken during initial stages and the following sequence of abating efforts will monotonically decrease. This observation follows from a case with $\rho = 1$ given time-horizon $T > 2$, where abatement decisions (10) are reduced to a decreasing geometric sequence with the scale factors equal to the initial allocation of the players' commitments $(\mathbf{q}^S, \mathbf{q}^F)$ and the common ratio of $1/2$. Additionally, as the value of future payoffs becomes more highly discounted, more immediate reductions should be undertaken by the signatories, while a rational reaction of the free-riders is to slow down emission reduction. An example below illustrates this property.

4 Dynamic Stability of Agreement

Analysis of the time consistency of an agreement is based on how the membership preference of players changes over time when the agreement is not fixed. Having accomplished its obligations for the moment t, a signatory i considers withdrawing from the coalition S, assuming that starting from that moment the abatement path will be continued along the restriction of the optimal solution of the game $\Gamma_0(S \setminus i)$. Upon such a scenario, the condition of a self-enforcing coalition receives the following presentation.

Definition 3. A self-enforcing coalition S is dynamically stable against a certain abatement scheme if for every moment $t = 1, \ldots, T$ the following conditions simultaneously hold:

(1) Dynamic internal stability

$$\pi_i^S(t) \geq \pi_i^{F \cup i}(t), \quad \forall i \in S,$$

where the strategies $(\mathbf{q}^S(t), \mathbf{q}^F(t))$ are a restriction over the period $[t, T]$ of the Stackelberg solution in the game $\Gamma_0(S)$ and $(\mathbf{q}^{S \setminus i}(t), \mathbf{q}^{F \cup i}(t))$ is the Stackelberg solution in the game $\Gamma_t(S \setminus i)$,

(2) Dynamic external stability

$$\pi_j^F(t) \geq \pi_j^{S\cup i}(t), \quad \forall j \in F,$$

where $(\mathbf{q}^{S\cup i}(t), \mathbf{q}^{F\setminus i}(t))$ is the Stackelberg solution in the game $\Gamma_0(S \cup i)$.

While the state variable $a(t)$ changes over time, parameters c, b are constant. The payoff for each player is given as follows:

$$\pi_i^S(t) = \frac{b}{N}\left(a(t)Q(t) - \frac{1}{2}Q^2(t)\right) - \frac{1}{2}c\left(q_i^S(t)\right)^2,$$
$$\pi_i^{F\cup i}(t) = \frac{b}{N}\left(a(t)Q^{-i}(t) - \frac{1}{2}(Q^{-i}(t))^2\right) - \frac{1}{2}c\left(q_i^{F\cup i}(t)\right)^2,$$

where $Q^{-i}(t)$ is the total emission to be reduced over $[t, T]$, if coalition S is abandoned by one of its members;

$$\pi_j^F(t) = \frac{b}{N}\left(a(t)Q(t) - \frac{1}{2}Q^2(t)\right) - \frac{1}{2}c\left(q_i^F(t)\right)^2,$$
$$\pi_j^{S\cup i}(t) = \frac{b}{N}\left(a(t)Q^{+i}(t) - \frac{1}{2}(Q^{+i}(t))^2\right) - \frac{1}{2}c\left(q_i^{S\cup i}(t)\right)^2,$$

where $Q^{+i}(t)$ is total emission to be reduced over $[t, T]$, if coalition S is accessed by one player. In the following example, we illustrate the depollution process (10) for different values of the discount factor and consider the evolution of a stable coalition with the state variable changing over time.

5 Example

We illustrate the results presented in Sects. 2, 3, and 4 by considering a numerical example that leads to a stable coalition of size $n = 3$. We assume $N = 10$, $a = 100$, $c = 0.2$, $b = 0.1$, which results in $\lambda = 0.5$, satisfying the interior solution constraint. In this game, the Stackelberg equilibrium is $(\mathbf{q}^S, \mathbf{q}^F) = (6.601, 2.97)$.

In Figs. 1 and 2, depollution schemes of the signatories and the free-riders are plotted against three different discount factors (see Tables 1 and 2). The dotted line represents discount $\rho = 0.85$, the dashed line represents discount $\rho = 0.95$, and the solid line describes the situation without discounting $\rho = 1$. Three emission reduction paths are plotted together to highlight the optimal behavior of the players in different situation. Both Figs. 1 and 2 suggest substantial pollution abatement in the early stage followed by a slower but steady decline in emissions. On the other hand, a change in the discount factor affects the optimal behavior of the coalition and the free-riders differently: Fig. 1 shows that if the discounting of future payoffs increase, coalition abatement targets should be more rapidly realized, while the free-riders, in

