Annals of the International Society of Dynamic Games
Volume 8

Series Editor
Tamer Başar

Editorial Board

Annals of the International Society of Dynamic Games

Advances in Dynamic Games

Applications to Economics, Management Science, Engineering, and Environmental Management

Alain Haurie
Shigeo Muto
Leon A. Petrosjan
T.E.S. Raghavan
Editors

Birkhäuser
Boston • Basel • Berlin

Alain Haurie
ORDECSYS
Place de l'Etrier 4
Chêne-Bougeries CH-1224
Switzerland

Shigeo Muto
Tokyo Institute of Technology
Graduate School of Decision Science
and Technology
2-12-1 Ookayama
Tokyo 152-8552
Japan

Leon A. Petrosjan
State University of St. Petersburg
Faculty of Applied Mathematics
and Control Processes
198504 St. Petersburg
Russia

T.E.S. Raghavan
University of Illinois at Chicago
Department of Mathematics, Statistics,
and Computer Science
Chicago, IL 60607-7045
USA

Mathematics Subject Classification: 49N70, 49N75, 49N90, 91-XX, 91Axx, 91A05, 91A06, 91A10, 91A12, 91A13, 91A15, 91A18, 91A20, 91A22, 91A23, 91A24, 91A25, 91A30, 91A35, 91A40, 91A43, 91A44, 91A46, 91A50, 91A55, 91A60, 91A65, 91A60, 91A80

Library of Congress Control Number: 2006924926

ISBN-10 0-8176-4500-4 e-ISBN: 0-8176-4501-2
ISBN-13 978-0-8176-4500-7

Printed on acid-free paper.

Printed in the United States of America. (TXQ/EB)

9 8 7 6 5 4 3 2 1

www.birkhauser.com

Contents

Part VI Management Science Applications

Preface

This edited volume contains a selection of chapters that are an outgrowth of the 10th International Symposium on Dynamic Games organized by the ISDG, in St. Petersburg, Russia, in July 2002, with a few additional contributed chapters. These fully reviewed chapters present an outlook of the current development of the theory of dynamic games and its applications to various domains, in particular, energy and the environment, economics, and management science.

It is now well established that the paradigms of dynamic games play an important role in the development of multi-agent models in engineering, economics, and management science. The ability of the concepts in providing insight for difficult real-life decision problems stems from their capacity to encompass situations with uncertainty, incomplete information, fluctuating coalition structure, and coupled constraints imposed on the strategies of all the players. The twenty-one chapters grouped in the six parts that constitute this volume cover these different aspects of modern dynamic game theory.

While the foundations of discrete dynamic game theory were laid down by L. Shapley and D. Blackwell, the continuous case in the form of differential games was initiated by R. Isaacs in the USA and by L.S. Pontryagin in Russia. Almost all of their efforts were directed towards the study of zero-sum two-person games. Since the seminal papers by J. Case and A.W. Starr and Yu chi Ho on nonzero-sum, m-player differential games, numerous applications to management science, and economics have been presented, particularly in the previous volumes of the Annals of the ISDG series. In this volume several problems pertaining to pursuit-evasion, marketing, finance, climate and environmental economics, resource exploitation, and auditing and tax evasions are addressed using dynamic game models of various sorts. The volume also includes some chapters on cooperative games, which are increasingly drawing dynamic approaches to their classic solutions. The contributions are grouped in six parts.

Part I deals with *zero-sum game theory* and contains three chapters.

Rosenberg, Solan, and Vieille consider zero-sum stochastic games where players do not observe the progress of the game fully but only in bits and pieces. The key theorem is to show the existence of max-min and min-max for all zero-sum stochastic games with imperfect monitoring which extends an earlier work of Coulomb for imperfect monitoring zero-sum stochastic games with all but one state absorbing. The zero-sum assumption is crucial, as a slight deviation from imperfection for nonzero-sum stochastic games may fail to give any equilibrium payoff.

Kumkov and Patsko consider a zero-sum linear differential game with fixed terminal time. Given two sets A and B, the set $A \overset{*}{-} B = \{x : B + x \subseteq A\}$. The set A is swept by the set B if $A = B + (A \overset{*}{-} B)$. For any constant c, the level set of a real-valued function f is the set $M_c = \{x : f(x) \leq c\}$. The function f is said to have the level sweeping property if for any $c < d$ the set M_d is swept by the set M_c. The main result of the work is the proof of the fact that if the payoff function depends on just two components of the phase vector and also possesses the level sweeping property, then so does the value function for the linear differential games. Such an inheritance of the level sweeping property by the value function is specific for the case when the payoff function depends on two components of the phase vector under very general regularity conditions for such differential games. If it depends on three or more components of the phase vector, this inheritance is generally lost. The latter is shown by a counterexample.

Serov considers the game of the generalized shortest path problem, where the task is to transit optimally from a fixed point through a system of intermediate sets in R^d to a fixed destination point (or set), with the condition that no point of the sets visited is visited again. The (combinatorial) cost to minimize is assumed additive or bottleneck. It becomes a zero-sum dynamic game when, say, player II decides to choose the order of the sets to visit or to terminate at each stage while player I chooses a point of the set to be visited. For this multistage game problem both open-loop and feedback settings are suggested. The feedback problem is posed in the class of feedback strategies and these strategies can change a route during a motion, depending on current moves of the opponent. They provide, in general, a strictly better value of the problem compared to the open-loop minimax setting. The author shows how to construct an optimal feedback minimax strategy, and some heuristics are also proposed.

Part II is concerned with *pursuit-evasion games.* It contains three chapters that address the now classical problems of pursuit-evasion and the related domain of zero-sum differential games.

Shinar and Glizer take up the problem of pursuit-evasion where the pursuer's information about the evader's lateral acceleration is delayed. The pursuer needs to estimate this, essentially based on the available measurement history during this delay period. This approach reduces the uncertainty set of the pursuer, due to the estimation delay, by considering in addition to the current (pure feedback) measurements also the available measurement history during the period of the estimation delay. The reduced uncertainty set is computed by solving two auxiliary optimization problems. By using the center of the new uncertainty set's convex hull as a new state variable, the original game is transformed into a nonlinear delayed dynamics game with perfect information for both players. The solution of this new game is obtained in pure strategies for the pursuer and mixed ones for the evader. The value of this game (the guaranteed miss distance) is substantially less than the one obtained in previous works by using only the current measurements.

Chikrii extends the well-known Pontryagin sufficiency conditions of capture in ordinary differential games to game problems for systems with fractional derivatives of arbitrary order. These are games with evolution described by equations with fractional derivatives, one for each player. Here player II strives for the state variables to get closer to the state variables of the opponent to within a specified distance. Player I wants them as far away as possible.

Petrov and Vagin study the problem of group pursuits and evasion. They provide necessary conditions for the capture of several evaders in a group pursuit problem, where all evaders use the same control. Necessary conditions for capture in such a group pursuit problem are also obtained for a special case called "soft" capture.

Part III contains four chapters concerned with *games of coalitions*.

Petrosjan introduces the notion of an n-person cooperative stochastic game. The Shapley value for cooperative n-person transferable utility (TU) games and the value in stationary strategies for zero-sum two person discounted stochastic games are central to the study of cooperative games and dynamic games, respectively. Petrosjan combines these two distinct value concepts: Players in a given coalition S might join together and play the game as though the rest are against them and treat this game as a zero-sum stochastic game. This game has a value and this induced value can be taken to be the worth of the coalition. This in turn determines its Shapley value. The subtree of cooperative trajectories maximizing the sum of expected players' payoffs is defined and the solution of the game along the paths of this tree is investigated. The new notion of cooperative payoff distribution procedure (CPDP) is introduced to show that the resultant Shapley value constructed is time consistent.

Funaki and Yamato study the problem of consistency for the core as the solution of a cooperative game with transferable utility. While one can have many reduced games using the same solution, not every one is capable of inherit-ing the solution for the reduced game. While an earlier characterization used a specific reduced game and used three axioms on this reduced game that characterized the core of a balanced game, here the authors provide a new set of four axioms to characterize the core via a new reduced game.

Scheffran proposes a framework for analyzing the interaction between individual players (actors) and collective players (coalitions) who mutually adapt the allocation of investment to their values and each other's decisions. The dynamic process of coalition formation is described by a coupled evolutionary game of allocation controls. An application to analyzing the management of energy and carbon emissions is discussed in some detail.

Raghavan and Sudhölter survey the geometric and algorithmic aspects of solutions, like the core and the nucleolus. While a cooperative TU game in general is defined by its characteristic function, some special classes of cooperative TU games are easily determined by a small amount of data. Assignment games belong to this category. They are models of two-sided markets. Players on one side, called sellers, supply exactly one unit of some indivisible good, say, a house in exchange for money, with players from the other side, called buyers. Each buyer has a demand for exactly one house. When a transaction between a seller and a buyer takes place, a certain profit accrues. The worth of any coalition is given by the total profit of an optimal assignment of players within the coalition. Therefore, the characteristic function is fully determined by the profits of the buyer-seller pairs. They study conditions for the core to be a stable set in the sense of von Neumann and Morgenstern. Another solution, called the modiclus, is a spin-off from the notion of nucleolus, taking into account the jealousy between coalitions for any given payoff configuration. Unlike the nucleolus, the modiclus is not in general a core element even when the core exists. However, from a computational point of view, when the modiclus is in the core one has hopes of computing the modiclus efficiently.

Part IV, which discusses *new concepts of equilibrium*, is composed of three chapters proposing new interpretations of the interdependence between different members of a social group.

Vasin shows that "natural" evolution of behavior in repeated games in human populations is a very unstable process which may be easily manipulated by outside forces. Any feasible and individually rational payoff of the game may be converted into a globally stable outcome by arbitrary small perturbation of the payoff functions in the repeated game. He shows that this result also holds for a trembling-hand perturbation of the game, and proves a new version of the Folk theorem for this case.

Morgan and Patrone study the Stackelberg equilibrium, which can be thought of as a subgame perfect equilibium in an extensive game when the leader makes the first move in the leader-follower games. The problems here often stem from the nonuniqueness of the best reply and the differences between weak and strong Stackelberg equilibria. In particular, the authors study Tikhonov regularization methods for seeking the so called weak or strong lower Stackelberg equilibria.

Hidano and Muto study the philosophical problem: What triggers two selfish individuals to unite? They treat this as a two-stage decision process, namely to unite or not to unite as the first-stage decision and in case one of the players prefers not to unite, then their aims are simply to maximize their individual utility levels. If they both choose to unite, then they both equally enjoy the maximum utility given by $\max\{h(s,t) \mid (s,t) \in R\}$. The authors discuss subgame perfect equilibria of the game in order to make clear under what conditions different selves unite. Since they have symmetric utility, symmetric equilibria are natural topics to be studied.

In **Part V** four chapters address original *applications to energy/environment economics.*

Kryazhimskii, Nikonov, and Minullin develop an explicit algorithm to approximate Nash equilibria for an earlier model of one of the authors on nonzero-sum games of timing for building gas pipelines. In the energy market, say for building gas pipelines, if there are no competitors, then any monopolistic pipeline builder can concentrate on the right time to stop construction and venture into supply that will maximize the rate of return on the initial investments on constructing pipelines. If there are no competitors, then when to start commercialization (stop construction) is often an optimal control problem. But with a competitor it is no longer an optimal control problem. These are games of timing, specifically nonzero-sum games. The Nash equilibria for these games of timing can be approximated. The key point in their approach is based on the observation that the best response commercialization times for all players concentrate at two time points, one corresponding to a fast investment policy and another corresponding to a slow investment policy.

McKelvey and Golubtsov study the dynamic fishery harvesting game in a stochastic environment, in order to examine the implications of incomplete and asymmetric information. The main emphasis is on a split-stream version of the game: At the beginning of each harvest season the initial fish stock (or recruitment) divides into two streams, each one accessible to harvest by just one of the two competing fishing fleets. The fleets simultaneously harvest down their streams, achieving net seasonal payoffs for the catch. After harvest, the residual sub-stocks reunite to form the brood-stock for the subsequent generation. The strength of this subsequent generation is determined by a specified stock-recruitment relation, and the cycle repeats. In this cyclic process, both natural environmental factors (stream-split proportions and stock-recruitment relation) and economic factors (harvest costs and benefits) incorporate Markovian stochastic elements. The implications of alternative knowledge structures are explored through dynamic programming and simulation.

Haurie, Moresino, and Viguier propose a two-level hierarchical differential game to represent the possible negotiations of GHG emissions quotas among different nations in a post-Kyoto era. The players are countries with growing economies. The quotas are determined noncooperatively but are globally constrained to satisfy a long-term limit on the discounted cumulative emissions. Once the quotas are determined, an international emissions trading system permits the country to realize the abatement program at the least global cost. The set of normalized equilibria is proposed as the solution concept that could be used to drive the negotiations in such a context. These equilibrium solutions are characterized.

Haurie proposes a model of intergenerational equity to deal with the long-term issue of climate change. The economic impacts of global climate change are far-reaching for the nations of the world. The myopic gain for one generation has to be balanced with the welfare for future generations. In his paper on a stochastic multigeneration game with application to global climate change, Haurie models this problem as a continuous-time, piecewise deterministic game where the players represent successive generations each with a random life duration. The intergenerational equilibrium concept is applied to a model of integrated assessment of climate change.

Part VI contains four chapters that deal with *management science applications.*

Bernhard, Farouq, and Thiery investigate a differential game motivated by a problem in mathematical finance, specifically in the theory of option pricing. The nature of one of the players, called the pursuer, is quite impulsive, resulting in unexpected jumps in the state variables. The authors' approach is to rigorously derive the optimal controls which consist of an initial impulse and a long static coasting followed by finitely many controls later. They investigate the convergence problem of an appropriate Isaacs equation for the discrete version and its convergence to the value function via an equivalent but nonimpulsive differential game.

Jørgensen, Taboubi, and Zaccour consider the problem of a manufacturer and a retailer allocating their advertisement budgets in national advertising and local advertising, respectively. It is modelled as a Stackelberg game with the manufacturer acting as the leader and the retailer as the follower. The manufacturer can create an incentive to influence the retailer's promotional strategy maximizing the total profit. He can also consider advertisement strategies to maximize one's own payoff. The authors study the Stackelberg equilibria for the two distinct models.

Suzuki and Muto study Cournot and Bertrand duopoly markets from the point of view of farsighted stability: players can alternately threaten to move from strategy to strategy objecting to the most recent threat by the opponent, and in each such alternate suggestion the threatening player has utility gain. When von Neumann and Morgenstern defined the notion of stable sets for cooperative TU games, the external stability concept they had via the notion of domination was somewhat of a myopic response by coalitions and was not farsighted. Allowing for the indirect domination that captures farsighted behavior, one can extend the stable sets definition to new stable sets via the indirect external stability. The authors study the consequences of such indirect stability notions for Cournot and Bertrand duopoly markets.

Raghavan concludes this volume by exploring the possibility of modeling tax evasion as a zero-sum two-person generalized stochastic game with incomplete information. The model incorporates the classical statistical classification procedures, the secrecy of the tax office, and the lack of information about the past history of the taxpayer.

We conclude by expressing our thanks to the many persons who have contributed to make this volume a success. First of all, we are indebted to the referees who generously gave their time to ensure the high quality of these Annals. Our special thanks go to Georges Zaccour, then President of the ISDG and Director of GERAD (HEC Montréal), who has generously assisted in dealing with the logistics of producing this edited book. We thank Jaime Brugueras and Tara Raghavan for their help in improving the quality of several manuscripts. Finally, we acknowledge the essential contribution of Francine Benoît, at GERAD, who produced the whole volume with considerable expertise and grace.

Alain Haurie
Shigeo Muto
Leon A. Petrosjan
T.E.S. Raghavan

Contributors

Pierre Bernhard, I3S, University of Nice-Sophia Antipolis and CNRS, France

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

Naïma El Farouq, I3S, University of Nice-Sophia Antipolis and CNRS, France

Yukihiko Funaki, School of Political Science and Economics, Waseda University, 1-6-1 Nishi-Waseda, Shinjuku-ku, Tokyo 169-8050, Japan

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

Peter V. Golubtsov, Physics Department, Moscow State Lomonosov University, Moscow, 119899, Russia

Alain Haurie, Logilab-HEC, University of Geneva, 40 Blvd du Pont d'Arve, CH-1211 Geneva 4, Switzerland

Noboru Hidano, Tokyo Institute of Technology, Department of Social Engineering, 2-12-1, Ookayama, Meguroku, Tokyo, Japan

Steffen Jørgensen, Department of Organization and Management, University of Southern Denmark, Odense, Denmark

Arkadii Kryazhimskii, Moscow State University, Russia

Sergey S. Kumkov, Institute of Mathematics and Mechanics, Ural Branch of Russian Academy of Sciences, S. Kovalevskaya str., 16, GSP-384, Ekaterinburg, 620219, Russia

Robert McKelvey, Department of Mathematical Sciences, University of Montana, Missoula, Montana 59812, USA

Yaroslav Minullin, Environmentally Compatible Energy Strategies Project, International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria

Francesco Moresino, Logilab-HEC, University of Geneva, 40 Blvd du Pont d'Arve, CH-1211 Geneva 4, Switzerland

Jacqueline Morgan, Università di Napoli "Federico II", Dipartimento di Matematica e Statistica, Via Cintia, 80126 Napoli, Italy

Shigeo Muto, Department of Social Engineering, Graduate School of Decision Science and Technology, Tokyo Institute of Technology, Japan

Oleg Nikonov, Urals State Technical University, Russia

Fioravante Patrone, Università di Genova, Dipartimento di Matematica, Via Dodecaneso 35, 16146 Genova, Italy

Valerii S. Patsko, Institute of Mathematics and Mechanics, Ural Branch of Russian Academy of Sciences, S. Kovalevskaya str., 16, GSP-384, Ekaterinburg, 620219, Russia

Leon A. Petrosjan, St. Petersburg State University, Faculty of Applied Mathematics and Control Processes, Bibliotechnaya pl. 2, St. Petersburg, 198504, Russia

N. N. Petrov, Udmurt State University, Izevsk, Russia

T. E. S. Raghavan, Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, USA

Dinah Rosenberg, Laboratoire d'Analyse Géométrie et Applications, Institut Galilée, Université Paris Nord, 93430 Villetaneuse, France

Jürgen Scheffran, University of Illinois at Urbana-Champaign, ACDIS, 505 East Armory Ave., Champaign, Illinois 61820, USA

Vyacheslav P. Serov, Ural Polytechnic Institute, Ekaterinburg, Russia

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

Eilon Solan, MEDS Department, Kellogg School of Management, Northwestern University, Evanston, Illinois 60208, USA, *and* the School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel

Peter Sudhölter, Department of Economics, University of Southern Denmark at Odense, Denmark

Akihiro Suzuki, Department of Public Policy and Social Studies, Faculty of Literature and Social Sciences, Yamagata University, Japan

Sihem Taboubi, GERAD and Department of Marketing, HEC Montréal, Montréal (Québec) Canada H3T 2A7

Stéphane Thiery, I3S, University of Nice-Sophia Antipolis and CNRS, France

D. A. Vagin, Udmurt State University, Izevsk, Russia

Alexander Vasin, Faculty of Computational Mathematics and Cybernetics, Moscow State University, Moscow 119899, Russia, and New Economic School, Moscow

Nicolas Vieille, Département Finance et Économie, HEC, 78351 Jouy-en-Josas, France

Laurent Viguier, EPFL REME, BP 16, CH-1015 Lausanne, Switzerland

Takehiko Yamato, Department of Value and Decision Science, Graduate School of Decision Science and Technology, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552, Japan

Georges Zaccour, GERAD and Department of Marketing, HEC Montréal, Montréal (Québec) Canada H3T 2A7

Stochastic Games with Imperfect Monitoring*

Dinah Rosenberg
Laboratoire d'Analyse Géométrie et Applications
Institut Galilée, Université Paris Nord
Villetaneuse, France
dinah@zeus.math.univ-paris13.fr

Eilon Solan
MEDS Department, Kellogg School of Management
Northwestern University
Evanston, IL, USA
The School of Mathematical Sciences
Tel Aviv University
Tel Aviv, Israel
e-solan@kellogg.northwestern.edu
eilons@post.tau.ac.il

Nicolas Vieille
Département Finance et Économie, HEC
Jouy-en-Josas, France
vieille@hec.fr

Abstract

This chapter provides an introductory exposition of stochastic games with imperfect monitoring. These are stochastic games in which the players imperfectly observe the play. We discuss at length a few basic issues, and describe selected contributions.