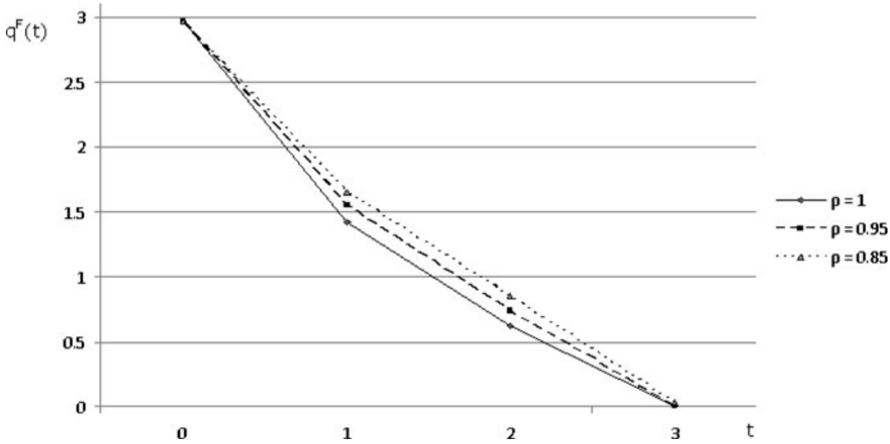


Fig. 1 Depollution path of coalition members

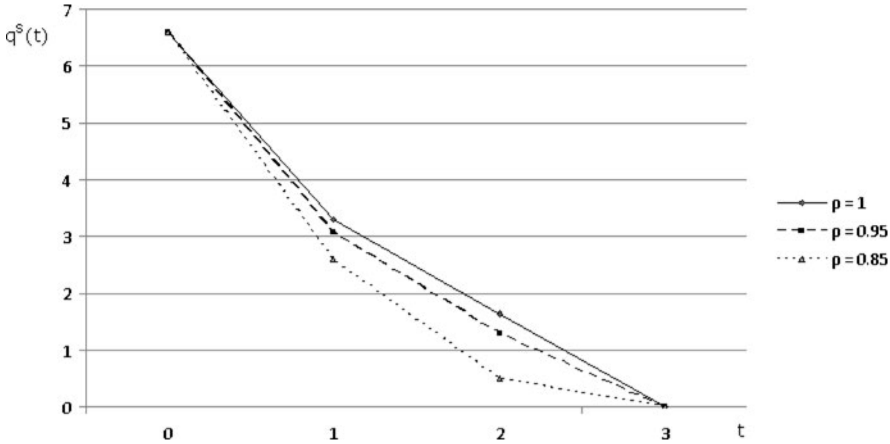


Fig. 2 Depollution path of free-riders

Table 1 Depollution scheme of coalition members

Discount factor	$\Delta q_i^F(1)$	$\Delta q_i^F(2)$	$\Delta q_i^F(3)$
$\rho = 1$	3.3	1.65	1.65
$\rho = 0.95$	3.523	1.785	1.292
$\rho = 0.85$	4.009	2.087	0.505

contrast to the signatories, should lean toward gradual emission reduction, as shown in Fig. 2.

The following question is whether such a depollution approach meets the conditions of dynamic stability of the agreement. It appears that with higher discounting, free-riding incentives become stronger. At $t = 1$, the internal and external dynamic

Table 2 Depollution scheme of free-riders

Discount factor	$\Delta q_i^F(1)$	$\Delta q_i^F(2)$	$\Delta q_i^F(3)$
$\rho = 1$	1.485	0.743	0.743
$\rho = 0.95$	1.462	0.749	0.759
$\rho = 0.85$	1.382	0.765	0.823

stability fails for the coalition of 3 players, e.g. if $\rho = 1$, then $\pi_i^S(1) - \pi_i^{F\cup i}(1) = -3.907$; if $\rho = 0.95$, then $\pi_i^S(1) - \pi_i^{F\cup i}(1) = -4.13$; if $\rho = 0.85$, then $\pi_i^S(1) - \pi_i^{F\cup i}(1) = -7.434$.

Similar to [19] and [22], coalition size declines to 2 players in all three cases. Thus, mathematically supporting the opinion expressed in the *Stern* review, we nevertheless should be aware that large abatement at an early stage leads to possible free-riding from the agreement.

6 Conclusion

In this paper, we present an analysis of the multistage realization of IEAs, which with increasing urgency have been attracting the attention of scientists, politicians and society during the past decades. It appears that the enforcement and achievement of the agreement targets face significant obstacles due to the economic self-interest of the participating countries, as expressed by their unwillingness to pay for environmental protection. To analyze countries' incentives and the results of their interactions, the strategic preferences of the countries received mathematical interpretation and agreement participation was studied from a game-theoretic point of view.

To contribute to the existing literature on environmental agreements, we include a realistic assumption of the time dimension. In contrast to the mentioned literature, which refrains from specific remarks on the depollution process, our interest is directed to the question of how rapid the emission reduction should be. An initial decision on emission reduction is determined by the Stackelberg equilibrium concept. We generalize Barrett's static 'emission' model to a dynamic framework with the pollution flow as a state variable, and analytically construct a time-consistent depollution path utilizing Bellman's principle of optimality. It appears that the depollution process should start with substantial reductions in the initial stages followed by much lower reductions at later stages. We mathematically support the idea expressed in the *Stern* review, and show that as discounting rises, more immediate reductions should be undertaken by the agreement parties.

The presented numerical simulations with different scales of the discount factor demonstrate that free-riding incentives become stronger as discounting grows, and in the considered model the coalition size declines to 2 players. We suggest that additional incentives from external organizations or governments are needed to deter free-riding when following the proposed time-consistent scheme.

Acknowledgements Funding from MTT Agrifood Research Finland and the COMAS Graduate School of the University of Jyväskylä is gratefully acknowledged.

References

1. d'Aspremont, C., Jacquemin, A., Weymark, J.A., Godszeweiz, J.: On the Stability of Collusive Price Leadership. Can. J. Econ. **16**, 17–25 (1983)
2. Carraro, C., Siniscalco, D.: Strategies for the International Protection of the Environment. J. Pub. Econ. **52**, 309–328 (1993)
3. Cline, W.R.: The Economics of Global Warming, Washington, Institute for International Economics (1992)
4. Bahn, O., Breton, M., Sbragia, L., Zaccour, G.: Stability of International Environmental Agreements: An Illustration with Asymmetrical Countries. Int. Trans. Oper. Res. **16**(3), 307–324 (2009)
5. Barrett, S.: Self-Enforcing International Environmental Agreements. Ox. Econ. Pap. **46**, 878–894 (1994)
6. Barrett, S.: Environment and Statecraft: The Strategy of Environmental Treaty-Making, Oxford University Press, Oxford (2003)
7. Dementieva, M., Pavlova, Y.: Emission Trading as an Expanding Mechanism of International Environmental Agreement. In: Neittaanmäki P., Périaux J., Tuovinen T. (eds.) Evolutionary and Deterministic Methods for Design, Optimization and Control, pp. 457–462, CIMNE, Barcelona, Spain (2007)
8. Diamantoudi, E., Sartzetakis, E.S.: Stable International Environmental Agreements: An Analytical Approach. J. Pub. Econ. Th. **8**(2), 247–263 (2006)
9. Eyckmans, J., Tulkens, H.: Simulating Coalitionally Stable Burden Sharing Agreements for the Climate Change Problem. Res. En. Econ. **25**(4), 299–327 (2003)
10. Finus, M.: Game Theory and International Environmental Cooperation (New Horizons in Environmental Economics), Edward Elgar Publishing Ltd., UK (2001)
11. Hoel, M.: International Environmental Conventions: The Case of Uniform Reductions of Emissions. Env. Res. Econ. **2**(3), 141–159 (1992)
12. Kelly, D.L., Kolstad, C.D.: Integrated Assessment Models For Climate Change Control. In: Folmer H., Tietenberg T. (eds.) International Yearbook of Environmental and Resource Economics 1999/2000: A Survey of Current Issues, Cheltenham, Edward Elgar, UK (1999)
13. McGinty, M.: International Environmental Agreements among Asymmetric Nations. Ox. Econ. Pap. **59**(1), 45–62 (2006)
14. Nordhaus, W.D.: A Review of the Stern Review on the Economics of Climate Change. J. Econ. Lit. **45**(3), 686–702 (2007)
15. Pavlova, Y.: A Multistage Game-Theoretic Model of International Environmental Agreements, (Russian). Vestnik of St. Peterburg University **10**(3), 85–97 (2008)
16. Pavlova, Y.: Multistage Coalition Formation Game of a Self-Enforcing International Environmental Agreement. Dissertation, Jyväskylä Studies in Computing 94 (2008)
17. Petrosjan, L.: Stability of Solutions in N-person Differential Games, (Russian). Bulletin of Leningrad University **19**, 46–52 (1977)
18. Petrosjan, L., Zaccour, G.: A Multistage Supergame of Downstream Pollution. Ann. Int. Soc. Dyn. Games **5**, 387–404 (2000)
19. Rubio, S., Ulph, A.: An Infinite-Horizon Model of Dynamic Membership of International Environmental Agreements. J. Env. Econ. Man. **54**, 296–310 (2007)
20. Stern, N.: The Economics of Climate Change: The Stern Review, Cambridge University Press, Cambridge, UK (2007)
21. Zakharov, V.: Regularization and Time-Consistency of Dynamic Solutions in Hierarchical Differential Games, (Russian). Vestnik of Leningrad University **1,2**(8), 45–53 (1988)
22. de Zeeuw, A.: Dynamic Effects on the Stability of International Environmental Agreements. J. Env. Econ. Man. **55**(2), 163–174 (2008)
23. Weikard, H.-P., Dellink, R.: Sticks and Carrots for the Design of International Climate Agreements with Renegotiations. Nota di lavoro 26.2008 (2008)
24. Weitzman, M.L.: A Review of the Stern Review on the Economics of Climate Change. J. Econ. Lit. **45**(3), 703–724 (2007)

Fish Wars with Changing Area for a Fishery

Anna Rettieva

Abstract In this paper, a discrete-time game model related to a bioresource management problem (fish catching) is considered. The center (referee) shares a reservoir between the competitors, the players (countries) harvest the fish stock on their territory. We consider power population's growth function and logarithmic players' profits. Both cases of finite and infinite planning horizon are investigated. The Nash and cooperative equilibria are derived. We investigate a new type of equilibrium – cooperative incentive equilibrium. Hence, the center punishes players for a deviation from the cooperative equilibrium by changing the harvesting territory. A numerical illustration is carried out and results are compared.

1 Introduction

We consider here a discrete-time game model related to a bioresource management problem (fish catching). The center (referee) shares a reservoir between the competitors. The players (countries) which harvest the fish stock are the participants of this game.

In this paper, we derive Nash and cooperative equilibria for an infinite planning horizon. We find the infinite–time problem solution by finding solutions of n–step games and letting the number of steps tend to infinity [3, 7].

The main objective of this work is to apply the approach with reserved territory, developed by the authors [4–6], to a bioresource sharing problem for two players. We investigate a new type of equilibrium – cooperative incentive equilibrium. This concept was introduced in [2] as a natural extension of Osborn's work [8] about cartel stability. In our model, the center punishes players for a deviation from the cooperative equilibrium by changing the harvesting territory.

A. Rettieva (✉)
Institute of Applied Mathematical Research, Karelian Research Center of RAS,
Pushkinskaya str. 11, Petrozavodsk, 185910, Russia
e-mail: annaret@krc.karelia.ru

M. Breton and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games 11, DOI 10.1007/978-0-8176-8089-3_28,

2 Discrete-Time Model

Let us divide the water area into two parts: s and $1-s$, where two countries exploit the fish stock. The center (referee) shares the reservoir. The players (countries) which exploit the fish stock on their territory are the participants of this game.

The fish population grows according to the biological rule

$$x_{t+1} = (\varepsilon x_t)^{\alpha}, \quad x_0 = x, \tag{1}$$

where $x_t \geq 0$ is the size of the population at a time t; $0 < \varepsilon < 1$ is the natural death rate, and $0 < \alpha < 1$ is the natural birth rate.

We suppose that the utility function of country i is logarithmic:

$$I_1 = \ln((1-s)x_t u_t^1), \quad I_2 = \ln(s x_t u_t^2),$$

where u_t^1, $u_t^2 \geq 0$ represent countries' fishing efforts at time t. We do not put a restriction $u_t^i \leq 1$, because the situation $u_t^i > 1$ is possible when country i enters the territory of the opponent.

We consider the problem of maximizing the sum of discounted utilities for two players:

$$J_1 = \sum_{t=0}^{n} \beta_1^t \ln((1-s)x_t u_t^1), \quad J_2 = \sum_{t=0}^{n} \beta_2^t \ln(s x_t u_t^2), \tag{2}$$

where $0 < \beta_i < 1$ is the discount factor for country i, $i = 1, 2$.

We also consider the problem with an infinite time horizon where:

$$\bar{J}_1 = \sum_{t=0}^{\infty} \beta_1^t \ln((1-s)x_t u_t^1), \quad \bar{J}_2 = \sum_{t=0}^{\infty} \beta_2^t \ln(s x_t u_t^2). \tag{3}$$

The next theorem characterizes Nash equilibrium and steady-state solution of the game.