1 Introduction

Our objective in this paper is to provide an introductory exposition of some recent work on zero-sum stochastic games with imperfect monitoring. We will try to avoid many of the technical subtleties inherent in this type of work by discussing at length some fundamental issues, before we introduce the basic

*We acknowledge the financial support of the Arc-en-Ciel/Keshet program for 2001/2002. The research of the second author was supported by the Israel Science Foundation (Grant No. 69/01-1).

insights of the known results. This introduction briefly recalls historical developments of the theory, discussed more extensively later, and describes the organization of the paper.

Stochastic games are played in stages. At every stage $n \in \mathbf{N}$ the players are to play one matrix game, taken from a finite set of possible games, called *states.* The matrix game played at stage n depends on the actions that were played at stage $n-1$ and on the previous state. In the present paper, we limit ourselves to zero-sum games, i.e., to the case where each component matrix game is a (two-player) zero-sum game. *Imperfect monitoring* refers to a situation where past moves of a player are imperfectly observed by his/her opponent, as opposed to perfect monitoring. Most work on stochastic games assumes perfect monitoring.

Stochastic games were introduced in a seminal paper of Shapley [26]. Shapley introduced discounted games in which each player i uses a discounted evaluation, that is, he wishes to maximize the discounted sum $\lambda \sum_{n=1}^{\infty}(1-\lambda)^{n-1} r_n^i$, where $\lambda \in (0,1)$ is the common discount factor, and r_n^i is the payoff to player i at stage n. He proved that any λ-discounted zero-sum stochastic game with perfect monitoring has a value v_λ. In addition, he proved that each player has an optimal strategy that is stationary: it depends only on the current state, and not on past history. Blackwell [3] analyzed one-player stochastic games, better known as Markov decision process or stochastic dynamic programming problems. For such games, Blackwell proved that there is a stationary strategy that is optimal for every discount factor λ sufficiently close to zero. This robustness, or uniformity, result was extended by Mertens and Neyman [18] to the class of zero-sum stochastic games with perfect monitoring. Specifically, Mertens and Neyman proved that, given any $\varepsilon > 0$, each player has a strategy that is ε-optimal in the λ-discounted game, for *every* $\lambda < \lambda(\varepsilon)$. Thus, a single strategy is approximately optimal, whatever the discount factor being used, provided it is sufficiently small. However, in contrast to Shapley's and Blackwell's results, in general this strategy cannot be taken to be stationary.

The consequences of imperfect monitoring have been widely explored within the framework of repeated games, see, e.g., Radner [19], Rubinstein and Yaari [25] and Lehrer [14–17]. Most of the interest has been focused on trying to provide a characterization of the set of equilibrium payoffs.

Stochastic games with imperfect monitoring were first analyzed in a series of papers by Coulomb [7,9,10]. In these papers, Coulomb analyzes absorbing stochastic games. These are stochastic games in which the state changes at most once along the play. Coulomb provides an example where the value does not exist, and proves that in this class of games the max-min and min-max values always exist (see Section 3 for definitions). This existence result was recently extended to all zero-sum stochastic games with imperfect monitoring independently by Coulomb [11] and Rosenberg et al. [23]. Flesch et al. [12] showed that a slight amount of imperfect monitoring in non-zero-sum games can prevent the existence of equilibrium payoffs.

This paper is organized as follows. In Section 2, we discuss a number of classical examples, in order to highlight a few fundamental issues related to imperfect monitoring. In particular, we will argue that the issue of imperfect monitoring is irrelevant both for zero-sum repeated games and for λ-discounted stochastic games. Alternatively, it is relevant only for zero-sum stochastic games, in connection with the uniformity property mentioned above. In addition, we illustrate with three examples the consequences of imperfect monitoring. The discussion in this section is mainly kept at a heuristic level. In Section 3, we will be more specific in introducing a formal model and in stating existence results. Section 4 contains a discussion of the proofs. It first summarizes the main insights of the proof of Mertens and Neyman [18]. It then explains how those insights are used in the analysis of games with imperfect monitoring. We conclude by discussing related work.

2 Basic Observations and Examples

2.1 Repeated Matching Pennies

We start with one of the simplest games, *Matching Pennies.* A version of the strategic form of this game is given by the table

	L	R
T	0	1
B	1	0

in which player 1 and player 2 are respectively the row and column players, and whose entries contain the payoff paid by player 2 to player 1.

The value of the one-shot Matching Pennies is 1/2. Each player has a unique optimal strategy, which is mixed and assigns probability 1/2 to both actions.

Suppose that the game is repeated over time. Consider the strategy σ of player 1 that tosses a fair coin at each stage, independent of previous tosses, and that plays T or B depending on the outcome. Let τ be any strategy of player 2. Such a strategy specifies, for each $n \in \mathbf{N}$, the mixed move (that is to say a probability distribution over the set $\{L, R\}$) to be used at stage n, as a function of all the information available at stage n. For each given n, the (conditional) distribution of player 1's move at stage n, given the information known to player 2, assigns probability 1/2 to each action. Therefore, the (conditional) expected payoff at stage n under (σ, τ), given player 2's information, is equal to 1/2. By averaging over all possible information sets of player 2 at stage n, this implies that the expected payoff under (σ, τ) at each stage n is 1/2.

As a consequence, the strategy σ guarantees that the (expected) payoff to player 1 in the repeated game is exactly $1/2$, whatever be the weights assigned to the different stages. A similar analysis holds true for player 2.

In other words, the value of the infinitely repeated Matching Pennies is the same as that of the one-shot Matching Pennies. Moreover, the strategy in the repeated game that repeats an optimal strategy of the one-shot game is optimal. Plainly, these conclusions are not specific to the Matching Pennies game, and hold more generally for every repeated two-player zero-sum game. Note that the preceding observation holds, no matter what information about past play is available to the players. Therefore, the nature of monitoring—perfect or imperfect—is irrelevant for the analysis of repeated zero-sum games.

To conclude this example, it may be helpful to realize why the conclusions are dramatically different for non-zero-sum repeated games. Let G be a given strategic form game that is repeated over time. Generalizing upon the above observation, the strategy profile that consists of repeating over time a given equilibrium profile x of the game G is a Nash equilibrium of the repeated game, whose payoff coincides with the payoff induced by x in G. In contrast to zero-sum games, this need not be the unique equilibrium payoff of the repeated game. When it comes to a characterization of the equilibrium payoffs in the repeated game, the nature of monitoring is crucial. Typical proofs of the *Folk Theorem* (see, e.g., Sorin [29] or Aumann and Shapley [2]) proceed along the following lines: given a payoff vector, a play path is identified that induces this payoff. A strategy profile is next designed, under which the players are required to follow the play path, and to "punish" reciprocally in case of deviations from this path. Clearly, whether or not this profile is an equilibrium depends on the extent to which deviations are observed and deviators identified. A complete characterization of the equilibrium payoffs is not yet available. A solution has been provided by Lehrer [14–17] for various notions of undiscounted equilibrium and/or under specific assumptions on the monitoring structure. Only partial results have been established for discounted games, see, e.g., the January 2002 special issue of the *Journal of Economic Theory*, and the references therein.

2.2 The Big Match

We here recall well-known results on the Big Match game, an example of an absorbing stochastic game introduced by Gillette [13] and later analyzed by Blackwell and Ferguson [4]. Although the formal model of stochastic games has yet to be introduced, this example will clarify why the issue of monitoring is irrelevant for the analysis of the discounted games, but not if one seeks to establish optimality properties which are uniform with respect to (w.r.t.) the discount factor.

The Big Match is described by the table

	L	R
T	0*	1*
B	1	0

Both players have two actions. As long as player 1 plays the Bottom row, his payoff is given by this table. As soon as he plays the Top row, say at stage θ, the payoff to player 1 at this stage and *all* subsequent stages is 0 or 1, depending on whether player 2 played the Left or the Right column at stage θ. Equivalently, at stage θ, the play moves to one of two possible trivial (absorbing) states, in which the payoff is constant.

As proven in Shapley [26], the value v_λ of the λ-discounted game satisfies a dynamic programming principle. Indeed, v_λ is uniquely characterized as the value of the following one-shot zero-sum game $\Gamma(v_\lambda)$:

	L	R
T	0	1
B	$\lambda + (1-\lambda)v_\lambda$	$(1-\lambda)v_\lambda$

Moreover, let x_λ denote an optimal mixed move of player 1 in the game $\Gamma(v_\lambda)$. Then the stationary strategy that plays x_λ at every stage is an optimal strategy in the λ-discounted stochastic game.[1]

Again, this property is not specific to the Big Match. More generally, the existence of stationary optimal strategies in λ-discounted games ensures that the λ-discounted value v_λ is independent of the type of monitoring, as long as both players are always informed of the current state. Hence the issue of monitoring is irrelevant for the analysis of λ-discounted games.

We now discuss the existence of strategies that are ε-optimal in all λ-discounted games, provided λ is close enough to zero. We shall discuss only the problem faced by player 1, since the stationary strategy of player 2 that assigns probability 1/2 to both actions is optimal for each $\lambda > 0$. Assume first that the assumption of perfect monitoring holds, that is, at each stage n, player 1 knows the complete sequence of actions selected by player 2 up to stage n.

Blackwell and Ferguson [4] devised a parametric family $(\sigma_N)_{N\in\mathbf{N}}$ of strategies. For $n \in \mathbf{N}$, define e_n to be the excess number of stages up to n in which player 2 selected the Left column: $e_n = l_n - r_n$, where l_n and r_n are respectively the number of stages up to stage n in which player 2 played the Left and the Right columns. The strategy σ_N plays the Top row with probability

[1]One can check that for the Big Match, $v_\lambda = 1/2$, player 1 has a unique optimal strategy that assigns probability $\lambda/(1+\lambda)$ to the Top row, and player 2 has a unique optimal strategy that assigns probability 1/2 to each column.

$1/(N+e_n)^2$ (assuming player 1 played the Bottom row in all previous stages). Then there is a constant λ_N such that for every strategy τ, the λ-discounted payoff induced by σ_N and τ is at least $(N-2)/2N$, provided $\lambda < \lambda_N$; for a proof of this result see Blackwell and Ferguson [4] and Coulomb [8].

The intuition behind this strategy is as follows. Suppose that $e_n > 0$, so that so far player 2 played the Left column more often than the Right column. In this case player 1 does not want the game to terminate: as long as the frequency of Left is higher, his payoff by playing the Bottom row is more than 1/2. Therefore, player 1 decreases the probability of playing the Top row. If, on the other hand, $e_n < 0$, player 1 would like the game to terminate, so he increases the probability of playing the Top row. The effect of any given stage on the probability of playing the Top row is small, so that any strategic manipulation of the future behavior of player 1 by player 2 comes at the cost of being absorbed to a bad payoff while the manipulation takes place. However, this effect is large enough so that, if $\liminf_{n\to\infty}(e_n) < 0$, player 1 will eventually play the Top row.

We postpone the discussion of the strategies devised by Mertens and Neyman [18] to Section 4.1. In a nutshell, according to Mertens and Neyman, at every stage n player 1 plays a stationary λ_n-discounted strategy x_{λ_n}, where λ_n is defined recursively as a function of λ_{n-1} and of the choice of player 2 in stage $n-1$.

To contrast with the full monitoring case, we now assume that player 1 receives no information about past moves of player 2. Since the game stops at the first stage θ in which player 1 chooses the Top row, a strategy of player 1 reduces here to a sequence $\mathbf{x} = (x_n)_{n\in\mathbf{N}}$, with the interpretation that x_n is the probability assigned to the Top row at stage n, assuming $\theta \geq n$.

Let such a strategy $\mathbf{x}$ be given. We shall now check that, given $\varepsilon > 0$, there is a reply τ of player 2 such that the expected λ-discounted payoff under $\mathbf{x}$ and τ is at most ε, provided λ is close enough to zero. Thus, by playing properly, player 2 can lower the expected payoff of player 1 as close to 0 as he wishes. This will prove that player 1 cannot guarantee a positive payoff in all discounted games with sufficiently low discount factor.

Let ε be given. As $\mathbf{x}$ is given, there is N sufficiently large such that the probability that the game reaches stage N and player 1 plays the Top row at least once after that stage is at most $\varepsilon/2$, that is, $\mathbf{P}(N \leq \theta < +\infty) \leq \varepsilon/2$. Let τ be the pure strategy that plays the Left column up to stage N, and the Right column afterwards. If player 1 plays the Top row at some stage $n \leq N$, the terminal payoff is 0. Since the probability that player 1 plays the Top row after stage N is at most $\varepsilon/2$, and since after stage N player 2 plays the Right column, the expected λ-discounted payoff from stage N on, when restricted to the event that play reaches stage N, is at most $\varepsilon/2$. Therefore, if λ is sufficiently small so that the contribution of the first N stages to the λ-discounted payoff is at most $\varepsilon/2$, we deduce that the λ-discounted payoff under $(\mathbf{x},\tau)$ is at most ε.

2.3 A Modified Big Match

We here discuss a striking example due to Coulomb [7], which is a modification of the Big Match. The game is given by the matrix

	L	M	R
T	0^*	1^*	γ
B	1	0	γ

where $\gamma \geq 1/2$ is arbitrary. If either the action combination (T, L) or the action combination (T, M) is played, the play moves to an absorbing state with constant payoff. Note that the current state can change at most once along any play, as in the Big Match. This game differs from the Big Match by the adjunction of the column R.

Assuming perfect monitoring, this extra column makes no difference, since $\gamma \geq 1/2$. Indeed, let σ be any strategy in the Big Match, and define a strategy $\widetilde{\sigma}$ in the present game as follows. Given a history $\widetilde{h}$, $\widetilde{\sigma}$ plays at $\widetilde{h}$ the mixed move that would be played by σ at h, where h is obtained from $\widetilde{h}$ by deleting all stages in which player 2 played R. It can be checked that $\widetilde{\sigma}$ guarantees $1/2 - \varepsilon$ in the λ-discounted game, as soon as σ guarantees $1/2 - \varepsilon$ in the λ-discounted Big Match.

We shall now assume that player l is only imperfectly informed of player 2's past choices. Specifically, we shall assume that, whenever player 1 plays B, he is "told" "L" if player 2 played L, and "M or R" otherwise. The information received by player 1 upon playing T is irrelevant for the present analysis, as well as the signals for player 2.

We now check that in this game player 2 can do much better than in the Big Match. Specifically, given any strategy σ of player 1 and any $\varepsilon > 0$, we shall exhibit a strategy τ of player 2 such that the expected λ-discounted payoff under (σ, τ) is at most ε, for every λ close enough to zero. This result is striking because the signalling structure and the payoffs are such that this game is a Big Match with perfect monitoring with an additional column for which payoffs can be as high as we want. Nevertheless, the highest quantity that player 1 can guarantee in any discounted game with small enough discount factor decreases from 1/2, which is the value of the Big Match with perfect monitoring to 0.

Define θ to be the first stage in which either (T, L) or (T, M) is played, so that θ is the stage at which the game effectively terminates.

Let y be the stationary strategy of player 2 that plays L and R with probabilities $\varepsilon/2$ and $1 - \varepsilon/2$ respectively, and let y' be the stationary strategy of player 2 that plays L and M with probabilities $\varepsilon/2$ and $1 - \varepsilon/2$ respectively.

If the probability that under (σ, y) the game terminates in finite time is 1, then the probability that under (σ, y) the game terminates by (T, L) is 1, so that for every λ sufficiently small, the λ-discounted payoff under (σ, y) is at most ε, as desired.

So assume that this probability is strictly less than 1. Therefore, there is N such that the probability that under (σ, y) stage N is reached and player 1 plays the Top row after stage N is at most $\varepsilon/2$. Let τ be the strategy that coincides with y up to stage N, and with y' afterwards. Since the signal to player 1 is "M or R", whether the action pair is (B, M) or (B, R), as long as player 1 follows σ and does not play the Top row, he *cannot* tell whether he plays against y or against τ. Since the probability that player 1 plays the Top row after stage N is at most $\varepsilon/2$, the probability that player 1 can distinguish between y and between τ is at most $\varepsilon/2$. This means that the probability that under (σ, τ) player 1 plays the Top row after stage N is at most $\varepsilon/2$, while the probability that player 2 plays L at any given stage after stage N is $1 - \varepsilon/2$, so that the expected λ-discounted payoff, restricted to the event that the game is not terminated before stage N, is at most ε. If λ is sufficiently small so that the contribution of the first N stages to the discounted payoff is at most ε, we deduce that the expected λ-discounted payoff under (σ, τ) is at most 2ε.

The two phases in the definition of τ have a natural interpretation. In the first phase, player 2 exhausts the probability that the play will end up in an absorbing state. In the second phase, player 2 switches to a mixed move that yields a low stage payoff. The fact that the mixed moves used by player 2 in the two phases cannot be distinguished by player 1 (as long as he plays B) guarantees that the probability that the play moves to an absorbing state in the second phase is very low. This two-part definition of a reply of player 2 to a given strategy of player 1 turns out to be a powerful tool; see Section 4.4.