Theorem 1. *The Nash equilibrium for the n-step game (1), (2) is*

$$u_1^n = \frac{\varepsilon\alpha\beta_2 \sum\limits_{j=0}^{n-1} (\alpha\beta_2)^j}{\left(\sum\limits_{j=0}^{n} (\alpha\beta_1)^j \sum\limits_{j=0}^{n} (\alpha\beta_2)^j - 1\right)(1-s)},$$

$$u_2^n = \frac{\varepsilon\alpha\beta_1 \sum\limits_{j=0}^{n-1} (\alpha\beta_1)^j}{\left(\sum\limits_{j=0}^{n} (\alpha\beta_1)^j \sum\limits_{j=0}^{n} (\alpha\beta_2)^j - 1\right)s}.$$

The Nash equilibrium for the infinite-time problem (1), (3) is

$$\bar{u}_1 = \frac{\varepsilon\beta_2(1-\alpha\beta_1)}{(\beta_1+\beta_2-\alpha\beta_1\beta_2)(1-s)},$$
$$\bar{u}_2 = \frac{\varepsilon\beta_1(1-\alpha\beta_2)}{(\beta_1+\beta_2-\alpha\beta_1\beta_2)s}.$$

The dynamics of the fishery becomes

$$x_t = \left(\frac{\varepsilon\alpha\beta_1\beta_2}{\beta_1+\beta_2-\alpha\beta_1\beta_2}\right)^{\sum_{j=1}^{t}\alpha^j} x_0^{\alpha^t},$$

and the steady-state size of the population under Nash policies is

$$\bar{x} = \left(\frac{\varepsilon\alpha\beta_1\beta_2}{\beta_1+\beta_2-\alpha\beta_1\beta_2}\right)^{\frac{\alpha}{1-\alpha}}.$$

Proof. We use the approach of Levhari and Mirman [3]. We first consider the problem in context of the finite horizon. Assume that if there were no future period, the countries would get the remaining fish in the ratio $(1-s):s$. Let the initial size of the population be x.

We first consider one-step game. Suppose that the players' fishing efforts are u_1 and u_2. These strategies must satisfy the inequality

$$u_1(1-s)+u_2 s<\varepsilon. \tag{4}$$

Then the first player's profit is

$$\begin{aligned} H_1^1(u_1,u_2) &= \ln((1-s)xu_1)+\beta_1\ln((1-s)(\varepsilon x-(1-s)xu_1-sxu_2)^\alpha) \\ &= \ln((1-s)xu_1)+\alpha\beta_1\ln(\varepsilon x-(1-s)xu_1-sxu_2)+\beta_1\ln(1-s), \end{aligned}$$

and, similarly, the second player' profit is

$$H_2^1(u_1,u_2) = \ln(sxu_2)+\alpha\beta_2\ln(\varepsilon x-(1-s)xu_1-sxu_2)+\beta_2\ln s.$$

Functions $H_1^1(u_1,u_2)$ and $H_2^1(u_1,u_2)$ are convex, so the Nash equilibrium exists and can be found from the system of equations $\partial H_1^1/\partial u_1=0$, $\partial H_2^1/\partial u_2=0$. The Nash equilibrium for one-step game is

$$u_1^1 = \frac{\varepsilon\alpha\beta_2}{(\alpha\beta_1+\alpha\beta_2+\alpha^2\beta_1\beta_2)(1-s)} = \frac{\varepsilon\alpha\beta_2}{((1+\alpha\beta_1)(1+\alpha\beta_2)-1)(1-s)},$$
$$u_2^1 = \frac{\varepsilon\alpha\beta_1}{(\alpha\beta_1+\alpha\beta_2+\alpha^2\beta_1\beta_2)s} = \frac{\varepsilon\alpha\beta_1}{((1+\alpha\beta_1)(1+\alpha\beta_2)-1)s}.$$

Now we verify that (4) is satisfied

$$u_1^1(1-s)+u_2^1 s=\frac{\varepsilon(\alpha\beta_1+\alpha\beta_2)}{\alpha\beta_1+\alpha\beta_2+\alpha^2\beta_1\beta_2}<\varepsilon.$$

The remaining stock is given by

$$\varepsilon x-(1-s)xu_1^1-sxu_2^1=\frac{\varepsilon\alpha^2\beta_1\beta_2}{(1+\alpha\beta_1)(1+\alpha\beta_2)-1}x.$$

Now we obtain Player 1's profit in this one-step game

$$\begin{aligned}H_1^1(u_1^1,u_2^1)&=\ln((1-s)xu_1^1)+\alpha\beta_1\ln(\varepsilon x-(1-s)xu_1^1-sxu_2^1)+\beta_1\ln(1-s)=\\&=(1+\alpha\beta_1)\ln x+A_1^1+\beta_1\ln(1-s),\end{aligned}$$

where A_1^1 is independent of x and is equal to

$$A_1^1=\ln\left[\frac{(\varepsilon\alpha\beta_2)(\varepsilon\alpha^2\beta_1\beta_2)^{\alpha\beta_1}}{(\alpha\beta_1+\alpha\beta_2+\alpha^2\beta_1\beta_2)^{1+\alpha\beta_1}}\right].$$

Similarly for Player 2

$$\begin{aligned}H_2^1(u_1^1,u_2^1)&=\ln(sxu_2^1)+\alpha\beta_2\ln(\varepsilon x-(1-s)xu_1^1-sxu_2^1)+\beta_2\ln s=\\&=(1+\alpha\beta_2)\ln x+A_2^1+\beta_2\ln s,\end{aligned}$$

where

$$A_2^1=\ln\left[\frac{(\varepsilon\alpha\beta_1)(\varepsilon\alpha^2\beta_1\beta_2)^{\alpha\beta_2}}{(\alpha\beta_1+\alpha\beta_2+\alpha^2\beta_1\beta_2)^{1+\alpha\beta_2}}\right].$$

Hence, the objective function of Player 1 for the two-step game is

$$\begin{aligned}H_1^2(u_1,u_2)=&\ln((1-s)xu_1)+\alpha\beta_1(1+\alpha\beta_1)\ln(\varepsilon x-(1-s)xu_1-sxu_2)+\\&+\alpha\beta_1(A_1^1+\beta_1\ln(1-s)),\end{aligned}$$

and for Player 2

$$\begin{aligned}H_2^2(u_1,u_2)=&\ln(sxu_2)+\alpha\beta_2(1+\alpha\beta_2)\ln(\varepsilon x-(1-s)xu_1-sxu_2)+\\&+\alpha\beta_2(A_2^1+\beta_2\ln s).\end{aligned}$$

Analogously, these functions are convex and we find the Nash equilibrium for the two-step game

$$u_1^2=\frac{\varepsilon\alpha\beta_2(1+\alpha\beta_2)}{((1+\alpha\beta_1+\alpha^2\beta_1^2)(1+\alpha\beta_2+\alpha^2\beta_2^2)-1)(1-s)},$$

$$u_2^2 = \frac{\varepsilon\alpha\beta_1(1+\alpha\beta_1)}{((1+\alpha\beta_1+\alpha^2\beta_1^2)(1+\alpha\beta_2+\alpha^2\beta_2^2)-1)s}.$$

The restriction (4) is satisfied

$$u_1^2(1-s)+u_2^2 s = \frac{\varepsilon\alpha(\beta_1+\beta_2+\alpha\beta_1^2+\alpha\beta_2^2)}{(1+\alpha\beta_1+\alpha^2\beta_1^2)(1+\alpha\beta_2+\alpha^2\beta_2^2)-1} < \varepsilon.$$

Analogously, we compute Player 1's profit in the two-step game

$$\begin{aligned} H_1^2(u_1^2,u_2^2) &= \ln((1-s)xu_1^2)+\alpha\beta_1(1+\alpha\beta_1)\ln(\varepsilon x-(1-s)xu_1^2-sxu_2^2)+ \\ &+\alpha\beta_1(A_1^1+\beta_1\ln(1-s)) = (1+\alpha\beta_1+\alpha^2\beta_1^2)\ln x + A_1^2 + \\ &+\alpha\beta_1(A_1^1+\beta_1\ln(1-s)), \end{aligned}$$

where A_1^2 is equal to

$$A_1^2 = \ln\left[\frac{(\varepsilon(\alpha\beta_2+\alpha^2\beta_2^2))^{1+\alpha\beta_1+\alpha^2\beta_1^2}(\alpha\beta_1+\alpha^2\beta_1^2)^{\alpha\beta_1+\alpha^2\beta_1^2}}{((1+\alpha\beta_1+\alpha^2\beta_1^2)(1+\alpha\beta_2+\alpha^2\beta_2^2)-1)^{1+\alpha\beta_1+\alpha^2\beta_1^2}}\right].$$

Similarly for Player 2

$$\begin{aligned} H_2^1(u_1^1,u_2^1) &= \ln(sxu_2^2)+\alpha\beta_2(1+\alpha\beta_2)\ln(\varepsilon x-(1-s)xu_1^2-sxu_2^2)+ \\ &+\alpha\beta_2(A_2^1+\beta_2\ln s) = (1+\alpha\beta_2+\alpha^2\beta_2^2)\ln x + A_2^2+\alpha\beta_2(A_2^1+\beta_2\ln s), \end{aligned}$$

where

$$A_2^2 = \ln\left[\frac{(\varepsilon(\alpha\beta_1+\alpha^2\beta_1^2))^{1+\alpha\beta_2+\alpha^2\beta_2^2}(\alpha\beta_2+\alpha^2\beta_2^2)^{\alpha\beta_2+\alpha^2\beta_2^2}}{((1+\alpha\beta_1+\alpha^2\beta_1^2)(1+\alpha\beta_2+\alpha^2\beta_2^2)-1)^{1+\alpha\beta_2+\alpha^2\beta_2^2}}\right].$$

The process can be repeated for the n-step game and we find the Nash equilibrium in the form, that is given in the statement. And the restriction (4) is satisfied because

$$\begin{aligned} \sum_{j=0}^{n}(\alpha\beta_1)^j\sum_{j=0}^{n}(\alpha\beta_2)^j-1 &= \sum_{j=1}^{n}(\alpha\beta_1)^j+\sum_{j=1}^{n}(\alpha\beta_2)^j+\sum_{j=1}^{n}(\alpha\beta_1)^j\sum_{j=1}^{n}(\alpha\beta_2)^j > \\ &> \sum_{j=1}^{n}(\alpha\beta_1)^j+\sum_{j=1}^{n}(\alpha\beta_2)^j = \alpha\beta_1\sum_{j=0}^{n-1}(\alpha\beta_1)^j+\alpha\beta_2\sum_{j=0}^{n-1}(\alpha\beta_2)^j. \end{aligned}$$

We can also compute Player 1's profit in the n-step game

$$H_1^n(u_1^n,u_2^n) = \sum_{j=0}^{n}(\alpha\beta_1)^j\ln x+\sum_{k=1}^{n}A_1^k(\alpha\beta_1)^{n-k}+(\alpha\beta_1)^{n-1}\beta_1\ln(1-s),$$

where A_1^k is equal to

$$A_1^k = \ln\left[\left(\frac{\varepsilon\sum_{j=1}^{k}(\alpha\beta_2)^j}{\sum_{j=0}^{k}(\alpha\beta_1)^j\sum_{j=0}^{k}(\alpha\beta_2)^j - 1}\right)^{\sum_{j=0}^{k}(\alpha\beta_1)^j}\left(\sum_{j=1}^{k}(\alpha\beta_1)^j\right)^{\sum_{j=1}^{k}(\alpha\beta_1)^j}\right].$$