3 The Model and the Results

3.1 Stochastic Games with Imperfect Monitoring

We proceed to the model of stochastic games with imperfect monitoring. Given a finite set K, $\Delta(K)$ will denote the set of probability distributions over K. An element $k \in K$ will be identified with the element of $\Delta(K)$ that assigns probability one to k.

A *two-person zero-sum stochastic game with imperfect monitoring* is described by (i) a set S of states, (ii) action sets A and B for the two players, (iii) a daily reward function $r : S \times A \times B \to \mathbf{R}$, (iv) signal sets M^1 and M^2 and (v) a transition function $\psi : S \times A \times B \to \Delta(M^1 \times M^2 \times S)$. The sets S, A, B, M^1 and M^2 are assumed to be finite.

The game is played in stages. An initial state s_1 is given and known to both players. At each stage $n \in \mathbf{N}$, (a) the players independently choose actions $a_n \in A$ and $b_n \in B$; (b) a triple (m_n^1, m_n^2, s_{n+1}) is drawn according to $\psi(s_n, a_n, b_n)$; (c) players 1 and 2 are *only* told m_n^1 and m_n^2 respectively and (d) the game proceeds to stage $n+1$.

We denote by ψ^1 the marginal of ψ on M^1. It stands for the distribution of player 1's signal, as a function of the current state and the current action choices. Unlike the previous examples, note that we allow for the case where the signal depends stochastically on (s, a, b). We will always assume that each player always knows the current state and has perfect recall (i.e., remembers his past choices and past information).

This implies that ψ^1 is such that $(s, a) = (s', a')$ as soon as $\psi^1(s, a, b)\,[m^1] > 0$ and $\psi^1(s', a', b')[m^1] > 0$, for some $m^1 \in M^1$.

We define accordingly ψ^2 as the marginal of ψ over M^2. We also denote by q the marginal of ψ on S. Thus, $q(s, a, b)[s']$ is the probability of moving from s to s', if the players play a and b.

In a sense, ψ^1 provides all the information player 1 has on player 2's current move. However, since the signals to the two players can be correlated, the pair (ψ^1, ψ^2) does not fully describe the information available to player 1 on player 2's signal. Therefore, our model is more general than a model in which, given (s, a, b), the next state and the signals are chosen independently.

A behavior strategy of player 1 is a sequence $\sigma = (\sigma_n)_{n\geq 1}$ of functions $\sigma_n : H_n^1 \to \Delta(A)$, where $H_n^1 = S \times (M^1)^{n-1}$ is the set of "private" histories of player 1 at stage n. A stationary strategy depends only on the current stage. Hence, a stationary strategy of player 1 is described by a vector $(x^s)_{s\in S}$ in $(\Delta(A))^S$, with the interpretation that x^s is the mixed move used whenever the current state is $s \in S$. Strategies τ of player 2 are defined analogously, with obvious changes. We let $H_\infty = (S \times A \times B \times M^1 \times M^2)^{\mathbf{N}}$ denote the set of *plays*. For $i = 1, 2$, $\mathcal{H}_n^i$ denotes the cylinder algebra over H_∞ induced by H_n^i, and we let $\mathcal{H}_\infty = \sigma(\mathcal{H}_n^1, \mathcal{H}_n^2, n \in \mathbf{N})$ denote the σ-algebra generated by all cylinder sets. A given strategy pair (σ, τ), together with an initial state $s \in S$, induces a probability distribution $\mathbf{P}_{s,\sigma,\tau}$ over $(H_\infty, \mathcal{H}_\infty)$. Expectations w.r.t. $\mathbf{P}_{s,\sigma,\tau}$ are denoted $\mathbf{E}_{s,\sigma,\tau}$. The initial state s is used here as a parameter.

Given $\lambda \in (0, 1)$, the λ-discounted payoff induced by a strategy pair (σ, τ) starting from state $s \in S$ is given by

$$\gamma_\lambda(s, \sigma, \tau) = \mathbf{E}_{s,\sigma,\tau}\left[\lambda \sum_{n=1}^{+\infty} (1-\lambda)^{n-1} r(s_n, a_n, b_n)\right].$$

The seminal result of Shapley [26] asserts that the λ-discounted game has a value v_λ that does not depend on ψ. That is, for each $s \in S$, the zero-sum game with payoff function $\gamma_\lambda(s, \cdot, \cdot)$ has a value. We now introduce the definitions of the uniform properties we will be dealing with.

Definition 3.1. Let $\phi \in \mathbf{R}^S$. Player 1 can *guarantee* ϕ if for every $\varepsilon > 0$ there exists a strategy σ and $\lambda_0 \in (0, 1)$ such that

$$\forall s \in S, \forall \tau, \forall \lambda \in (0, \lambda_0),\ \gamma_\lambda(s, \sigma, \tau) \geq \phi(s) - \varepsilon.$$

We then say that the strategy σ guarantees $\phi - \varepsilon$.

Definition 3.2. Let $\phi \in \mathbf{R}^S$. Player 2 can *defend* ϕ if for every $\varepsilon > 0$ and every strategy σ of player 1 there exists a strategy τ of player 2 and $\lambda_0 \in (0,1)$, such that

$$\forall s \in S, \forall \lambda \in (0, \lambda_0), \gamma_\lambda(s, \sigma, \tau) \leq \phi(s) + \varepsilon.$$

We then say that the strategy τ defends $\phi + \varepsilon$ against σ.

The definitions of a vector guaranteed by player 2, and defended by player 1, are similar, with the roles of the two players exchanged.

Definition 3.3. A vector $v \in \mathbf{R}^S$ is
- the (uniform) *value* of Γ if both players can guarantee v;
- the (uniform) max-min if player 1 can guarantee v and player 2 can defend v;
- the (uniform) min-max if player 1 can defend v and player 2 can guarantee v.

Assume that player 1 cannot guarantee ϕ. Then, for every strategy σ and every $\lambda_0 \in (0,1)$, there is a strategy τ such that $\gamma_\lambda(s, \sigma, \tau) < \phi(s) - \varepsilon$, for *some* $\lambda \in (0, \lambda_0)$ and some $s \in S$. Plainly, it does not follow that player 2 can defend ϕ. Therefore the existence of the max-min is not at all a trivial matter.

Note that the value coincides with $\lim_{\lambda \to 0} v_\lambda$, as soon as it exists. Also, if both the max-min and the min-max exist, one has max-min $\leq$ min-max.

3.2 The Results

We first quote two known results. A stochastic game has perfect monitoring if the signal received by a player always reveals the current state and the action choices. Formally, given $(s, a, b) \neq (s', a', b')$, the supports of the probability distributions $\psi^i(s, a, b)$ and $\psi^i(s', a', b')$ are disjoint, for $i = 1, 2$.

Theorem 3.1 (Mertens and Neyman [18]). *Every two-player zero-sum stochastic game with perfect monitoring has a value.*

Actually, the proof of Theorem 3.1 is valid as soon as the signals the players receive at each stage n contain the new state s_{n+1} and the payoff r_n at stage n.

Let Γ be a stochastic game. A state $s \in S$ is *absorbing* if $q(s, a, b)\,[s] = 1$, for every $(a, b) \in A \times B$. The game Γ is *absorbing* if all states but one are absorbing.

Theorem 3.2 (Coulomb [7,9,10]). *Every two-player zero-sum absorbing stochastic game has a max-min and a min-max. The max-min (resp. the min-max) depends on ψ only through ψ^1 (resp. through ψ^2).*

In the rest of this paper, we will report on the following theorem, obtained independently by Coulomb [11] and Rosenberg et al. [23]. Our goal is to identify the main ideas of the proof and to strip the exposition from details. The interested reader should consult Coulomb [11] or Rosenberg et al. [23].

Theorem 3.3 (Coulomb [11], Rosenberg, Solan and Vieille [23]). *Every two-player zero-sum stochastic game with imperfect monitoring has a max-min and a min-max. The max-min (resp. the min-max) depends on ψ only through ψ^1 (resp. through ψ^2).*

The proof of Theorem 3.3 is quite related to the proof of Theorem 3.1. It is independent of the proof of Theorem 3.2.

W.l.o.g. we focus on the existence of the max-min value, and assume that payoffs are non-negative and bounded by 1.

4 Existence of the Max-Min: Highlights

4.1 On Mertens and Neyman's [18] Proof

The proof of Theorem 3.3 builds upon the proof of Theorem 3.1. We therefore start by recalling the main insights of Mertens and Neyman [18] (hereafter MN). We will next single out the main computational step in their proof, and discuss the additional issues that arise in games with imperfect monitoring.

MN offer a wide class of ε-optimal strategies σ for player 1. All share the following structure. The play is divided into blocks of random finite-length L_k. On each block k, the strategy requires to play an optimal strategy in the λ_k-discounted game. Both L_k and λ_k depend on an auxiliary parameter, z_k: $L_k = L(z_k)$ and $\lambda = \lambda(z_k)$. In a sense, $z_k \in \mathbf{R}$ is a statistic that summarizes all the relevant aspects of the play, up to the beginning of block k.

To be specific, given two functions $L : [0, +\infty) \to \mathbf{N}$ and $\lambda : [0, +\infty) \to \mathbf{R}$, and $M \in \mathbf{R}$, the sequences $(z_k), (L_k), (\lambda_k)$ are defined recursively by

$$z_1 \geq Z, B_1 = 1, \; \lambda_k = \lambda(z_k), L_k = L(z_k),$$

$$B_{k+1} = B_k + L_k, z_{k+1} = \max\left\{ Z, z_k + \frac{\varepsilon}{2} + \sum_{B_k \leq n < B_{k+1}} (r_n - v(s_{B_{k+1}})) \right\}, \tag{1}$$

where $v = \lim_{\lambda \to 0} v_\lambda$ and $r_n = r(s_n, a_n, b_n)$ is the payoff in stage n. In a first approximation, the new value z_{k+1} of the statistic is obtained by adding to the previous value z_k the excess of payoffs received over the values of the states visited along the block.

MN provide sufficient conditions on the functions $\lambda(\cdot)$ and $L(\cdot)$ under which the above strategy σ is ε-optimal, for M large. These conditions are in particular satisfied for each of the two following simple functions.

Case 1: $\lambda(z) = z^{-\beta}$ and $L(z) = \lceil \lambda(z)^{-\alpha} \rceil$, where $\alpha \in (0, 1)$ satisfies $\|v_\lambda - v\|_\infty < \lambda^{1-\alpha}$ for every λ sufficiently close to 0, and $\beta > 1$ satisfies $\alpha\beta < 1$;

Case 2: $\lambda(z) = 1/(z \ln^2 z)$ and $L(z) = 1$.

Given an appropriate choice for $\lambda(\cdot)$ and $L(\cdot)$, MN prove that $\gamma_\lambda(s,\sigma,\tau) \geq v(s) - \varepsilon$ for each τ, and that $\mathbf{E}_{s,\sigma,\tau}[\sum_{k=1}^{+\infty} \lambda_k L_k] < +\infty$.

The proof relies on the semi-algebraicity of the map $\lambda \mapsto v_\lambda$, due to Bewley and Kohlberg [5], and on inequality (2) below, which holds for every τ, since during block k, σ follows an optimal strategy in the λ_k-discounted game:

$$\mathbf{E}_{s,\sigma,\tau}\left[\lambda_k \sum_{n=B_k}^{B_{k+1}-1} (1-\lambda_k)^{n-B_k} r_n + (1-\lambda_k)^{L_k} v_{\lambda_k}(s_{B_{k+1}}) | \mathcal{H}_{B_k}\right] \geq v_{\lambda_k}(s_{B_k}). \tag{2}$$

We conclude this section by listing some standing issues that need to be addressed in order to adapt MN's proof to games with imperfect monitoring. This list is not exhaustive.

- In games with imperfect monitoring, the max-min need not be equal to the limit of the λ-discounted values. The above proof asserts that player 1 can guarantee $\lim_{\lambda\to 0} v_\lambda$. Therefore, we will have to define *auxiliary* discounted games. The proof when imperfect monitoring is present will assert that the max-min is equal to the limit of the solutions to these auxiliary discounted games. The definition of these auxiliary games will take into account the structure ψ of signals.
- The solution w_λ of these auxiliary games will have to be semi-algebraic as a function of λ.
- In (1), the updating formula for z_k involves $\sum_{B_k \leq n < B_{k+1}} r_n$, the payoffs received in the previous block. Since this quantity is not available to player 1, we will have to estimate it using only the information that is available to player 1. In effect, we will use a measure of the worst payoff that is consistent with the distribution of the signals received in the elapsed block.
- Finally, the ε-optimality computation will have to be adapted.

As it turns out, the last issue is easy. Specifically, replace in (1) the term $\sum_{n=B_k}^{B_{k+1}-1} r_n$ by an $\mathcal{H}^1_{B_k}$-measurable variable $L_k \widehat{r}_k$, and let $\lambda \mapsto w_\lambda$ be an $\mathbf{R}^S$-valued semi-algebraic function, with $w := \lim_{\lambda\to 0} w_\lambda$. Let $\lambda(\cdot)$, $L(\cdot)$ satisfy MN's sufficient conditions, and let M be large enough. A close inspection of MN's proof reveals that the following theorem holds.

Theorem 4.1. *There exists* $\lambda_0 \in (0,1)$ *such that the following holds. Let* (σ,τ) *be a strategy pair such that*

$$\mathbf{E}_{s,\sigma,\tau}\left[\lambda_k L_k \widehat{r}_k + (1-\lambda_k L_k) w_{\lambda_k}(s_{B_{k+1}}) | \mathcal{H}^1_{B_k}\right] \geq w_{\lambda_k}(s_{B_k}) - \frac{\varepsilon}{12}\lambda_k L_k, \tag{3}$$

$\mathbf{P}_{s,\sigma,\tau}$*-a.s. for every* k*. Then for each* $\lambda \in (0,\lambda_0)$*,*

$$\mathbf{E}_{s,\sigma,\tau}\left[\lambda \sum_{n=1}^{+\infty} (1-\lambda)^{n-1} \widehat{R}_n\right] \geq w(s) - \varepsilon, \tag{4}$$

where $\widehat{R}_n = \widehat{r}_k$ *for* $B_k \leq n < B_{k+1}$. *Moreover,*

$$\mathbf{E}_{s,\sigma,\tau}\left[\sum_{k=1}^{+\infty} \lambda_k L_k\right] < +\infty. \tag{5}$$

4.2 Auxiliary Discounted Games

We let a stochastic game $\Gamma = (S, A, B, M^1, M^2, \psi, r)$ be given. We here define an auxiliary family of stochastic games. The stage payoff of these games incorporates the structure of the signals.

A preliminary comment is in order. Consider first a *repeated* game with imperfect monitoring. Assume that player 1 and player 2 consider using mixed moves $x \in \Delta(A)$ and $y \in \Delta(B)$ in some given stage. If player 2 replaces y by another mixed move $y' \in \Delta(B)$, this replacement can possibly have an effect on the future behavior of player 1 only if it alters the distribution of signals to player 1 at that stage. In other words, if y' is *indistinguishable* from y, in the sense that the distributions $\psi^1(x,y)$ and $\psi^1(x,y')$ of signals to player 1 coincide, then switching from y to y' while player 1 is using x has no incidence whatever on player 1's future behavior.[2] This suggests that a proper modified payoff function for player 1 is $\widetilde{r}(x,y) = \inf r(x,y')$, where the infimum is taken over all $y' \in \Delta(B)$ that are indistinguishable from y given x. That is, $\widetilde{r}(x,y)$ is the worst payoff to player 1, given that player 1's signals are consistent with y.

This equivalence relation and the corresponding modified payoff function have played an important role in the analysis of games with imperfect monitoring, see Aumann and Maschler [1], Lehrer [14–17] and Coulomb [9,10].

However, this relation is not well suited for general stochastic games with imperfect monitoring. Indeed, a mixed move y' can be practically indistinguishable from y if the probability that player 1 can distinguish between y and y' is quite small compared to the discount factor. We therefore amend it as follows. Given a discount factor $\lambda \in (0,1)$, a state $s \in S$, a mixed move $x \in \Delta(A)$ and an additional parameter $\varepsilon \in (0,1)$, we say that $y \in \Delta(B)$ and $z \in \Delta(B)$ are indistinguishable, written $y \sim_{\lambda,\varepsilon,s,x} z$ if

$$\psi^1(s,a,y) = \psi^1(s,a,z) \quad \text{for every } a \text{ such that } x[a] \geq \lambda/\varepsilon.$$

Accordingly, we set

$$\widetilde{r}^{\varepsilon}_{\lambda}(s,x,y) = \min_{z \sim_{\lambda,\varepsilon,s,x} y} r(s,x,z). \tag{6}$$

As above, it can be thought of as the worst payoff consistent with a given distribution of signals to player 1. The specific role of the parameter ε will be clarified later.

---

[2]This is equivalent to the requirement that $\psi^1(a,y) = \psi^1(a,y')$ for every action $a \in A$ that is played with positive probability under x.

We briefly mention some basic properties of $\widetilde{r}^{\varepsilon}_{\lambda}$. Note first that $\widetilde{r}^{\varepsilon}_{\lambda} \leq r$. Since $\sim_{\lambda,\varepsilon,s,x}$ is an equivalence relation, one has $\widetilde{r}^{\varepsilon}_{\lambda}(s,x,y) = \widetilde{r}^{\varepsilon}_{\lambda}(s,x,z)$ whenever $z \sim_{\lambda,\varepsilon,s,x} y$. In addition, it can be checked that, for fixed λ, ε and s, the function $\widetilde{r}^{\varepsilon}_{\lambda}(s,\cdot,\cdot)$ is continuous with respect to y and upper semi-continuous in the pair (x,y). Finally, the map $(\varepsilon,\lambda,x,y) \mapsto \widetilde{r}^{\varepsilon}_{\lambda}(s,x,y)$ is semi-algebraic.

We now proceed to introducing a vector $v^{\varepsilon}_{\lambda}$, which will play the role of the "value" of the auxiliary discounted game. Specifically, define $v^{\varepsilon}_{\lambda} \in \mathbf{R}^S$ as the unique solution to the fixed-point equation

$$v^{\varepsilon}_{\lambda}(s) := \max_{x\in\Delta(A)} \min_{y\in\Delta(B)} \left\{\lambda \widetilde{r}^{\varepsilon}_{\lambda}(s,x,y) + (1-\lambda)\mathbf{E}_{q(s,x,y)}[v^{\varepsilon}_{\lambda}(\cdot)]\right\}, \quad w \in \mathbf{R}^S, \quad (7)$$

where $\mathbf{E}_{q(\cdot|s,x,y)}$ is the expectation w.r.t. $q(s,x,y)$.[3] It follows from this fixed-point property that the map $(\lambda,\varepsilon) \mapsto v^{\varepsilon}_{\lambda}(s)$ is semi-algebraic.