And for Player 2

$$H_2^n(u_1^n, u_2^n) = \sum_{j=0}^{n}(\alpha\beta_2)^j \ln x + \sum_{k=1}^{n} A_2^k(\alpha\beta_2)^{n-k} + (\alpha\beta_2)^{n-1}\beta_2 \ln s,$$

where

$$A_2^k = \ln\left[\left(\frac{\varepsilon\sum_{j=1}^{k}(\alpha\beta_1)^j}{\sum_{j=0}^{k}(\alpha\beta_1)^j\sum_{j=0}^{k}(\alpha\beta_2)^j - 1}\right)^{\sum_{j=0}^{k}(\alpha\beta_2)^j}\left(\sum_{j=1}^{k}(\alpha\beta_2)^j\right)^{\sum_{j=1}^{k}(\alpha\beta_2)^j}\right].$$

Letting the horizon tend to infinity yields the Nash equilibrium for the infinite-horizon problem (1), (3) (see [7]):

$$\begin{aligned}
\bar{u}_1 &= \frac{\varepsilon\beta_2(1-\alpha\beta_1)}{(\beta_1+\beta_2-\alpha\beta_1\beta_2)(1-s)},\\
\bar{u}_2 &= \frac{\varepsilon\beta_1(1-\alpha\beta_2)}{(\beta_1+\beta_2-\alpha\beta_1\beta_2)s},\\
\varepsilon x - \bar{u}_1(1-s)x - \bar{u}_2 s x &= \frac{\varepsilon\alpha\beta_1\beta_2}{\beta_1+\beta_2-\alpha\beta_1\beta_2}x,
\end{aligned}$$

and inequality (4) is satisfied

$$\bar{u}_1(1-s) + \bar{u}_2 s = \frac{\varepsilon(\beta_1+\beta_2-2\alpha\beta_1\beta_2)}{\beta_1+\beta_2-\alpha\beta_1\beta_2} < \varepsilon.$$

Now we find the steady-state solution. Under Nash equilibrium strategies, the dynamics of the fish population is

$$x_{t+1} = (\varepsilon x_t - \bar{u}_1(1-s)x_t - \bar{u}_2 s x_t)^\alpha = \left(\frac{\varepsilon\alpha\beta_1\beta_2}{\beta_1+\beta_2-\alpha\beta_1\beta_2}\right)^{\sum_{j=1}^{t}\alpha^j} x_0^{\alpha^t}.$$

And the steady-state is (as $n \to \infty$)

$$\bar{x} = \left(\frac{\varepsilon \alpha \beta_1 \beta_2}{\beta_1 + \beta_2 - \alpha \beta_1 \beta_2} \right)^{\frac{\alpha}{1-\alpha}}.$$

□

We find the cooperative equilibrium using the same approach of transfer from finite to infinite resource management problem. Here, the players wish to maximize the discounted sum of their total utilities on finite and infinite time horizon:

$$\max_{u_t^1, u_t^2} \sum_{t=0}^{n} \rho^t (\mu_1 \ln((1-s)x_t u_t^1) + \mu_2 \ln(s x_t u_t^2)), \tag{5}$$

$$\max_{u_t^1, u_t^2} \sum_{t=0}^{\infty} \rho^t (\mu_1 \ln((1-s)x_t u_t^1) + \mu_2 \ln(s x_t u_t^2)), \tag{6}$$

where $0 < \rho < 1$ is the common discount factor, and $0 < \mu_1, \mu_2 < 1$ are the weighting coefficients ($\mu_1 + \mu_2 = 1$).

Theorem 2. *The cooperative equilibrium for n-step game (1), (5) is*

$$u_1^{dn} = \frac{\varepsilon \mu_1}{\sum\limits_{j=0}^{n} (\alpha\rho)^j (1-s)}, \qquad u_2^{dn} = \frac{\varepsilon \mu_2}{\sum\limits_{j=0}^{n} (\alpha\rho)^j s}.$$

The cooperative equilibrium for the problem (1), (6) is

$$u_1^d = \frac{\varepsilon \mu_1 (1-\alpha\rho)}{1-s}, \qquad u_2^d = \frac{\varepsilon \mu_2 (1-\alpha\rho)}{s}.$$

The dynamics of the fishery becomes

$$x_t^d = (\varepsilon\alpha\rho)^{\sum\limits_{j=1}^{t} \alpha^j} x_0^{\alpha^t},$$

and the steady-state size of the population under cooperative policies is

$$\bar{x}^d = (\varepsilon\alpha\rho)^{\frac{\alpha}{1-\alpha}}.$$

Proof. The proof of this statement is analogous to the one of Theorem 1. We remark that the restriction (4) is satisfied for finite:

$$u_1^{dn}(1-s) + u_2^{dn} s = \frac{\varepsilon}{\sum\limits_{j=0}^{n} (\alpha\rho)^j} < \varepsilon,$$

and for infinite horizon:

$$u_1^d (1-s) + u_2^d s = \varepsilon(1-\alpha\rho) < \varepsilon.$$

We can also compute the equilibrium profit in the n-step game

$$H^n(u_1^{dn}, u_2^{dn}) = \sum_{j=0}^{n} (\alpha\rho)^j \ln x + \sum_{k=1}^{n} A_k (\alpha\rho)^{n-k},$$

where

$$A_k = \ln\left[\left(\frac{\varepsilon}{\sum\limits_{j=0}^{k} (\alpha\rho)^j}\right)^{\sum\limits_{j=0}^{k} (\alpha\rho)^j} \left(\sum_{j=1}^{k} (\alpha\rho)^j\right)^{\sum\limits_{j=1}^{k} (\alpha\rho)^j} \mu_1^{\mu_1} \mu_2^{\mu_2}\right].$$

□

Here, we consider the problem of maintaining the agreement achieved at the beginning of the game and punishing the player who deviates. We consider the incentive equilibrium concept as the method of ensuring cooperation, where the center (referee) punishes the players for any deviation.

Denote s^d the water area sharing rule under cooperation.

Following [2] we assume that the strategy of Player i is a causal mapping $\gamma_i : D_j \to D_i$ $(u_i \in D_i = [0, \infty))$, $i, j = 1, 2, i \neq j$.

Definition 1. A strategy pair (γ_1, γ_2) is called the incentive equilibrium for the problem (1), (5), (6) if

$$\begin{aligned} u_1^d = \gamma_1(u_2^d), \quad u_2^d &= \gamma_2(u_1^d), \\ J_1(u_1^d, u_2^d) &\geq J_1(u_1, \gamma_2(u_1)), \quad \forall u_1 \in D_1, \\ J_2(u_1^d, u_2^d) &\geq J_2(\gamma_1(u_2), u_2), \quad \forall u_2 \in D_2. \end{aligned}$$

We now find the incentive equilibrium for a n-step game and for an infinite-time game.

Theorem 3. *The incentive equilibrium for problem (1), (5) is*

$$\gamma_1^n(u_2) = \frac{\varepsilon\mu_1}{\sum\limits_{j=0}^{n} (\alpha\rho)^j (1 - s_2^{*n})}, \qquad \gamma_2^n(u_1) = \frac{\varepsilon\mu_2}{\sum\limits_{j=0}^{n} (\alpha\rho)^j s_1^{*n}},$$

where

$$s_2^{*n} = s^d - \frac{s^d}{u_2^{dn}}(u_2 - u_2^{dn}), \quad s_1^{*n} = s^d + \frac{1 - s^d}{u_1^{dn}}(u_1 - u_1^{dn}).$$

The incentive equilibrium for problem (1), (6) is

$$\gamma_1(u_2) = \frac{\varepsilon\mu_1(1 - \alpha\rho)}{1 - s_2^*}, \qquad \gamma_2(u_1) = \frac{\varepsilon\mu_2(1 - \alpha\rho)}{s_1^*},$$

where

$$s_2^* = s^d - \frac{s^d}{u_2^d}(u_2 - u_2^d), \;\; s_1^* = s^d + \frac{1 - s^d}{u_1^d}(u_1 - u_1^d).$$

Proof. We assume that the punishment for a deviation from the equilibrium point is done by the center, and not by the players themselves, as it was in [2]. Assume that, if the first player deviates the center increases s^d, but if the second player deviates, it decreases s^d.

Consider the second player's deviation for the n-step game

$$u_2^n = u_2^{dn} + \Delta.$$

We give the center's punishment strategy the following form

$$s^{*n} = s^d - \eta^n (u_2^n - u_2^{dn}).$$

Then the first player's punishment strategy is

$$\gamma_1^n(u_2^n) = \frac{\varepsilon \mu_1}{\sum_{j=0}^{n} (\alpha\rho)^j (1 - s^{*n})}.$$

Next we should determine the coefficient η^n. For that we solve the problem for Player 2 assuming that Player 1 uses the punishment strategy, so

$$\sum_{t=0}^{n} \rho^t \ln(s^{*t} x u_2^t) \to \max,$$

where

$$s^{*t} = s^d - \eta^t (u_2^t - u_2^{dt}),$$

and the dynamics of the fishery becomes

$$x_{t+1} = (\varepsilon x_t - \gamma_1^t(u_2^t)(1 - s^{*t})x_t - s^{*t} x_t u_2^t)^\alpha.$$

For η^n to be the incentive equilibrium, the solution of this optimization problem must be achieved on u_2^{dn}.

To solve that problem, we use the same approach as above. So, we first consider one-step game

$$u_2 = u_2^{d1} + \Delta, \;\; s^{*1} = s^d - \eta^1 (u_2 - u_2^{d1}),$$
$$\gamma_1^1(u_2) = \frac{\varepsilon \mu_1}{(1 + \alpha\rho)(1 - s^{*1})}.$$

The objective function for Player 2 becomes

$$\max_{u_2 \geq 0} \{\ln(s^{*1} x u_2) + \alpha\rho \ln(\varepsilon x - \gamma_1^1(u_2)(1 - s^{*1})x - s^{*1} x u_2)\}$$

and the solution of this problem u_2 must be equal to

$$u_2^{d1} = \frac{\varepsilon\mu_2}{(1+\alpha\rho)s^d},$$

from which we find that

$$\eta^1 = \frac{s^d}{u_2^{d1}}.$$

Continuing this process for a n-step game yields

$$\eta^n = \frac{s^d}{u_2^{dn}}.$$

Here, inequality (4) is again satisfied, because

$$\gamma_1^t(u_2^t)(1-s^{*t}) + s^{*t}u_2^t = \frac{\varepsilon\mu_1}{\sum\limits_{j=0}^{t}(\alpha\rho)^j} + u_2^t s^d\left(2 - \frac{u_2^t}{u_2^{dt}}\right) =$$

$$= \frac{\varepsilon\mu_1}{\sum\limits_{j=0}^{t}(\alpha\rho)^j} + 2u_2^t s^d - \frac{(u_2^t)^2(s^d)^2\sum\limits_{j=0}^{t}(\alpha\rho)^j}{\varepsilon\mu_2}$$

$$= \frac{\varepsilon}{\sum\limits_{j=0}^{t}(\alpha\rho)^j} - \frac{\left(u_2^t s^d\sum\limits_{j=0}^{t}(\alpha\rho)^j - \varepsilon\mu_2\right)^2}{\varepsilon\mu_2\sum\limits_{j=0}^{t}(\alpha\rho)^j} < \varepsilon.$$

Analogously, for Player 1 we find the center's strategy in the following form

$$s^{*n} = s^d + \theta^n(u_1^n - u_1^{dn}).$$

Then the second player's punishment strategy is

$$\gamma_2^n(u_1^n) = \frac{\varepsilon\mu_2}{\sum\limits_{j=0}^{n}(\alpha\rho)^j s^{*n}}.$$

To determine the coefficient θ^n, we solve the problem for Player 1 assuming that Player 2 uses the punishment strategy, so

$$\sum_{t=0}^{n} \rho^t \ln((1 - s^{*t})x_t u_1^t) \to \max,$$

where

$$s^{*t} = s^d + \theta^t (u_1^t - u_1^{d\,t}),$$

and the dynamics of the fishery becomes

$$x_{t+1} = (\varepsilon x_t - u_1^t (1 - s^{*t})x_t - \gamma_2^t (u_1^t) s^{*t} x_t)^\alpha,$$

yielding

$$\theta^n = \frac{1 - s^d}{u_1^{dn}}.$$

The incentive equilibrium of problem (1), (6) is obtained by letting the horizon tend to infinity.