One can relate $v^{\varepsilon}_{\lambda}$ to the sup inf of some non-standard λ-discounted game. Indeed, define the (ε,λ)-game to be a λ-discounted game, in which the stage payoff is $\widetilde{r}^{\varepsilon}_{\lambda}$. The (ε,λ)-game differs from standard stochastic games in several respects. At each stage, the players choose *mixed moves* in $\Delta(A)$ and $\Delta(B)$ (and not actions in A and B). In addition, the stage payoff function depends on the discount factor being used. It can be checked that $v^{\varepsilon}_{\lambda}$ coincides with the supinf of the (ε,λ)-game, when players are restricted to pure strategies.

We conclude this section by offering a candidate for the max-min. Since the map $\lambda \mapsto v^{\varepsilon}_{\lambda}(s)$ is semi-algebraic for fixed ε, the limit $\lim_{\lambda\to 0} v^{\varepsilon}_{\lambda}(s)$ exists for every $\varepsilon > 0$. In addition, the auxiliary reward $\widetilde{r}^{\varepsilon}_{\lambda}$ is non-decreasing w.r.t. ε, hence so is $\lim_{\lambda\to 0} v^{\varepsilon}_{\lambda}(s)$. As a consequence, the limit $v := \lim_{\varepsilon\to 0}\lim_{\lambda\to 0} v^{\varepsilon}_{\lambda}$ exists. It turns out that v is the max-min of the game Γ, as we explain in the next two sections.

4.3 Guaranteeing *v*

We here explain why player 1 can guarantee v. We shall rely on the tools introduced in Section 4.1, and we first introduce the function w_λ that will be used. Using the theory of semi-algebraic sets, there is a semi-algebraic function $\lambda \in (0,1) \mapsto \varepsilon(\lambda) \in (0,1)$ such that $\lambda \leq \varepsilon(\lambda)^2$ for each λ, and $\lim_{\lambda\to 0} v^{\varepsilon(\lambda)}_{\lambda} = v$. We set $w_\lambda := v^{\varepsilon(\lambda)}_{\lambda}$. Besides, there is a semi-algebraic map $\lambda \in (0,1) \mapsto x_\lambda = (x^s_\lambda)_{s\in S} \in \Delta(A)^S$, such that, for each $s \in S$, x^s_λ achieves the maximum in the definition of $v^{\varepsilon(\lambda)}_{\lambda}$, see (7). By semi-algebraicity again, the set $\overline{A}(s) = \{a \in A : x^s_\lambda[a] \geq \lambda/\varepsilon(\lambda)\}$is, for λ close enough to zero, independent of λ.

We now define the estimate $\widehat{r}_k$ that is used by player 1 at the end of block k to update the statistic z_k. At the end of block k, player 1 collects the signals he received during the block. For each state $s \in S$, player 1 computes a mixed move $\widehat{y}^s \in \Delta(B)$ that is "most likely" given the signals he received in state s. Specifically, for each state $s \in S$, and each action $a \in \overline{A}(s)$, player 1 computes

[3] The justification of why the max and min in (7) are achieved is omitted.

the empirical distribution $\rho_{s,a}$ of signals that he received in those stages in which he played a while at state s (if there was no such stage, the definition of $\rho_{s,a}$ is irrelevant). The mixed move $\widehat{y}^s$ is chosen to minimize over $y \in \Delta(B)$ the maximal discrepancy $\max_{a\in\overline{A}(s)} \|\rho_{s,a} - \psi^1(s,a,y)\|_\infty$. Finally, player 1 sets

$$\widehat{r}_k = \frac{1}{L_k}\sum_{s\in S} N_s \widetilde{r}_{\lambda_k}^{\varepsilon(\lambda_k)}(s, x_{\lambda_k}^s, \widehat{y}^s),$$

where N_s is the number of times the state s was visited during block k. In a sense, $\widehat{r}_k$ is the worst (average) payoff in block k, given that player 2 played a stationary strategy that is consistent with the signals to player 1.

The strategy of player 1 is defined as in Section 4.1, taking Case 1 specifications for $\lambda(\cdot)$ and $L(\cdot)$. To be precise, we first choose $d > 0$ such that $\varepsilon(\lambda) \le \lambda^d$ for λ close enough to zero. We next choose $\alpha \in (1-d, 1)$, $\beta \in (1, 1/\alpha)$ and we set $\lambda(z) = z^{-\beta}$, $L(z) = \lceil \lambda(z)^{-\alpha} \rceil$.

We turn to the intuition of the proof. The crucial part is to show that the inequality (3) is satisfied, provided M is large enough. To this end, we introduce, for each $s \in S$, the average mixed move $\overline{y}^s$ used by player 2 in state s.[4] This average mixed move $\overline{y}$ can be related to the strategy $\widehat{y}$ that is reconstructed by player 1 at the end of the block. Indeed, fix a state $s \in S$, and some action $a \in \overline{A}(s)$. By the definition of $\overline{A}(s)$, at any visit to the state s, the action a is played with probability at least $\lambda/\varepsilon(\lambda) \ge \lambda^{1-d}$, which is much larger than $1/L_k$, provided M is large enough. Thus, provided the number of visits to s exceeds a small fraction of L_k, the action a will typically be played many times. Since the action choices of the two players are independent (conditional on past play), it is quite likely that the empirical distribution of signals $\rho_{s,a}$ will be very close to $\psi^1(s,a,\overline{y}^s)$. As a result, provided the state s is visited more than a negligible fraction of L_k stages, the reconstructed strategy $\widehat{y}$ will be such that $\|\psi^1(s,a,\overline{y}^s) - \psi^1(s,a,\widehat{y}^s)\|_\infty$ is close to zero. By continuity, this will imply that $\widetilde{r}_{\lambda_k}^{\varepsilon(\lambda_k)}(s, x_{\lambda_k}^s, \overline{y}^s)$ is close to $\widetilde{r}_{\lambda_k}^{\varepsilon(\lambda_k)}(s, x_{\lambda_k}^s, \widehat{y}^s)$.[5] On the other hand, states that are visited less than a negligible fraction of L_k stages hardly contribute to $\widehat{r}_k$. Therefore, the expectation $\mathbf{E}_{s,\sigma,\tau}[L_k\widehat{r}_k \mid \mathcal{H}_{B_k}^1]$ of $\widehat{r}_k$ given the past history is close to $\mathbf{E}_{s,\sigma,\tau}[\sum_{s\in S} N_s\widetilde{r}_{\lambda_k}^{\varepsilon(\lambda_k)}(s, x_{\lambda_k}^s, \overline{y}^s) \mid \mathcal{H}_{B_k}^1]$.

Using the optimality of x_{λ_k}, it can be checked—although this is not a trivial observation—that the difference

$$\mathbf{E}_{s,\sigma,\tau}\left[\lambda_k \sum_{s\in S} N_s \widetilde{r}_{\lambda_k}^{\varepsilon(\lambda_k)}(s, x_{\lambda_k}^s, \overline{y}^s) + (1-\lambda_k L_k) v_{\lambda_k}^{\varepsilon(\lambda_k)}(s_{B_{k+1}}) | \mathcal{H}_{B_k}^1\right] - v_{\lambda_k}^{\varepsilon(\lambda_k)}(s_{B_k})$$

is bounded from below by an amount of the order $\varepsilon\lambda_k L_k$. As a consequence, (3) holds.

[4]It is given by $\overline{y}^s = (1/N_s)\sum_{n: s_n = s} y_n$, where the summation runs over all stages of block k, and where $y_n = \tau(h_n^2)$ is the mixed move used by player 2 in stage n.

[5]The formal proof involves many technical complications.

4.4 Defending v

We here deal with the other side of the analysis. We will prove that player 2 can defend $v^\varepsilon := \lim_{\lambda\to 0} v^\varepsilon_\lambda$ for each $\varepsilon > 0$. Let $\varepsilon > 0$, and a strategy σ of player 1 be given. Generalizing upon the example of Section 2.3, we shall define a reply τ in two steps.

First, we use the tools of Section 4.1 to construct a strategy τ_1 that yields a low discounted payoff against σ, when measured in terms of $\widetilde{r}^\varepsilon_\lambda$. It is convenient here to use the specifications of Case 1 for the functions $\lambda(\cdot)$ and $L(\cdot)$: $L(z) = 1$ and $\lambda(z) = 1/(z\ln^2 z)$. In effect, player 2 updates his summary z_k at every stage. We define simultaneously and inductively the strategy τ_1 and the estimate $\widehat{r}_k$.

Consider a given stage $n \in \mathbf{N}$, and assume that τ_1 has already been defined for the first $n-1$ stages, together with $\widehat{r}_1, \dots, \widehat{r}_{n-1}$. Consequently, player 2 has in mind a fictitious discount factor $\lambda_n = \lambda(z_n)$, as determined by (1). At stage n, player 2 computes the conditional distribution of player 1's action choice in stage n, given the past sequence of states. To be specific, we set $\xi_n[\cdot] = \mathbf{P}_{s,\sigma,\tau_1}(a_n = \cdot \mid s_1, \dots, s_n)$. Note that this distribution involves only the restriction of τ_1 to the first $n-1$ stages, and the observation of past states, so that player 2 is indeed in a position to compute ξ_n. The strategy τ_1 recommends playing a mixed move $y_n \in \Delta(B)$ that satisfies

$$\lambda_n \widetilde{r}^\varepsilon_{\lambda_n}(s_n, \xi_n, y_n) + (1-\lambda_n)\mathbf{E}_{q(s_n,\xi_n,y_n)}\left[v^\varepsilon_{\lambda_n}\right] \le v^\varepsilon_{\lambda_n}(s_n). \tag{8}$$

We set $\widehat{r}_n = \widetilde{r}^\varepsilon_{\lambda_n}(s_n, \xi_n, y_n)$. This completes the definition of τ_1.

By the choice of y_n and the definition of $\widehat{r}_n$, (3) trivially holds (with the inequality reversed), which implies that

$$\mathbf{E}_{s,\sigma,\tau_1}\left[\lambda \sum_{n=1}^{+\infty} (1-\lambda)^{n-1} \widetilde{r}^\varepsilon_{\lambda_n}(s_n, \xi_n, y_n)\right] \le v^\varepsilon(s) + \varepsilon, \tag{9}$$

provided λ is close enough to zero.

The inequality (9) says that, if the payoff at every given stage were defined as the *worst* payoff $\widetilde{r}^\varepsilon_{\lambda_n}$, consistent with the actual choice of player 2, then the discounted payoff would be low. Since $\widetilde{r}^\varepsilon_{\lambda_n} \le r$, it however fails to imply $\gamma_\lambda(s,\sigma,\tau_1) \le v(s) + 2\varepsilon$. We now address this issue.

Given a stage n, we let $z_n \in \Delta(B)$ be a mixed move such that $z_n \sim_{\lambda_n,\varepsilon,s_n,\xi_n} y_n$ and $r(s_n,\xi_n,z_n) = \widetilde{r}^\varepsilon_{\lambda_n}(s_n,\xi_n,y_n)$. Hence, z_n achieves the minimal payoff against ξ_n, among all the mixed moves that are indistinguishable from y_n. By the definition of the equivalence relation $\sim_{\lambda_n,\varepsilon,s_n,\xi_n}$, the probability (given the sequence of states) that at stage n, player 1 plays an action that might possibly distinguish y_n from z_n is at most $|A|\,\lambda_n/\varepsilon$. We next make use of the fact that

$$\mathbf{E}_{s,\sigma,\tau_1}\left[\sum_{n=1}^{+\infty} \lambda_n\right] < +\infty$$

(see Theorem 4.1) to choose $N \in \mathbf{N}$ such that $\mathbf{E}_{s,\sigma,\tau}[\sum_{n=N}^{+\infty} \lambda_n] < \varepsilon^2/|A|$. Finally, we let τ be the strategy that coincides with τ_1 up to stage N, and plays z_n rather than y_n in each subsequent stage $n \geq N$.

By the choice of N, the probability that player 1 will ever, from stage N on, play an action that might possibly distinguish τ from τ_1 is at most $(|A|/\varepsilon)\mathbf{E}_{s,\sigma,\tau}[\sum_{n=N}^{+\infty} \lambda_n] \leq \varepsilon$. This implies that the probability distributions $\mathbf{P}_{s,\sigma,\tau}$ and $\mathbf{P}_{s,\sigma,\tau_1}$ induced over the sequences of states differ by at most ε. Therefore,

$$\begin{aligned} \mathbf{E}_{s,\sigma,\tau}\left[r_n\right] = \mathbf{E}_{s,\sigma,\tau}\left[r(s_n,\xi_n,z_n)\right] &= \mathbf{E}_{s,\sigma,\tau}\left[\widetilde{r}^{\varepsilon}_{\lambda_n}(s_n,\xi_n,y_n)\right] \\ &\leq \mathbf{E}_{s,\sigma,\tau_1}\left[\widetilde{r}^{\varepsilon}_{\lambda_n}(s_n,\xi_n,y_n)\right] + \varepsilon. \end{aligned} \tag{10}$$

The first equality simply states that the payoff at stage n is the payoff function evaluated at the current state, with current mixed actions. The second equality follows by the choice of z_n. The inequality follows from the previous claim.

Together with (9), (10) implies that $\gamma_\lambda(s,\sigma,\tau) \leq v^\varepsilon(s) + 2\varepsilon$, provided ε is small enough.

Note that the strategy τ uses only the sequence of states, and not any additional signal that player 2 may receive. It is important to observe that z_n need not satisfy (8) since $q(s_n,\xi_n,y_n) \neq q(s_n,\xi_n,z_n)$. Hence, the two-part definition of τ cannot be avoided.

5 Concluding Comments

The results discussed in this chapter raise additional questions. We will mention just a few.

Within the framework of this survey, it would be useful to characterize the games that have a value. More precisely, given S, A, B, M^1 and M^2, it is interesting to know for which signalling structures ψ the game has a value, for every payoff function r.

The examples of Flesch et al. [12] suggest that the analysis of non-zero-sum stochastic games with imperfect monitoring will need additional insights. This field is yet unexplored.

Finally, challenging problems arise as soon as one drops the assumption that the current state is observed. The one-player case has been investigated in Rosenberg et al. [22]. They prove that the value exists, in the sense that the player can guarantee $\lim_{\lambda\to 0} v_\lambda$. However, they leave unanswered basic questions on the nature of ε-optimal strategies.

In the two-player case, the model is related to stochastic games with incomplete information (see Sorin [27,28,30], Sorin and Zamir [31], Rosenberg and Vieille [24] and Rosenberg et al. [22]). Most work in this area has focused on the case where the state is a pair (k,ω), and (i) the k-component is fixed at the outset of the game and known to one player only, while (ii) the ω-component can change from stage to stage, but is observed by both players. A recent exception is the paper by Renault [20], in which the state s follows a Markov chain, that is, the evolution of s is unaffected by action choices, and is observed only by one player. In this framework, Renault proves the existence of the value. This paper assumes that actions are observed.

REFERENCES

[1] Aumann R.J. and Maschler M., with the collaboration of R.E. Stearns, *Repeated Games with Incomplete Information*, MIT Press, Cambridge, MA, 1995.

[2] Aumann R.J. and Shapley L.S., Long-term competition—a game-theoretic analysis, in *Essays in Game Theory (Stony Brook, NY, 1992)*, 1–15, Springer, New York, 1994.

[3] Blackwell D., Discounted dynamic programming, *Ann. Math. Stat.*, 36, 226–235, 1965.

[4] Blackwell D. and Ferguson T., The Big Match, *Ann. Math. Stat.*, 39, 159–163, 1968.

[5] Bewley T. and Kohlberg E., The asymptotic theory of stochastic games, *Math. Oper. Res.*, 1, 197–208, 1976.

[6] Bochnak J., Coste M. and Roy M.F., *Real Algebraic Geometry*, Springer-Verlag, Berlin, 1998.

[7] Coulomb J.M., Repeated games with absorbing states and no signals, *Internat. J. Game Th.*, 21, 161–174, 1992.

[8] Coulomb, J.M., A note on "Big Match", *ESAIM PS*, 1, 89–93, 1996.

[9] Coulomb J.M., Generalized Big-Match, *Math. Oper. Res.*, 24, 795–816, 1999.

[10] Coulomb J.M., Absorbing games with a signaling structure, *Math. Oper. Res.*, 26, 286–303, 2001.

[11] Coulomb J.M., Stochastic games without perfect monitoring, *Internat. J. Game Th.*, 32, 17–40, 2003.

[12] Flesch J., Thuijsman F. and Vrieze O.J., Stochastic games with non-observable actions, Preprint, 2000.

[13] Gillette D., Stochastic games with zero stop probabilities, *Contributions to the Theory of Games*, 3, Princeton University Press, Princeton, NJ, 1957.

[14] Lehrer E., Lower equilibrium payoffs in two-player repeated games with nonobservable actions, *Internat. J. Game Th.*, 18, 57–89, 1989.

[15] Lehrer E., Nash equilibria of n-player repeated games with semi-standard information, *Internat. J. Game Th.*, 19, 191–217, 1990.

[16] Lehrer E., Correlated equilibria in two-player repeated games with nonobservable actions, *Math. Oper. Res.*, 17, 175–199, 1992a.

[17] Lehrer E., On the equilibrium payoffs set of two player repeated games with imperfect monitoring, *Internat. J. Game Th.*, 20, 211–226, 1992b.

[18] Mertens J.F. and Neyman A., Stochastic games, *Internat. J. Game Th.*, 10, 53–66, 1981.

[19] Radner R., Monitoring cooperative agreements in repeated principle-agent relationship, *Econometrica*, 49, 1127–1148, 1981.

[20] Renault J., The value of Markov chain games with lack of information on one side, TR 2003-18, Cahiers du CEREMADE, Université Paris Dauphine, Paris, 2003.

[21] Rosenberg D., Solan E. and Vieille N., Blackwell optimality in Markov decision processes with partial observation, *Ann. Stat.*, 30, 1178–1193, 2002a.

[22] Rosenberg D., Solan E. and Vieille N., Stochastic games with a single controller and incomplete information, Discussion Paper 1346, Center for Mathematical Studies in Economics and Management Science, Northwestern University, Evanston, IL, 2002b, *SIAM J. Cont. & Optim.*, 43, 86–110.

[23] Rosenberg D., Solan E. and Vieille N., The maxmin value of stochastic games with imperfect monitoring, *Internat. J. Game Th.*, 32, 133-150, 2003.

[24] Rosenberg D. and Vieille N., The maxmin of recursive games with incomplete information on one side, *Math. Oper. Res.*, 25, 23–35, 2000.

[25] Rubinstein A. and Yaari M., Repeated insurance contracts and moral hazard, *J. Econ. Th.*, 30, 74–97, 1983.

[26] Shapley L.S., Stochastic games, *Proc. Nat. Acad. Sci. U.S.A.*, 39, 1095–1100, 1953.

[27] Sorin S., Big Match with lack of information on one side (Part 1), *Internat. J. Game Th.*, 13, 201–255, 1984.

[28] Sorin S., Big Match with lack of information on one side (Part 2), *Internat. J. Game Th.*, 14, 173–204, 1985.

[29] Sorin S., Supergames, in *Game Theory and Applications* (Columbus, OH, 1987), 46–63, *Econom. Theory Econometrics Math. Econom.*, Academic Press, San Diego, CA, 1990.

[30] Sorin S., A first course on zero-sum repeated games, *Mathématiques et Applications*, 37, Springer, Berlin, 2002.

[31] Sorin S. and Zamir S., Big Match with lack of information on one side II, in T.E.S. Raghavan et al. (eds), *Stochastic Games and Related Topics*, Kluwer, Dordrecht, 101–112, 1991.

Level Sweeping of the Value Function in Linear Differential Games*

Sergey S. Kumkov
Institute of Mathematics and Mechanics
Ural Branch of Russian Academy of Sciences
Ekaterinburg, Russia
2445@mail.ur.ru

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

Abstract

In this chapter one considers a linear antagonistic differential game with fixed terminal time T, geometric constraints on the players' controls, and continuous quasi-convex payoff function φ depending on two components x_i, x_j of the phase vector x. Let $\mathcal{M}_c = \{x : \varphi(x_i, x_j) \leqslant c\}$ be a level set (a Lebesgue set) of the payoff function. One says that the function φ possesses the level sweeping property if for any pair of constants $c_1 < c_2$ the relation $\mathcal{M}_{c_2} = \mathcal{M}_{c_1} + (\mathcal{M}_{c_2} \overset{*}{-} \mathcal{M}_{c_1})$ holds. Here, the symbols $+$ and $\overset{*}{-}$ mean algebraic sum (Minkowski sum) and geometric difference (Minkowski difference). Let $\mathcal{W}_c$ be a level set of the value function $(t, x) \mapsto \mathcal{V}(t, x)$. The main result of this work is the proof of the fact that if the payoff function φ possesses the level sweeping property, then for any $t \in [t_0, T]$ the function $x \mapsto \mathcal{V}(t, x)$ also has the property: $\mathcal{W}_{c_2}(t) = \mathcal{W}_{c_1}(t) + \big(\mathcal{W}_{c_2}(t) \overset{*}{-} \mathcal{W}_{c_1}(t)\big)$. Such an inheritance of the level sweeping property by the value function is specific to the case where the payoff function depends on two components of the phase vector. If it depends on three or more components of the vector x, the statement, generally speaking, is wrong. This is shown by a counterexample.

Key words. Linear differential games, value function, level sets, geometric difference, complete sweeping

*This work was supported by the Russian Foundation for Basic Research under Grants No. 06-01-00414 and No. 04-01-96099.

1 Introduction

The central theme for this work is the operation of the geometric difference (Minkowski difference). Its definition and basic properties are given, for example, in [5]. At the early stage of developing the theory of differential games, the geometric difference was applied in [13,14] to solve games with linear dynamics. After that, the concept of the geometric difference was intensively used in the theory of control and differential games (see, for example, [10,3,2,9]).

As usual, the *algebraic sum* (Minkowski sum) of two sets A and B is the set $A + B = \{a + b : a \in A,\, b \in B\}$.

Definition 1.1. The geometric difference of two sets A and B, where $B \neq \varnothing$, is the set $A \overset{*}{-} B = \{x : B + x \subset A\}$. In other words, the geometric difference of the sets A and B is the set of elements such that each of them shifts the set B into the set A.

Let us give some planar examples (Figure 1). The example a) shows the geometric difference of a large square and a small circle. The result is a square with the sides shorter than the original ones by the diameter of the circle. The example b) demonstrates the geometric difference of two circles. The result is also a circle with the radius equal to the difference of the radii of the original circles.

If the set A is convex, then the set $A \overset{*}{-} B$ is convex too. In general the following relation holds:

$$B + (A \overset{*}{-} B) \subset A,$$

that is, the subtrahend set after summation with the geometric difference gives, generally speaking, only a subset of the original set. For instance, in the first

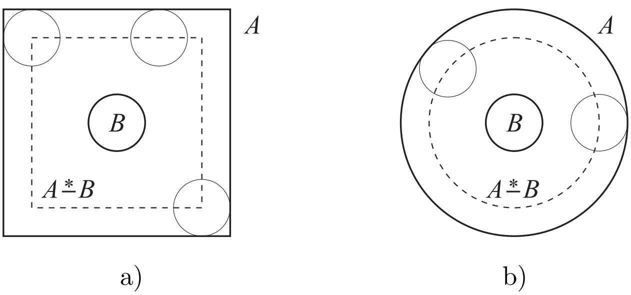


Figure 1: Examples of geometric difference: a) the geometric difference of a square and a circle; b) the geometric difference of two circles. The geometric difference is shown by dashed lines. Thin lines denote some extreme lays of the subtrahend set.

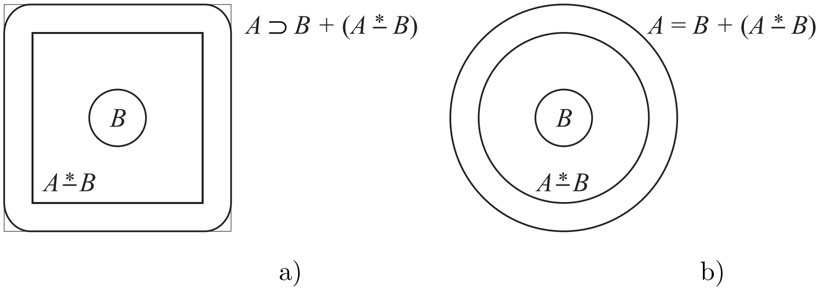


Figure 2: Pictures of the summation of the geometric difference and the subtrahend set for the examples in Figure 1.

of the above examples, after such a summation a square with round corners is obtained (Figure 2a). In the example b), such a summation gives exactly the original circle (Figure 2b).

Definition 1.2. The situation, when the equality

$$B + (A \overset{*}{-} B) = A$$

holds, is called the *complete sweeping* of the set A by the set B.

The notion of "complete sweeping" was originally introduced in [4]. The preceding example a) shows the possibility of absence of the complete sweeping property, whereas example b) shows its possible presence.

As a good illustrative analogy, one can imagine the set A as a room and the set B as a broom. So, the situation of complete sweeping corresponds to a good hostess who sweeps the whole room and does not miss any corner.

Let us give an equivalent definition of the complete sweeping.

Definition 1.3. A set A is completely swept by a set B if $\forall a \in A \, \exists \, x : 1)\, a \in B + x$ and $2)\, B + x \subset A$.

Let M_c be the level set (the Lebesgue set) of a function f corresponding to a constant c: $M_c = \{x : f(x) \leqslant c\}$.

Definition 1.4. A function f possesses the *level sweeping* property if for any pair of constants $c_1 < c_2$ such that $M_{c_1} \neq \varnothing$, the set M_{c_1} sweeps completely the set M_{c_2}, that is, the relation $M_{c_2} = M_{c_1} + (M_{c_2} \overset{*}{-} M_{c_1})$ holds.

Note that the convexity of a function is neither necessary nor sufficient for presence of the level sweeping property. This is demonstrated by the example shown in Figure 3. Here we consider a function whose graph is a hemisphere cut by two planes such that some smaller level set is a circle and some greater one is a circle with a "roof." It is evident that the smaller level set does not completely sweep the greater one: the corner of the latter cannot be covered.

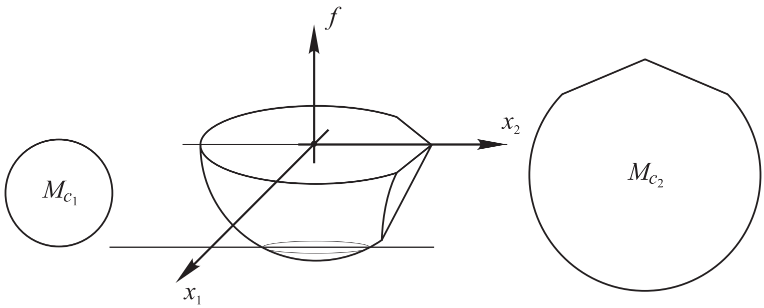


Figure 3: Example of a convex function which does not possess the level sweeping property.

2 Description of the Main Result

Let us consider a linear antagonistic differential game

$$\begin{aligned} &\dot{x} = A(t)x + B(t)u + C(t)v, \quad t \in [t_0, T],\ x \in \mathbb{R}^n,\ u \in P,\ v \in Q, \\ &\varphi\big(x_i(T), x_j(T)\big) \to \min_u \max_v \end{aligned} \tag{1}$$

with fixed terminal time T, convex compact constraints P, Q for controls of the first and second players, and continuous quasi-convex payoff function φ depending on two components x_i, x_j of the phase vector x at the terminal time. (A function is quasi-convex if each of its level sets (Lebesgue sets) is convex.) The first player minimizes the payoff, and the interests of the second one are opposite. It is assumed that every level set $M_c = \big\{(x_i, x_j) : \varphi(x_i, x_j) \leqslant c\big\}$ of the payoff function φ is bounded in the coordinates x_i, x_j.

Using a change of variable $y(t) = X_{i,j}(T, t)x(t)$ ([7, p. 354], [8, pp. 89–91]), which is provided by a matrix combined of two rows of the fundamental Cauchy matrix of system (1), one can pass to the equivalent game

$$\begin{aligned} &\dot{y} = D(t)u + E(t)v, \\ &t \in [t_0, T], \quad y \in \mathbb{R}^2, \quad u \in P, \quad v \in Q, \quad \varphi\big(y_1(T), y_2(T)\big), \\ &D(t) = X_{i,j}(T, t)B(t), \quad E(t) = X_{i,j}(T, t)C(t). \end{aligned} \tag{2}$$

Here, the new phase variable y is two dimensional. The right-hand side of the dynamics does not contain the phase variable. The game interval, the constraints for controls, and the payoff function are the same as in the original game (1) (except that the payoff function now depends on components of the vector y).

Let $(t, y) \mapsto V(t, y)$ be the value function of the differential game (2). The function V is continuous. For any $t \in [t_0, T]$, the function $y \mapsto V(t, y)$ is quasi-convex with compact level sets.

Suppose that the payoff function φ possesses the level sweeping property, that is, for two arbitrary constants $c_1 < c_2$ the corresponding level sets M_{c_1}

and M_{c_2} of the function φ (such that $M_{c_1} \neq \varnothing$) obey the relation

$$M_{c_2} = M_{c_1} + (M_{c_2} \overset{*}{-} M_{c_1}). \tag{3}$$

It turns out that the value function inherits the level sweeping property from the payoff function. Namely, let $W_c(t) = \{y : V(t,y) \leqslant c\}$ be a time section at the instant t of the level set $W_c = \{(t,y) : V(t,y) \leqslant c\}$ of the value function V. In the paper, it is shown that the relation (3) with an additional condition $W_{c_1}(t) \neq \varnothing$, $t \in [t_0, T]$, gives

$$W_{c_2}(t) = W_{c_1}(t) + \big(W_{c_2}(t) \overset{*}{-} W_{c_1}(t)\big), \qquad t \in [t_0, T]. \tag{4}$$

The main result can be reformulated in the following way.

Theorem 2.1. *If the payoff function of the game* (2) *is such that any of its smaller level sets completely sweeps any larger one, then the time sections of level sets of the value function at any fixed time instant t from the game interval have the same property.*

Since the sections of a level set of the value function in the original and equivalent coordinates are connected by the relation $\mathcal{W}_c(t) = \{x \in \mathbb{R}^n : X_{i,j}(T,t)x \in W_c(t)\}$, $t \in [t_0, T]$, the statement about inheritance of the level sweeping property by the value function from the payoff function is also true for the original game (1). In this form, the fact was formulated in the abstract.

3 Backward Procedure for Constructing Level Sets

To prove the theorem, now a backward procedure will be described, which constructs approximately a level set of the value function in game (2). A level set corresponding to a number c is built as a collection of time sections $\{\boldsymbol{W}_c(t_i)\}$ in a grid of instants $\{t_i\}$. Here, the bold notation $\boldsymbol{W}$ is used instead of W to emphasize that approximate sets are used. Construction is started from a level set M_c of the payoff function taken at the terminal instant T. The set M_c is processed by means of a procedure to the instant $T - \Delta$ giving the section $\boldsymbol{W}_c(T - \Delta)$. Then by means of the same procedure on the basis of the set $\boldsymbol{W}_c(T - \Delta)$, a new set $\boldsymbol{W}_c(T - 2\Delta)$ is computed for the instant $T - 2\Delta$, and so on until the given time $t_* \in [t_0, T)$ (Figure 4).

The procedure for constructing a section $\boldsymbol{W}_c(t_i)$ uses the previous section $\boldsymbol{W}_c(t_{i+1})$ of the level set, the matrices $D(t_i)$ and $E(t_i)$ from the game dynamics (2), and the sets P and Q constraining the players' controls. It is described by the following formula [14,15,9]:

$$\boldsymbol{W}_c(t_i) = \Big(\boldsymbol{W}_c(t_{i+1}) + \Delta\big(-D(t_i)P\big)\Big) \overset{*}{-} \Delta E(t_i)Q. \tag{5}$$

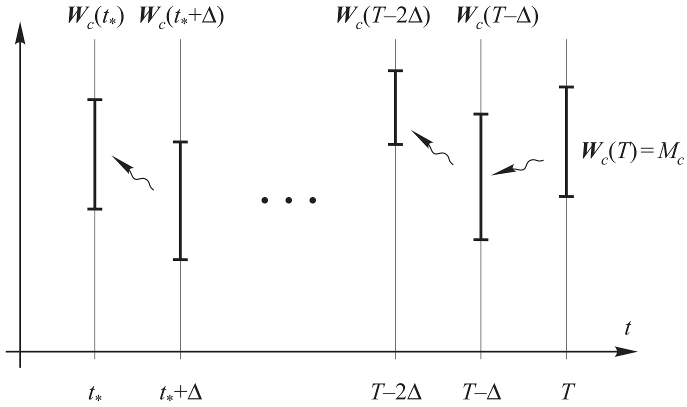


Figure 4: Scheme of the backward procedure of constructing a level set of the value function.

Suppose that $\operatorname{int} W_c(t) \neq \varnothing$ for any $t \in [t_*, T]$. Here, $\operatorname{int} A$ means the interior of a set A. It is known that when decreasing the step size Δ of the discrete scheme, the approximately built section $\boldsymbol{W}_c(t_*)$ of a level set converges to the ideal one $W_c(t_*)$ in the Hausdorff metric [12,1,11].

So, to prove the inheritance of the level sweeping property by the value function it is necessary to prove that the property of complete sweeping is conserved after operations of algebraic sum and geometric difference and after passing to the limit when decreasing the step size Δ.

4 Additional Properties of the Geometric Difference

The following statement concerns the conservation of the complete sweeping property after the operations of algebraic sum and geometric difference.

Lemma 4.1. *Let convex compact sets A, B, and C in the plane be such that the set A is completely swept by the set B, that is, $A = B + (A \overset{*}{-} B)$. Then*
1) $(A + C) = (B + C) + \big((A + C) \overset{*}{-} (B + C)\big)$;
2) *if $B \overset{*}{-} C \neq \varnothing$, then $(A \overset{*}{-} C) = (B \overset{*}{-} C) + \big((A \overset{*}{-} C) \overset{*}{-} (B \overset{*}{-} C)\big)$.*

Proof. The first fact is proved directly with the help of equivalent Definition 1.3 of the complete sweeping. So, let us show that for any $a' \in A + C$ there is an element $x \in \mathbb{R}^2$ such that $a' \in (B + C) + x$ and $(B + C) + x \subset (A + C)$.

Fix $a' \in A + C$. Then one can find $a \in A$ and $c \in C$ such that $a' = a + c$. According to the complete sweeping of the set A by the set B, there is an element $x \in \mathbb{R}^2$ such that $a \in B + x$ and $B + x \subset A$. Prove that this element x is also acceptable for establishing the complete sweeping of the set $A + C$ by the set $B + C$.

Since $a \in B + x$, it follows that $a + c = a' \in B + x + c \subset (B + C) + x$.

Because $B + x \subset A$, then $(B + C) + x \subset (A + C)$.

So, the conservation of the complete sweeping after the algebraic sum is proved. Note that this proof does not demand any compactness, or convexity, or dimension restriction of the sets A, B, and C. Therefore, statement 1) of Lemma 4.1 also holds under more general conditions.

Let us proceed to statement 2) of Lemma 4.1. We use the support functions of the sets under consideration. Recall that every convex compact set A produces a finite positively homogeneous convex function by the formula $\rho_A(l) = \max\big\{l'a : a \in A\big\}$. This function is called the *support function* of the set A. And vice versa, for any finite positively homogeneous convex function ρ, a convex compact set can be found such that ρ is its support function [16].

Let us establish a correspondence between set operations and operations over support functions. Let $A \leftrightarrow \rho_A$, $B \leftrightarrow \rho_B$. Then $\rho_{A+B} = \rho_A + \rho_B$. It is also known that if $A \overset{*}{-} B \neq \varnothing$, then $\rho_{A \overset{*}{-} B} = \operatorname{conv}\{\rho_A - \rho_B\}$ [2,9]. When $A \overset{*}{-} B = \varnothing$, it is supposed that $\rho_{A \overset{*}{-} B} \equiv -\infty$.

Let the set A be completely swept by the set B, that is, $A = B + (A \overset{*}{-} B)$. Then $\rho_A = \rho_B + \operatorname{conv}\{\rho_A - \rho_B\}$, or $\rho_A - \rho_B = \operatorname{conv}\{\rho_A - \rho_B\}$. Hence, if the set A is completely swept by the set B, then the difference of their support functions is convex.

Using the language of support functions, the statement about conservation of the complete sweeping property after the geometric difference can be formulated as follows.

2^*) Let some convex compact sets A, B, and C be such that the difference $\rho_A - \rho_B$ is convex and the function $\operatorname{conv}\{\rho_B - \rho_C\}$ has finite value everywhere in $\mathbb{R}^2$. Then the difference $\operatorname{conv}\{\rho_A - \rho_C\} - \operatorname{conv}\{\rho_B - \rho_C\}$ is also convex.

Assume $f = \rho_A - \rho_C$, $g = \rho_B - \rho_C$.