□

3 Modelling

We consider a numerical illustration with the set of parameter values

$$\rho = 0.1, \quad \varepsilon = 0.8, \quad s^d = 0.5,$$
$$\alpha = 0.3, \ \mu_1 = 0.55, \ \mu_2 = 0.45.$$

The initial size of the population is $x = 0.8$. The number of steps is 12. The time when the second player deviates is $n_0 = 5$ and the size of the deviation is $\Delta = 0.1$.

The steady-state size of the population under cooperative policies is 0.2022, which is larger than the steady-state size of the population under Nash policies, 0.1512.

The figures show results in the case of cooperative behavior (dark line) and in the case of deviation (light line). Figure 1 presents population's dynamics. Figure 2 shows water area sharing (s), where s decreases from 0.5 to 0.43. Figures 3 and 4 present the first and the second player's catch, respectively ($v_1^t = (1 - s)x_t u_t^1$, $v_2^t = s x_t u_t^2$). We notice that the first player's catch increases slightly on time interval [5,7]. But the second player's catch decreases on time interval [5,6] and increases when he returns to cooperation.

The first player's gain is $0.806\ 10^{-7}$ and the second player's loss is $0.1263\ 10^{-6}$.

If we compare the players profits as a percentage of the cooperative equilibrium result, then:

Player 1 wins 0.1% when the second player deviates, loses 8,6% when players use Nash equilibrium strategies.

Fig. 1 Size of population x_t

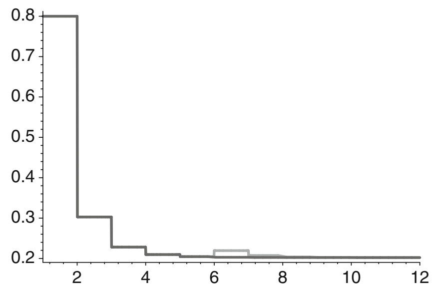


Fig. 2 Water area sharing s

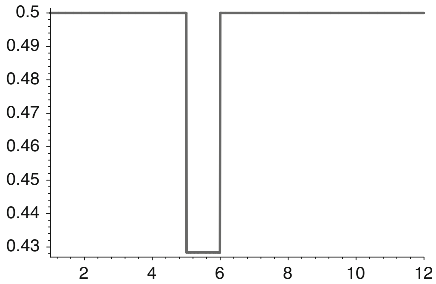


Fig. 3 The catch of Player 1

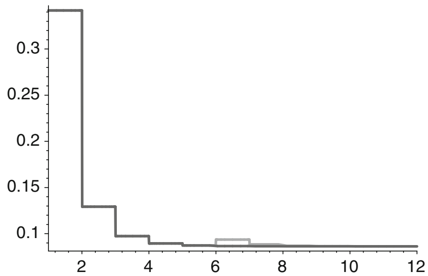
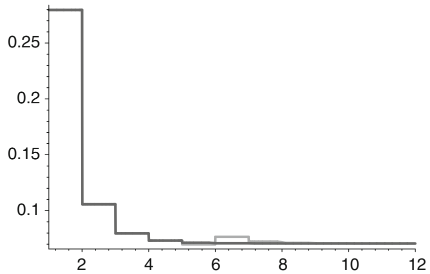


Fig. 4 The catch of Player 2



Player 2 loses 0.008% when he deviates, wins 7,3% when players use Nash equilibrium strategies.

Comparing the sum of the players' profits, then under deviation of the second player they lose 0.001%, while under Nash equilibrium they lose 0.1% .

4 Conclusions

We introduce the discrete-time bioresource management problem. The Nash, cooperative and incentive equilibria are derived for power population's growth function and logarithmic players' profits. Here, in contrast to traditional schemes for maintaining cooperation, we investigate the case when the center punishes deviating players.

The center's strategy here is the territory sharing. A player who breaks the agreement is punished by gradually decreasing the harvesting territory. This scheme can be easily realized in practice.

The approach can be applied for various biological growth rules for fish population in the reservoir. In particular, the authors investigated the cases when the growth rate depends on nature-conservative measures of one of the players, that is:

$$x_{t+1} = (\varepsilon s x_t)^{\alpha}, \quad 0 < \alpha < 1$$

and

$$x_{t+1} = (\varepsilon x_t)^{\alpha s}, \quad 0 < \alpha < 1.$$

Acknowledgements The research was supported by the Russian Fund for Basic Research, projects 10-01-00089-a and 08-01-98801-r-sever-a.

References

1. Basar, T., Olsder G.J.: Dynamic noncooperative game theory. Academic, New York (1982)
2. Ehtamo, H., Hamalainen, R.P.: A cooperative incentive equilibrium for a resource management problem. J. Econ. Dyn. Control **17**, 659–678 (1993)
3. Levhari, D., Mirman, L.J.: The great fish war: an example using a dynamic Cournot-Nash solution. Bell J. Econ. **11**(1), 322–334 (1980)
4. Mazalov, V.V., Rettieva, A.N.: A fishery game model with migration: reserved territory approach. Int. J. Math. Game Theor. Algebra **15**(3), 243–256 (2005)
5. Mazalov, V.V., Rettieva, A.N.: Nash equilibrium in bioresource management problem. Math. Model. **18**(5), 73–90 (2006) (in russian)
6. Mazalov, V.V., Rettieva, A.N.: Bioresource management problem with changing area for fishery. Game Theor. Appl. **13**, 101–110 (2008)
7. Nowak, A.: A note on an equilibrium in the great fish war game. Econ. Bull. **17**(2), 1–10 (2006)
8. Osborn, D.K.: Cartel problems. Am. Econ. Rev. **66**, 835–844 (1976)