The function $f - g = (\rho_A - \rho_C) - (\rho_B - \rho_C) = \rho_A - \rho_B$ is convex. Convexity of the function $\operatorname{conv} f - \operatorname{conv} g = \operatorname{conv}\{\rho_A - \rho_C\} - \operatorname{conv}\{\rho_B - \rho_C\}$ is shown in the next lemma. $\square$

Lemma 4.2. *Let functions f and $g : \mathbb{R}^2 \to \mathbb{R}$ be positively homogeneous, continuous, the difference $f - g$ be convex, and the function* $\operatorname{conv} g$ *have finite value everywhere in $\mathbb{R}^2$. Then the difference* $\operatorname{conv} f - \operatorname{conv} g$ *is a convex function.*

Before the proof of Lemma 4.2, let us formulate some auxiliary propositions. They are quite simple, so no proofs are given.

Let us denote the boundary of a set D by ∂D. Restriction of f to a set D will be written as $f\big|_D$. By $\operatorname{conv}\big|_D f$ we mean the convex hull of the function f computed in a convex set D.

1° Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. Also let $D \subset \mathbb{R}^n$ be a closed convex set and let the function $\tilde{f}$ be convex in the set D. Let us suppose that $\tilde{f}(x) = f(x)$

when $x \in \partial D$ and $\tilde{f}(x) \geqslant f(x)$ when $x \in \operatorname{int} D$. Then the function

$$g(x) = \begin{cases} \tilde{f}(x), & x \in D, \\ f(x), & x \notin D \end{cases}$$

is convex in $\mathbb{R}^n$.

2° Let $f : \mathbb{R}^n \to \mathbb{R}$ and $D \subset \mathbb{R}^n$ be a closed convex set. Let us suppose that $(\operatorname{conv} f)(x) = f(x)$ when $x \in \partial D$. Then $\operatorname{conv}\big|_D f = (\operatorname{conv} f)\big|_D$.

3° Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous, positively homogeneous function. Then for any vector $l_* \neq 0$ a vector $p \in \{x : l'_* x \geqslant 0\}$ exists such that $f(p) = (\operatorname{conv} f)(p)$.

4° Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a continuous, positively homogeneous function, and let C be a closed cone of angle not greater than π. Let us suppose that $f(x) = (\operatorname{conv} f)(x)$ if $x \in \partial C$ and $f(x) \neq (\operatorname{conv} f)(x)$ if $x \in \operatorname{int} C$. Then the function $\operatorname{conv} f$ is linear in the cone C.

Now, Lemma 4.2 will be proved.

Proof. 1) Let us denote $\tilde{g} = \operatorname{conv} g$, $S = \{x \in \mathbb{R}^2 : \tilde{g}(x) = g(x)\}$. By virtue of the continuity of the functions $\tilde{g}$ and g, the set S is closed. Thus, the set $\mathbb{R}^2 \backslash S$ can be presented as at most a countable join of non-overlapping open cones C_i^0, $i = \overline{1, m}$, $m \leqslant \infty$. Following proposition 3°, each of these cones is of angle not greater than π. Let C_i be the closure of the cone C_i^0.

Using proposition 2°, one can establish that for any i, the equality $\operatorname{conv}\big|_{C_i} g = (\operatorname{conv} g)\big|_{C_i}$ holds.

2) The process of constructing the convex hull of the function g can be considered as a stepwise one: $g = g_0 \rightsquigarrow g_1 \rightsquigarrow g_2 \rightsquigarrow \ldots$ Here, each next function g_i is obtained from the previous one g_{i-1} by changing the latter in the cone C_i by a linear function l_i. One has $l_i(x) = g_{i-1}(x)$ when $x \in \partial C_i$ and $l_i(x) < g_{i-1}(x)$ when $x \in \operatorname{int} C_i$. Also according to proposition 4°, $l_i = (\operatorname{conv} g)\big|_{C_i}$.

Simultaneously, the function f is also corrected: $f = f_0 \rightsquigarrow f_1 \rightsquigarrow f_2 \rightsquigarrow \ldots$ such that $f_i\big|_{C_i} = \operatorname{conv}\big|_{C_i} f_{i-1}$, $f_i\big|_{\mathbb{R}^2 \backslash C_i} = f_{i-1}\big|_{\mathbb{R}^2 \backslash C_i}$. That is, f_i is obtained from f_{i-1} by convexification of the latter in the cone C_i.

3) Let $h_i = f_i - g_i$, $i \geqslant 0$. We will prove by induction on i that for any i the function h_i is convex.

When $i = 0$, the function $h_0 = f_0 - g_0 = f - g$ is convex by the condition of the lemma.

Suppose that for any $0 \leqslant i - 1 < m$, the function h_{i-1} is convex. We will show that in this case the function h_i is also convex.

When $x \in \mathbb{R}^2 \backslash C_i$, one has $g_i(x) = g_{i-1}(x)$ and $f_i(x) = f_{i-1}(x)$. Therefore, $h_i = h_{i-1}$ in $\mathbb{R}^2 \backslash C_i$.

We have $g_i(x) \leqslant g_{i-1}(x)$ when $x \in C_i$. Thus, in the cone C_i the relation $f_{i-1} - g_i \geqslant f_{i-1} - g_{i-1} = h_{i-1}$ holds, and, therefore, $f_{i-1} \geqslant g_i + h_{i-1}$. Because g_i is linear in C_i, then the sum $g_i + h_{i-1}$ is convex in C_i. Consequently, it follows that in C_i the relation $f_i = \operatorname{conv}\big|_{C_i} f_{i-1} \geqslant g_i + h_{i-1}$ holds, that is, $h_i = f_i - g_i \geqslant h_{i-1}$.

Since in the cone C_i the function f_i is convex and g_i is linear, the function $h_i = f_i - g_i$ is convex in C_i.

Applying proposition 1°, one obtains that the function h_i is convex in $\mathbb{R}^2$.

4) The sequence of the continuous functions g_i is nonincreasing. With that $\lim g_i = \operatorname{conv} g$. The sequence of the continuous functions f_i is nonincreasing and is bounded from below by the function $\operatorname{conv} f$. Thus, this sequence has a pointwise limit $\tilde{f}$. The sequence of convex functions h_i converges pointwise to a convex function $\tilde{h} = \tilde{f} - \operatorname{conv} g$. Hence, the function $\tilde{f} = \tilde{h} + \operatorname{conv} g$ is convex.

Let us prove that $\tilde{f} = \operatorname{conv} f$. One has that $\tilde{f}(x) = f(x) \geqslant (\operatorname{conv} f)(x)$ when $x \in S$. For any $x \in \mathbb{R}^2 \backslash S$ an index $i \geqslant 1$ exists such that $x \in C_i$, and, therefore,

$$\tilde{f}(x) = f_i(x) = \left(\operatorname{conv}\big|_{C_i} f_{i-1}\right)(x) = \left(\operatorname{conv}\big|_{C_i} f\right)(x) \geqslant (\operatorname{conv} f)(x).$$

Hence, $\tilde{f} \geqslant \operatorname{conv} f$. Because $f \geqslant \tilde{f}$ and the function $\tilde{f}$ is convex, then $\tilde{f} = \operatorname{conv} f$.

By this, it is shown that the difference $\operatorname{conv} f - \operatorname{conv} g$ is convex in $\mathbb{R}^2$. $\square$

5 Counterexamples to Generalizations of Lemma 4.2

Note that Lemma 4.2 holds only for positive homogeneous functions of two variables. Generally speaking, the lemma does not hold if the function does not possess positive homogeneity or the dimension of its argument is higher than two.

Let us show this by some counterexamples. At first, an example of convex compact three-dimensional sets A, B, and C will be given such that the set B completely sweeps the set A, but the difference $B \overset{*}{-} C$ does not completely sweep the set $A \overset{*}{-} C$. Let us take the set A as a hemisphere cut by two planes (Figure 5). The set B is homothetic to the set A with coefficient of homothety less than 1. The set C is taken as an interval, where the length is less than the horizontal side of the cut part of the set A, but larger than the cut part of the set B.

Since the set C is an interval, the geometric difference $B \overset{*}{-} C$ ($A \overset{*}{-} C$) is the intersection of two copies of the set B (correspondingly, A) shifted by the length of the interval C. According to this, the difference $B \overset{*}{-} C$ looks like a cap: the cut part disappeared. At the same time, the difference $A \overset{*}{-} C$ keeps the cut part. The sections of the flat sides of the geometric differences are shown at the right in Figure 5. It is evident that the sharp point of the "roof" of the set $A \overset{*}{-} C$ cannot be covered by the circle $B \overset{*}{-} C$. Therefore, there is no complete sweeping between the sets $A \overset{*}{-} C$ and $B \overset{*}{-} C$.

Thus, a counterexample for a possible generalization of statement 2) of Lemma 4.1 is constructed for the case when the sets A, B, C are of dimension

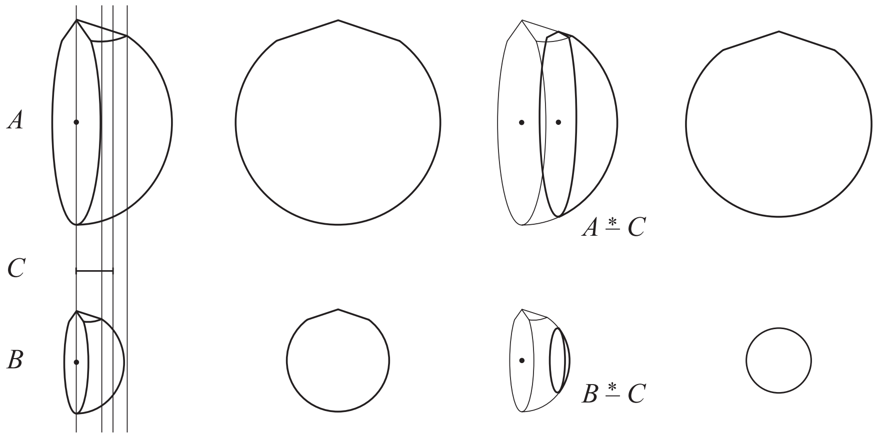


Figure 5: Counterexample for conservation of the complete sweeping after the operation of geometric difference of three-dimensional sets.

higher than two. Support functions of the sets considered give a counterexample for a generalization of statement 2*) and, therefore, for Lemma 4.2 in the case when the positively homogeneous functions have their arguments of dimension three or higher.

Violation of Lemma 4.2 in the case of functions of the general kind (not positively homogeneous) is demonstrated by the following example.

Let the functions f and g be piecewise linear. The graph of the function f can be obtained from a quadrahedral pyramid by cutting it by two planes parallel to the diagonal of the base (Figure 6a). Something looking like a "chisel" appears.

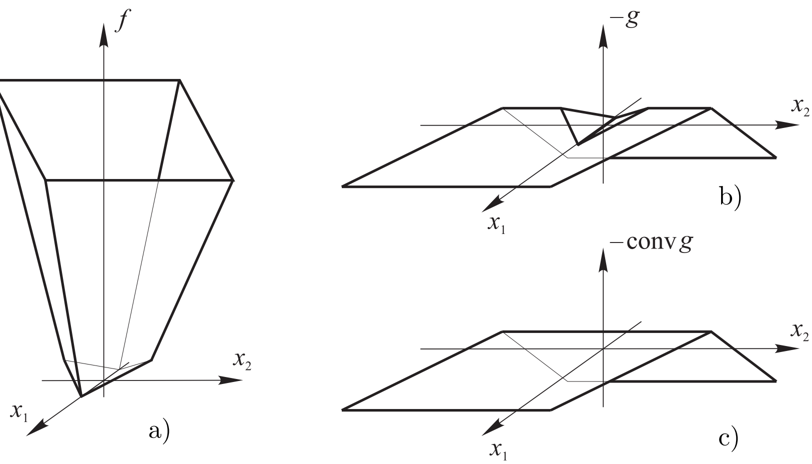
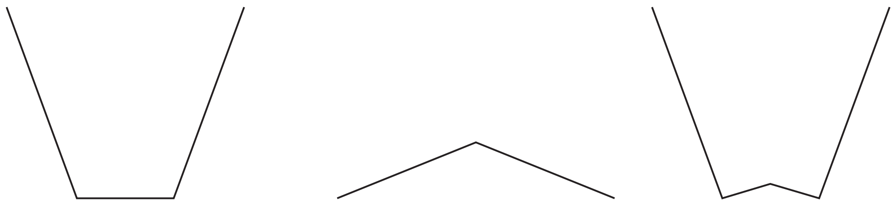


Figure 6: Graphs of the functions f (a), $-g$ (b), and $-\operatorname{conv} g$ (c).



a) b) c)

Figure 7: Sections of the graphs of $\operatorname{conv} f = f$ (a), $-\operatorname{conv} g$ (b), and $\operatorname{conv} f - \operatorname{conv} g$ (c).

The graph of the function $-g$ (it is more demonstrative to imagine the function $-g$) looks like a “roof” having a cavity of the same form as the bottom of the graph of f (Figure 6b). The origin is placed at the middle of the bottom of the graph of f and at the middle of the cavity of $-g$. Then the graph of $f - g = f + (-g)$ looks like the graph of f. The slope of the bottom outshoot becomes “sharper” and the slope of the side faces becomes, conversely, “flatter” in comparison with the graph of f. The original slopes can be chosen such that the graph of $f - g$ will be convex. (Namely, it is necessary to take the side faces of f quite “sharp” and the faces of g and the bottom outshoot of f quite “flat.”)

Let us consider the graph of the function $\operatorname{conv} f - \operatorname{conv} g = f + (-\operatorname{conv} g)$. The convex hull $\operatorname{conv} f$ coincides with f itself because the function f is convex. The graph of $-\operatorname{conv} g$ (or of the concave hull of $-g$) looks like a “roof” without any cavities (Figure 6c). Let us take the sections of the graphs made by a vertical plane containing the bottom line of “chisel” f. Since the section of the function $\operatorname{conv} f - \operatorname{conv} g$ is non-convex (Figure 7), the function $\operatorname{conv} f - \operatorname{conv} g$ itself is non-convex.

6 Conservation of Level Sweeping to the Limit

Fix an arbitrary instant $t_* \in [t_0, T)$ and choose a sequence $\{\vartheta_k\}$ of subdivisions of the time interval $[t_*, T]$: $\vartheta_k = \{t_* = t_*^{(k)} < \cdots < t_{N_k}^{(k)} = T\}$. With $k \to 0$ diameter Δ_k of subdivision ϑ_k goes to 0. Denote by $\boldsymbol{W}_{c_1}^{(k)}(t_*)$ and $\boldsymbol{W}_{c_2}^{(k)}(t_*)$ the results of applying the backward procedure (5) on the subdivision ϑ_k with starting sets $\boldsymbol{W}_{c_1}(T) = M_{c_1}$ and $\boldsymbol{W}_2(T) = M_{c_2}$.

Because the starting sets $\boldsymbol{W}_{c_1}(T)$ and $\boldsymbol{W}_{c_2}(T)$ have the complete sweeping, then according to the results on conservation of the complete sweeping after algebraic sum and geometric difference from Section 4, each pair of sets $\boldsymbol{W}_{c_1}^{(k)}(t_i)$ and $\boldsymbol{W}_{c_2}^{(k)}(t_i)$ has the complete sweeping. Consequently, for any k the set $\boldsymbol{W}_{c_1}^{(k)}(t_*)$ completely sweeps the set $\boldsymbol{W}_{c_2}^{(k)}(t_*)$.

1) Under the assumption that for any $t \in [t_*, T]$ the section $W_{c_1}(t)$ of ideal level set W_{c_1} of the value function has a non-empty interior (that is, $\operatorname{int} W_{c_1}(t) \neq \varnothing$),

one has the following convergence $\boldsymbol{W}_{c_1}^{(k)}(t_*) \to W_{c_1}(t_*)$ and $\boldsymbol{W}_{c_2}^{(k)}(t_*) \to W_{c_2}(t_*)$ in the Hausdorff metric with $k \to \infty$.

Therefore, to prove the complete sweeping of the set $W_{c_2}(t_*)$ by the set $W_{c_1}(t_*)$ under the additional condition $\operatorname{int} W_{c_1}(t) \neq \varnothing$, $t \in [t_*, T]$, it is necessary to justify the following simple fact. Let two sequences $\{A_k\}$ and $\{B_k\}$ of compact sets converge in the Hausdorff metric to compact sets A and B respectively. Suppose that for any k the set B_k completely sweeps the set A_k. Then the limit sets have the same property: the set B completely sweeps the set A.

Let us show that for the sets A and B, the properties, which stipulate the complete sweeping of the first set by the second one, hold: 1) $\forall a \in A\ \exists\, x : a \in B + x$ and 2) $B + x \subset A$ (see Definition 1.3).

Fix an arbitrary element $a \in A$. Due to the convergence $A_k \to A$, one can choose a sequence $\{a_k\}$, $a_k \in A_k$, such that $a_k \to a$. Since the set A_k is completely swept by the set B_k, it implies $\forall k\ \exists x_k : a_k \in B_k + x_k$ and $B_k + x_k \subset A_k$.

Consider the sequence $\{x_k\}$. It is bounded. Therefore, a converging subsequence can be extracted from it. Without loss of generality, let us suppose that the sequence $\{x_k\}$ itself converges to an element x. This limit is just the desired element, which figures in the properties giving the complete sweeping. Let us show this fact.

The first property: $a \in B + x$. We have that $\forall k\ a_k \in B_k + x_k$. Choose $b_k \in B_k : a_k = b_k + x_k$. Since $a_k \to a$ and $x_k \to x$, it follows $b_k \to b = a - x$. Taking into account the convergence $B_k \to B$, one can obtain that $b \in B$. Therefore, there is an element $b \in B$ such that $a = b + x$. Consequently, $a \in B + x$.

The second property: $B + x \subset A$. Let us take an arbitrary element $b \in B$. Due to the convergence $B_k \to B$, one can take a sequence $\{b_k\}$, $b_k \in B_k$, such that $b_k \to b$. Since $B_k + x_k \subset A_k$, it implies $b_k + x_k \in A_k$. Therefore, $\forall k\ \exists\, a_k \in A_k : b_k + x_k = a_k$. Because $b_k \to b$ and $x_k \to x$, then a_k tends to an element $\bar{a} = b + x$. Taking into account the convergence $A_k \to A$, one can obtain that $\bar{a} \in A$. This shows that $\forall\, b \in B\ \ b + x \in A$. Consequently, $B + x \subset A$.

Hence, the set B completely sweeps the set A.

2) Now let $W_{c_1}(t_*) \neq \varnothing$, but $\operatorname{int} W_{c_1}(\bar{t}) = \varnothing$ at an instant $\bar{t} \in [t_*, T]$. From the continuity of the value function, it follows that $\operatorname{int} W_c(\bar{t}) \neq \varnothing$ for $c > c_1$. Then also $\operatorname{int} W_c(t) \neq \varnothing$ for $c > c_1$ when $t \in [t_*, T]$. According to the fact proved above, the set $W_c(t_*)$ completely sweeps the set $W_{c_2}(t_*)$ for $c \in (c_1, c_2)$. It follows from this that the set $W_{c_1}(t_*)$ completely sweeps the set $W_{c_2}(t_*)$.

7 Is It Possible to Weaken the Dimension Assumption?

Theorem 2.1 is formulated for the case when the payoff function φ depends on two components of the phase vector at the terminal instant T. Let us show that, generally speaking, the theorem does not hold if the payoff function is defined by three or more components of the phase vector.

Let us consider a differential game

$$\begin{aligned} &\dot{x} = u + v, \quad t \in [t_0, T],\ x \in \mathbb{R}^3,\ u \in \{0\},\ v \in Q, \\ &\varphi\big(x(T)\big) = \min\big\{c : x(T) \in cM\big\} \end{aligned} \tag{6}$$

with fixed terminal time T, a fictitious first player (actually, the first player is absent) and the payoff function, which is the Minkowski function of a compact convex set M. The set M is taken as the set A shown in Figure 5. The payoff function depends on full a three-dimensional phase vector and, evidently, possesses the level sweeping property. As the set Q constraining the control of the second player, let us take the interval shown in Figure 5 and denoted there by C.

Because the right-hand side of the game dynamics does not depend on time and does not contain the phase variable, then for any t and any c the section $W_c(t)$ of the level set of the value function is defined by the formula $W_c(t) = W_c(T) \overset{*}{-} (T-t)Q$. Let $t = T - 1$. Take $c_2 = 1$ and $c_1 < 1$ such that the set $M_{c_1} = c_1 M$ coincides with the set B drawn in Figure 5. Then $W_{c_1}(t) = M_{c_1} \overset{*}{-} Q = B \overset{*}{-} C$ and $W_{c_2}(t) = M_{c_2} \overset{*}{-} Q = A \overset{*}{-} C$. As described in Section 5 in the text relating to Figure 5, the set $A \overset{*}{-} C$ is not completely swept by the set $B \overset{*}{-} C$. Therefore, the set $W_{c_2}(t)$ is not completely swept by the set $W_{c_1}(t)$.

Thus, the condition of Theorem 2.1 connected to the number of arguments of the payoff function is essential.

8 Conclusion

In this chapter, a linear antagonistic differential game with fixed terminal time, geometric constraints on the players' controls, and continuous quasi-convex terminal payoff function depending on two components of the phase vector is considered. A level sweeping property of a quasi-convex function is defined. This property consists of the condition that any non-empty smaller level set completely sweeps any larger one. The term "complete sweeping" is based on the concept of geometric difference (Minkowski difference) and is known in convex analysis and in differential game theory. It is proved that, in the game class considered, the level sweeping property is inherited by the value function. That is, if the payoff function possesses the level sweeping property, then the same property is true for the constriction of the value function to any time instant from the game interval. It is shown (by a counterexample) that this holds only when the payoff function depends on at most two components of the phase vector.

The level sweeping property of the value function can be useful, for example, when analyzing singular surfaces appearing in linear differential games with fixed terminal time. Namely, under the presence of this property, the structure of singular surfaces has some patterns absent in the general situation. In this case, numerical algorithms for constructing and classifying singular surfaces become essentially easier.

REFERENCES

[1] Botkin N.D., Evaluation of numerical construction error in differential game with fixed terminal time, *Problems of Control and Information Theory*, 11, 4, 283–295, 1982.

[2] Chikrii A.A., *Conflict controlled processes*, Naukova dumka, Kiev, 1992 (in Russian).

[3] Grigorenko N.L., *Mathematical methods for control of a number of dynamic processes*, Moscow State University, Moscow, 1990 (in Russian).

[4] Gusjatnikov P.B. and Nikol'skii M.S., Pursuit time optimality, *Doklady Akad. Nauk SSSR*, 184, 3, 518–521, 1969 (in Russian); Transl. as *Sov. Math. Doklady*, 10, 103–106, 1969.

[5] Hadwiger H., *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie*, Springer-Verlag, Berlin, 1957.

[6] Isaacs R., *Differential Games*, John Wiley & Sons, New York, 1965.

[7] Krasovskii N.N., *Games Problems about Contact of Motions*, Nauka, Moscow, 1970 (in Russian); Transl. as *Rendez-vous Game Problems*, Nat. Tech. Inf. Serv., Springfield, VA, 1971.

[8] Krasovskii N.N. and Subbotin A.I., *Game-Theoretical Control Problems*, Springer-Verlag, New York, 1988.

[9] Kurzhanski A.B. and Valyi I., *Ellipsoidal Calculus for Estimation and Control*, Birkhäuser, Boston, 1997.

[10] Nikol'skii M.S., *The First Direct Method of L.S. Pontryagin in Differential Games*, Moscow State University, Moscow, 1984 (in Russian).

[11] Polovinkin E.S., Ivanov G.E., Balashov M.V., Konstantiov R.V., and Khorev A.V., An algorithm for the numerical solution of linear differential games, *Matematicheskiĭ sbornik*, 192, 10, 95–122, 2001 (in Russian); Transl. as *Sbornik: Mathematics*, 192, 10, 1515–1542, 2001.

[12] Ponomarev A.P. and Rozov N.H., Stability and convergence of alternating Pontryagin sums, *Vestnik Moskov. Univ., Ser. XV Vyčisl. Mat. Kibernet.*, 1, 82–90, 1978 (in Russian).

[13] Pontryagin L.S., On linear differential games, 1, *Doklady Akad. Nauk SSSR*, 174, 6, 1278–1280, 1967 (in Russian); Transl. as. *Sov. Math. Doklady*, 8, 769–771, 1967.

[14] Pontryagin L.S., On linear differential games, 2, *Doklady Akad. Nauk SSSR*, 175, 4, 764–766, 1967 (in Russian); Transl. as *Sov. Math. Doklady*, 8, 910–912, 1967.

[15] Pschenichnyi B.N. and Sagaidak M.I., Differential games of prescribed duration, *Kibernetika*, 2, 54–63, 1970 (in Russian); Transl. as *Cybernetics*, 6, 2, 72–80, 1970.

[16] Rockafellar R.T., *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

Optimal Feedback in a Dynamic Game of Generalized Shortest Path*

Vyacheslav P. Serov
Ural Polytechnic Institute
Ekaterinburg, Russia
serov@imm.uran.ru

Abstract

We consider the problem of generalized shortest path. The task is to transit optimally from the origin through a system M_i, $i \in \overline{1,m}$, of intermediate sets in $\mathbb{R}^d$ to a fixed destination point (or set), under conditions that only one node in M_i can be chosen for passing. Any returns to the sets that have already been passed, are prohibited. The (combinatorial) cost function to minimize is either additive or bottleneck. The visiting nodes $x_i \in M_i$, $i \in \overline{1,m}$, are either governed by an antagonistic nature or by a rational antagonist. For this multistage game problem both open-loop and feedback settings are suggested. The feedback problem is posed in the class of feedback strategies which can change route during motion, depending on the current moves of the opponent. They provide, in general, a strictly better value of the problem, with respect to the open-loop minimax setting. The optimal feedback minimax strategy is constructed, and some (polynomial) heuristics are given.

Key words. Game generalized shortest path problem, feedback strategy

1 Introduction

The shortest path problem (SPP) is one of the main problems of combinatorial optimization, and it has many applications. The classical one-to-one SPP is to find the shortest path from a fixed origin x_0 to a fixed destination $x_f \triangleq x_{m+1}$ through a network of given nodes x_i, $i \in \overline{1,m}$ ($\overline{1,m} \triangleq \{1, 2, \ldots, m\}$). The cost of transition from x_i to x_j is denoted by c_{ij}. The cost functional to minimize is usually $\sum_i c_{r_{i-1}r_i}$ (additive) or $\max_i c_{r_{i-1}r_i}$ (bottleneck), where $(r_0, r_1, \ldots, r_s, m+1)$, $r_0 \triangleq 0$, is a route of size s, $s \leqslant m$. The SPP has a vast literature (see, e.g., the surveys [5,12,14,15,22,23,28,64]).

*This research was supported by the RF Ministry of Education under Grant E02-1.0-232.

We reformulate the SPP as an optimal open-loop (program) control problem where the set of admissible controls is the set of all possible routes of all possible sizes s, $s \leqslant m$. In [56] the generalized shortest path problem (GSPP) to transit optimally from x_0 to x_f through a system M_i, $i \in \overline{1,m}$, of given intermediate sets was considered. The nodes x_i within given sets $M_i \subset \mathbb{R}^d$ ($i \in \overline{1,m}$) were allowed to vary so as to minimize the total costs of transition. The GSPP can be regarded as an optimal open-loop control problem where admissible controls are pairs $\{r(s),\, (x_i)_{i\in\overline{1,s}}\}$, $x_i \in M_{r_i}$, $i \in \overline{1,s}$, $s \leqslant m$.

In this chapter we introduce the game (GGSPP) as a variant of the GSPP. Here, a control pair $\{r(s),\, (x_i)_{i\in\overline{1,s}}\}$ consists of controls: $r(s)$ for player 1 and $(x_i)_{i\in\overline{1,s}}$ for his antagonist, player 2 (who maximizes the cost). Further, for the case of the dynamic GGSPP we pose this routing problem under conditions of (set membership) uncertainties in the class of feedback strategies, analogous to those suggested in [55] for the game of the dynamic generalized traveling salesman problem (TSP). This formulation exploits ideas from control theory and dynamic games ([2,4,19,20,24,27,30,31,33,44,46,49,58]), such as open-loop control, feedback (positional) strategy, motion generated by strategy, multiple minimax, and others.

2 GSPP

In this section we consider the GSPP [56], and give some notation.

Put $M_0 \triangleq \{x_0\}$, $M_{m+1} \triangleq \{x_f\}$. The directed graph of the problem is given in the form of arc lists [7,12,23]. That is, for each number i of a set M_i, $i \in \overline{0,m}$, the arc list FS(i) ("forward star" of i) is given of those arcs (i,j) for which the transition from M_i to M_j or to x_f is allowed. Let A be a set of (directed) arcs, $A \subseteq \{(i,j) : i \in \overline{0,m},\ j \in \overline{1,m+1}\}$. Let n denote the cardinality $|A|$ of A. For sparse graphs (which are typical for applications), one has $n \ll m^2$.

For $x \in M_i$, $y \in M_j$ let $c_j(x,y)$ and $c_{m+1}(x,x_f)$ be the appropriate cost functions. In general, $c_j(x,y) \neq c_i(y,x)$. Note that the direct transition $x_0 \to x_f$ is also possible. The sets M_i are assumed to be just arbitrary nonempty subsets of $\mathbb{R}^d$. The functions c_j and c_{m+1} are any scalar functions (in particular, c_j and c_{m+1} may be discontinuous and may take negative values).

Revisiting sets M_i already passed is forbidden, and only one node in every set M_i can be used. Let $(r_1,\dots,r_s)$ be s distinct integers from $\overline{1,m}$ (so $s \leqslant m$). We call

$$r(s) \triangleq (0, r_1,\ r_2, \dots, r_s,\ m+1)$$

a *route* (of size s) from x_0 to x_f passing respectively through the sets $M_{r_1}, \dots, M_{r_s}$. A route $r(s)$ is *possible*, if all arcs

$$(0,r_1),\ (r_1,r_2),\ \dots,\ (r_{s-1},r_s),\ (r_s,m+1)$$

are admissible (i.e., belong to A). Note that there are no loops in $r(s)$. In every set M_{r_i} choose some visit node x_i, and determine a *path along a possible route*

$r(s)$ as a sequence

$$(x_0, x_1, \ldots, x_s, x_f)$$

(for an example, see Fig. 1). The total transition cost along this path is

$$I(r(s), (x_i)_{i \in \overline{1,s}}) = \sum_{i=1}^{s} c_{r_i}(x_{i-1}, x_i) + c_{m+1}(x_s, x_f) \tag{1}$$

or

$$I(r(s), (x_i)_{i \in \overline{1,s}}) = \max \Big\{ \max_{i \in \overline{1,s}} c_{r_i}(x_{i-1}, x_i);\ c_{m+1}(x_s, x_f) \Big\}. \tag{2}$$

Denote by $\{r\}$ the set of all possible routes $r(s)$ of all possible sizes s through the system of sets $M_1, \ldots, M_m$. The set $\{r\}$ is determined by the graph of the problem. Consider the problem

$$I(r(s), (x_i)_{i \in \overline{1,s}}) \to \min_{r(s) \in \{r\}} \inf_{(x_i)_{i \in \overline{1,s}} \in \prod_{i=1}^{s} M_{r_i}}, \tag{3}$$

where $\prod_{i=1}^{s} M_{r_i} = M_{r_1} \times \cdots \times M_{r_s}$ is the Cartesian product. The value of this problem (the "shortest path length") is denoted by V. Note that since x_0 and x_f are fixed, it is convenient to represent the minimization in paths $(x_0, x_1, \ldots, x_s, x_f)$ by an equivalent minimization in s-tuples $(x_i)_{i \in \overline{1,s}}$.

So, (3) is the problem of joint minimization in all possible routes $r(s)$ of all possible sizes $s \leqslant m$, and in all possible s-tuples $(x_i)_{i \in \overline{1,s}}$ corresponding to these routes. Problem (3) is a mixture of both discrete and continuous optimization problems, and its decomposition into a discrete (combinatorial) subproblem (to choose $r(s)$) and a continuous one (to choose $x_1, \ldots, x_s$) is impossible because the solutions to these subproblems depend on one another ($r(s)$ and $(x_i)_{i \in \overline{1,s}}$ are the bound variables: $(x_i)_{i \in \overline{1,s}} = (x_i)_{i \in \overline{1,s}}(r(s))$).

In problem (3) we seek an optimal open-loop control $\{r^0(s^0), (x_i^0)_{i \in \overline{1,s^0}}\}$, and it can be calculated in advance, before a process starts. In [56] the generalized Bellman–Ford–Moore (BFM) algorithm was proposed to solve this problem; this algorithm is the analog of the BFM Label Correcting algorithm [3,17,37] for the ordinary SPP.

The GSPP has applications to many areas including routing, location-routing, information theory, communication, transportation, network design, VLSI design, and robot motion planning.

In the Label Correcting methods (which originate from the BFM algorithm) the labels of nodes are updated at each iteration until the last needed one. In the BFM algorithm the number of intermediate nodes in a path is iterated.

Now consider briefly solving the GSPP. All functions c_i, $i \in \overline{1, m+1}$, are assumed to be bounded below. Forget for a while the point x_0, and consider the following set of auxiliary generalized SP subproblems. For every fixed index set $H \subset \overline{1,m}$ such that $1 \leqslant |H| \leqslant m-1$, and fixed $x \in M_i$, $i \in \overline{1,m} \setminus H$,

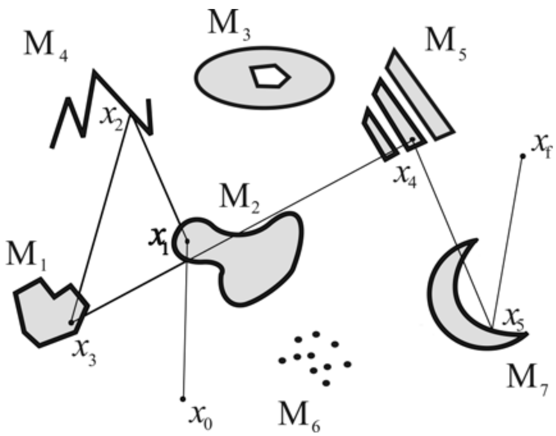


Figure 1: Example of a path ($m = 7$, $s = 5$, $r = (0, 2, 4, 1, 5, 7, 8)$).

consider the problem to find the shortest path from x to x_f through the system of intermediate sets $\{M_j,\, j \in H\}$. The value of this SPP is defined analogously to (3) (where $s \leqslant |H|$ and $\forall i \in \overline{1,s} : r_i \in H$) with "initial conditions" (H, x) instead of $(\overline{1,m}, x_0)$. Denote this value by $V(H; x)$. So, in every such auxiliary problem we ignore the node x_0, and pose the SPP only for possible paths within the system $\{M_j,\, j \in H\}$ that start at x and end at x_f.

Let

$$W_H \triangleq \bigcup_{j \in \overline{1,m} \setminus H} M_j \qquad (H:\ 0 \leqslant |H| \leqslant m-1),$$

$$W_{\overline{1,m}} \triangleq x_0$$

($|H| = 0$ corresponds to $H = \varnothing$). In the case of cost (1), define the equation on the space of admissible positions (H, x) ($|H| \geqslant 1$, $x \in W_H$):

$$J(H; x) = \min\left\{c_{m+1}(x, x_f);\ \min_{i \in H} \inf_{y \in M_i} [c_i(x, y) + J(H \setminus \{i\}; y)]\right\} \tag{4}$$

with the boundary condition

$$J(\varnothing; x) = c_{m+1}(x, x_f) \qquad (x \in M_i,\ i \in \overline{1,m}). \tag{5}$$

So, $J(\varnothing; x)$ is the value of the direct path from x to x_f. When the graph of the problem is not a complete one, then the minimization in i in (4) is taken over the set $\tilde{FS}(i^*) \cap H$, where i^* is such that $x \in M_{i^*}$, and

$$\tilde{FS}(i^*) \triangleq \{j : (i^*, j) \in FS(i^*)\}.$$

The solution to Equation (4) with condition (5) is obtained recursively [56] with the recursion in layers of sets H of constant cardinality $|H| = k$, where k

varies from 1 to $m-1$ (see Section 4, where an analogous procedure is described). For any $H : |H| = k$, $k \in \overline{1, m-1}$, and for any $x \in W_H$ the value $J(H; x)$ is the value $V(H; x)$ of the shortest path from x to x_f via at most k intermediate sets (namely, $\{M_j,\ j \in H\}$). So, in this algorithm the number of intermediate sets is iterated, and that is why it was named the generalized BFM in [56].

Practically all existing one-to-one SP algorithms are based on the implementation of some single origin (one-to-all) or single destination (all-to-one) SP algorithm. We have the same situation here. Namely, at the last stage of the above recursive procedure we have determined the optimal values $J(H; x)$ for $H \subset \overline{1, m}$ with $|H| = m-1$, and for $x \in M_j$, $j \in \overline{1, m} \setminus H$. The set of these (auxiliary) SP problems can be considered as one generalized all-to-one SP problem (in the sense that we find the shortest paths from every node of every set M_i to destination x_f through the system of intermediate sets $\{M_j,\ j \in \overline{1, m} \setminus \{i\}\}$).

Then, finally, calculate (again by (4)) the "full" value $J(\overline{1, m}; x_0)$, and it equals the value of the initial problem (3): $J(\overline{1, m}; x_0) = V$. So, we have calculated (and stored) via the backward dynamic programming (DP) procedure the values $J(H; x)$ for all possible positions (H, x). Now we can calculate the desired shortest path from x_0 to x_f. Namely, moving forward, step by step, from x_0 to x_f, and choosing at each kth step the number r_k^0 and node $x_k^0 \in M_{r_k^0}$ which give the minimums in (4), obtain the optimal route $(0, r_1^0, \ldots, r_{s^0}^0, m+1)$ and the shortest path $(x_0, x_1^0, \ldots, x_{s^0}^0, x_f)$, where $x_k^0 \in M_{r_k^0}$, $k \in \overline{1, s^0}$. If at some k^*th step the first of two compared values in (4) is lesser, then we go to x_f, and the procedure terminates (and so $s^0 = k^* - 1$); otherwise we continue the procedure, going to the next intermediate set M_i.

Remark 2.1. The solutions to problem (3) cannot produce a path with cycles (since repetitions are excluded a priori), and so the standard conditions of lack of negative or zero cycles are of no importance for problem (3). The algorithm is valid for any situation.

The time complexity of the algorithm is $O(N^2 m^2 2^m)$, and the space complexity is $O(Nm2^m)$, where $N = |\tilde{M}_i|$, $i \in \overline{1, m}$, and $\tilde{M}_i$ is a grid on M_i (see Section 5).

Note that the use of argument (H, x) instead of x for optimal function J is a crucial point in the problem without returns. Indeed, let us define the optimal function in the form $J_i(x)$, $x \in M_i$, $i \in \overline{1, m}$, i.e., without argument H, and let $x, \ldots, y, \ldots, x_f$ be the above shortest path from $x \in M_i$ to x_f, where $y \in M_k$ is some intermediate node. Then, considering (independently of $x \in M_i$ and $J_i(x)$) the analogous SPP from y to x_f, we can face the situation when the shortest path with value $J_k(y)$ from y to x_f can pass through some node $z \in M_i$. So, the shortest path from $x \in M_i$ to x_f can return to the set M_i. Moreover, the shortest path in such a problem can visit some sets M_i many times! The corresponding examples are simple to construct. Thus, in order to formalize properly our SPP without returns, we must consider the current argument of the optimal function J in the recursive equation as the pair (H, x), where H is an index set of remaining possible sets to visit (or, equivalently, an index set of all sets already passed). This leads to larger time and space complexities, but it is a price to pay for our hard combinatorial restriction "each M_i can be visited only once." The set H is none other than the history (memory) of a process. So, to produce every next step of a route, the solving procedure needs a preceding history H. And it is well known that the prohibition of repetitive moves in combinatorial problems can make the solution much harder ([50], [52]–[54], [60], see also [11,16,26,34,57,63] on relative questions).

Consider, for comparison of complexities, the "degenerate" GSPP, when both returns and intra-set arcs are allowed. In this case the GSPP becomes the ordinary one-to-one SPP with origin x_0, destination x_f, and a large amount $\bigcup_{i=1}^{m} \tilde{M_i}$ of intermediate points (namely, mN points). Further, let the standard condition of lack of negative cycles be fulfilled. Hence, solving by the ordinary BFM algorithm has the (worst-case) time complexity $O(m^3N^3)$ (for sparse graphs $O(nmN^3)$), and memory space $4mN + 2nN^2$ (since there are now nN^2 arcs instead of n). So, one can see that this problem is far less expensive in comparison with GSPP (and far less interesting and important for applications).

The advantage of generalized combinatorial optimization problems (COPs), such as the GSPP, is that one obtains an additional freedom to vary the possible locations of facilities within some preassigned sets, and so to improve the characteristics of a system under design. Further, the generalized (set membership) settings of COPs give a very convenient way to model and handle the game of COPs with uncertainties, when nature or a real opponent plays against us, having at its (his) disposal the errors, disturbances, noises, failures, or reasonable controls. Namely, instead of a single fixed ith vertex in a graph of a problem one uses a set M_i by which the possible realizations of an unknown ith disturbance or control are modeled. An example is the dynamic network problems under set membership uncertainties in node locations or in edge lengths, when the network administrator has the ability to make on-line corrections, and so needs some mathematical algorithm to optimize a process.

3 The Feedback Problem

3.1 Problem Statement

Now consider the feedback setting of the game GSPP. As stated in the Introduction, we divide the control $\{r(s), (x_i)_{i\in\overline{1,s}}\}$ in two parts. Recall that only one node in every M_i can be chosen.

Consider the following antagonistic multistage positional game of two players. At the initial "time" $k = 0$ a path is at the state x_0. Player 1 chooses the number $\{m + 1\}$, and so ends the game, or r_1, a number of the first intermediate set to go. After that, player 2 chooses a point x_1 within a set M_{r_1}, and so a path comes to a point x_1. Player 1 receives an information about x_1 ("measures" it), and then, knowing r_1 and x_1, chooses $\{m+1\}$, and so ends the game, or a number r_2 of the second set to go. After that, player 2 chooses a point $x_2 \in M_{r_2}$, and so on. At the $(k+1)$th step player 1 on the basis of information on $r_1, \ldots, r_k$ and $x_k \in M_{r_k}$ chooses $\{m+1\}$, and so ends the game, or a number r_{k+1} of the next set to go; then player 2 chooses a point $x_{k+1} \in M_{r_{k+1}}$, so a path arrives at a point x_{k+1}. At the mth step player 1 chooses $\{m+1\}$, and so ends the game, or the last remaining number r_m to go; then player 2 chooses a point $x_m \in M_{r_m}$. At the $(m+1)$th step player 1 (automatically) chooses the single remaining number $\{m+1\}$, i.e., the point x_f, and the game is over. As a result, obtain a route $(0, r_1, \ldots, r_s, m+1)$ and a path $(x_0, x_1, \ldots, x_s, x_f)$, where s, $s \leqslant m$, is a moment at which player 1 has decided to end the game. Under such a game formalization, player 1 at each step, depending on what uncertainty x_k has been realized, can generate various continuations of the carved-out part of a route until then. By receiving the above information on points x_k he may use the mistakes of the opponent to improve himself.

The two-person game has informational lag for the first player (in the sense that at every step he "plays first" and does not know what will be the opponent's move at this step, while player 2, who "plays second," knows the move of player 1 at this step). So, the players move alternately, with player 1 playing first.

It is assumed that the first player's control at the $(k+1)$th step is designed on the basis of available information on the current state of the system. By available information we mean the following. Locating at the set M_{r_k}, player 1 selects $\{m+1\}$ or the next visit number r_{k+1} after he gets the value $x_k \in M_{r_k}$; besides, he stores the numbers $r_1, \ldots, r_k$ of all sets M_i he has passed (which is equivalent to him knowing the numbers $r_{k+1}, \ldots, r_m$ of all remaining sets M_j). By *position* we mean a pair (H, x), where H is a set of indices j of all sets M_j which remain to be visited, and x is a point at which a path locates. By *feedback control* at the $(k+1)$th step we mean a function $R_{k+1}(H; x)$ defined for all H and x which may arise at this step, and producing r_{k+1} or $\{m+1\}$. All H and x which may appear during a motion satisfy the conditions $H \in 2^{\overline{1,m}} \cup \{\varnothing\}$ and $x \in x_0 \cup \left\{\bigcup_{j=1}^{m} M_j\right\}$.

Definition 3.1. A *feedback (positional) strategy* R of player 1 is any function $R : (H, x) \to R(H; x)$ defined for all possible positions (H, x), and producing for every current position a number of the next set to reach, or the number $\{m+1\}$ to end the game.

Definition 3.2. A *motion* $x(\cdot)$ *generated by strategy* R from initial position $(\overline{1,m}, x_0)$ is any pair

$$
\begin{gathered}
x(\cdot) = \{(0, r_1, \ldots, r_s, m+1),\ (x_0, x_1, \ldots, x_s, x_f)\}, \\
x_k \in M_{r_k}, \quad r_k = R(\overline{1,m} \setminus \{r_i : i \in \overline{1, k-1}\}; x_{k-1}), \quad k \in \overline{1,s}, \quad s \leqslant m, \\
\{m+1\} = R(\overline{1,m} \setminus \{r_i : i \in \overline{1,s}\}; x_s).
\end{gathered} \tag{6}
$$

Here Definition 3.2 is correct, since (see Definition 3.1) any strategy R sooner or later brings the corresponding motion to the destination x_f; this may occur, e.g., at the first step of the game ($s = 0$), or at the last (possible) step ($s = m$), or at some intermediate step ($0 < s < m$).

It is also convenient to represent R as a collection of functions $R = \{R_1, \ldots, R_m\}$, where $R_k(H; x)$ is determined for all H and x which may arise at the kth step ($k \in \overline{1,m}$). Here $k = m - |H| + 1$. Note that for $k = 1$ in (6) we have $\{r_i : i \in \overline{1,0}\} = \{r_i : i \in \varnothing\} = \varnothing$, so $r_1 = R(\overline{1,m}; x_0)$.

For these feedback motions use the notation $I(x(\cdot)) \triangleq I(r(s), (x_i)_{i\in\overline{1,s}})$. For every strategy R denote by $X(R)$ the bundle of all possible motions $x(\cdot)$ generated by R. So, to every motion $x(\cdot)$ from $X(R)$ corresponds some fixed route $r(s)$ and ordered s-tuple of uncertainties. The value

$$\sup_{x(\cdot)\in X(R)} I(x(\cdot))$$

is the result which the first player can guarantee himself under strategy R. His aim is to choose a strategy minimizing this guaranteed amount.

Problem 3.1. Find a strategy $\tilde{R}$ satisfying

$$\sup_{x(\cdot)\in X(\tilde{R})} I(x(\cdot)) = \min_{R} \sup_{x(\cdot)\in X(R)} I(x(\cdot)) = \gamma^0. \tag{7}$$

The minimum in (7) is really attained (see Theorems 4.1–4.3). By ε^0 denote the value of the problem in the class of program controls $r(s)$:

$$\varepsilon^0 \triangleq \min_{r(s)\in\{r\}} \sup_{(x_i)_{\overline{1,s}}\in\prod_{i=1}^{s} M_{r_i}} I(r(s), (x_i)_{i\in\overline{1,s}});$$

in this setting the first player is not allowed to change the route during a motion.

Note that, of course, one can transit from node x_0 to node x_f immediately (if $c_{m+1}(x_0, x_f) < \infty$), but it may turn out that the value $\varepsilon^0 = \varepsilon^0(\overline{1,m}; x_0)$ and, all the more, the value $\gamma^0 = \gamma^0(\overline{1,m}; x_0)$, are more profitable than $c_{m+1}(x_0, x_f)$ (e.g., $\gamma^0 \ll c_{m+1}(x_0, x_f)$), and so it is natural to use the feedback setting.

Some combinatorial problems, such as the TSP, GTSP, SPP, and GSPP, carry a certain duality as concerns their dynamic character. Formally, they are dynamic problems, since one moves from one node to another. But in fact they are static a priori problems for choosing a permutation or a route and a tuple of nodes (for the GTSP and GSPP). Analogously, in the game case, one may consider the program settings of the GGSPP and GGTSP as static problems as well. In contrast, the feedback settings of the GGSPP and GGTSP [55] are really dynamic routing problems, with nontrivial decision-making during a route.

Let us mention some possible applications of Problem 3.1.

3.2 Examples in Communication Systems

Suppose we have to transmit a signal (message, packet, command, *etc.*) from a device x_0 to a remote device x_f via the system of intermediate clusters M_i, $i \in \overline{1,m}$, of receiving-transmitting devices (RTDs). The clusters M_i can be compact zones (office buildings with many independent RTDs) or geographically dispersed industrial areas of RTDs (say, clusters of relay stations). The transmission can be wireless or cable (in this case, e.g., all RTDs y of a local area network (LAN) M_i can terminate at the ports of the ith multiplexer, which automatically parallelizes an external signal to all y in M_i, and transmits back a signal from any $y \in M_i$ to the global backbone cable network). We can exploit in each cluster M_i only one station y. Each station $y \in M_i$ has its own tariffs $c_i(x, y)$ for servicing the signal received from preceding stations x. We do not know in advance which station y in M_i will service our signal (it can be any of them), but all the above tariffs are known. The feedback transmission goes as follows. We transmit the signal from supervising node x_0 to some cluster M_{r_1}, and it is received by some station $x_1 \in M_{r_1}$. Then RTD x_1 gets in touch with us, and we give it the next number r_2 of the area to which our signal must be sent. At area M_{r_2} the signal is received by some RTD $x_2 \in M_{r_2}$, which gets in touch with us, and so on. Our aim is to minimize the total costs of transmission, and it makes sense to use the feedback setting (of course, if the a priori value γ^0 is satisfactory at all).

The variety of costs $c_i(x, y)$ can depend on wire length, physical types of cables, speed of processing of information by a device, different tariffs, and so on up to the human factors.

One particular case of this problem is that of finding the *most reliable transmission in a communication network* [6,18,42]. Consider the generalized version of this problem of finding the most reliable path between two given stations in a communication system. Let M_i, $i \in \overline{1,m}$, be the clusters of relay stations. The problem is to transmit a signal or a message from x_0 to a remote point x_f via the system of intermediate areas M_i, $i \in \overline{1,m}$, in the most reliable way. Only one RTD can be used in a cluster M_i, $i \in \overline{1,m}$. Let $p_i(x, y)$ be the reliability of a link (channel) (x, y) from node x to node $y \in M_i$, i.e., $p_i(x, y)$ is the probability of a link (x, y) being operative (i.e., not destroyed or jammed). All probabilities are independent, that is, links fail independently of one another.

If $r(s)$, where $s \leqslant m$, is some route, and $(x_0, x_1, \ldots, x_s, x_f)$, $x_i \in M_{r_i}$, $i \in \overline{1,s}$, is some path along a route $r(s)$, then its reliability is given by the product of reliabilities of its links:

$$I(r(s), (x_i)_{i \in \overline{1,s}}) = \prod_{i=1}^{s} p_{r_i}(x_{i-1}, x_i) \cdot p_{m+1}(x_s, x_f). \tag{8}$$

The problem is to find a route and a path which maximize (8). Now reduce the maximum reliability problem to the GSPP. Consider the same system of sets

M_i, $i \in \overline{1,m}$, and let the cost of a link (x,y) be now $c_i(x,y) = -\ln p_i(x,y)$ instead of $p_i(x,y)$. Then the maximization of (8) is equivalent to minimization of the "path length"

$$\sum_{i=1}^{s} c_{r_i}(x_{i-1}, x_i) + c_{m+1}(x_s, x_f).$$

We do not know in advance which relay stations in the sets M_i will be used, and hence we pose the feedback problem in order to obtain a certain guaranteed reliability γ^0.

Another communication problem is the *optimal wireless transmission via the system of mobile retranslators*. There is one mobile retranslator in every local geographical area M_i, $i \in \overline{1,m}$. The costs $c_i(x,y)$ (e.g., power of a signal) depend on geographical coordinates x and y because of landscape and atmospheric factors. The signal is transmitted consequently from one retranslator to another. The current locations of retranslators within M_i, $i \in \overline{1,m}$, are not known, so it is supposed that at any time moment the ith retranslator can occur at any point of M_i. Thus, to find a (guaranteed) shortest path, one can apply the above feedback posing.

4 Optimal Minimax Strategy

In this section one designs the strategy which solves Problem 3.1 when the cost is given as in (1). Assume that all functions c_i, $i \in \overline{1,m}$, are bounded above. On the set of all possible positions (H, x) define the Bellman equation

$$J(H;x) = \min\Big\{c_{m+1}(x,x_f);\ \min_{i\in H}\sup_{y\in M_i}\left[c_i(x,y) + J(H\setminus\{i\};y)\right]\Big\},$$

or, equivalently, and more conveniently,

$$J(H;x) = \min_{i\in H\cup\{m+1\}}\ \sup_{y\in M_i}\Big\{c_i(x,y) + J(H\setminus\{i\};y)\Big\} \tag{9}$$

(where $\forall\ H : J(H\setminus\{m+1\};x_f) \triangleq 0$) with the boundary condition (5).

If, e.g., the functions $c_i(x,y)$, $i \in \overline{1,m}$, are jointly upper semicontinuous (u.s.c.) on $\mathbb{R}^d \times \mathbb{R}^d$, the function $c_{m+1}(x,x_f)$ is u.s.c. on $\mathbb{R}^d$, and the sets M_i, $i \in \overline{1,m}$, are compact, then the maximum in (9) is always attained, since $\forall\ H \subseteq \overline{1,m}$ the function $J(H;x)$ (9) is u.s.c. on $\mathbb{R}^d$. This fact follows from the well-known (see, e.g., [1]) u.s.c. property of functions $\varphi(x) = \sup_{y\in Y}\varphi(x,y)$ and $\psi(x) = \min_{i\in H}\psi_i(x)$, where φ is jointly u.s.c., ψ_i are u.s.c., and Y is compact.

The solution to equation (9) with condition (5) is obtained as follows. The recursion (9) is in layers of sets H with constant cardinality $|H| = k$, where

k varies from 1 to m. In view of the specificities of the problem, one should calculate the values $J(H;x)$ only for the pairs (H,x) such that

$$H \in 2^{\overline{1,m}} \cup \{\varnothing\}, \quad x \in W_H.$$

We begin at the boundary condition (5), and at the first stage calculate $J(H;x)$ for all single-element sets $H = \{k\}$ and points $x \in (\bigcup_{j=1}^m M_j) \setminus M_k$ $(k \in \overline{1,m})$. At the second stage calculate $J(H;x)$ for all two-element sets $H = \{k,q\}$ and points $x \in (\bigcup_{j=1}^m M_j) \setminus (M_k \bigcup M_q)$ $(k,q \in \overline{1,m},\ k \neq q)$, and so on. At the last but one stage calculate $J(H;x)$ for all $(m-1)$-element sets $H = \overline{1,m} \setminus \{k\}$ and $x \in M_k$ $(k \in \overline{1,m})$. At the last stage calculate the value $J(\overline{1,m};x_0)$. Note that in comparison with Section 3 the reasoning "in steps" k is replaced by reasoning "in layers" of sets H of cardinality $m-k+1$ (so $k = m-|H|+1$).

In parallel with the calculation of $J(H;x)$, we find and store the number

$$R_k^0(H;x) \triangleq i^0(H;x), \tag{10}$$

which provides the minimum in (9), for all (H,x) which may occur at this step. If such an $i^0(H;x)$ is nonunique, choose and store any one of them.

Theorem 4.1. *The strategy R^0 defined by (9), (5), (10) satisfies the condition*

$$\sup_{x(\cdot) \in X(R^0)} I(x(\cdot)) = J(\overline{1,m};x_0).$$

Then, evidently, $\gamma^0 \leq J(\overline{1,m};x_0)$. In fact, R^0 is an optimal minimax strategy.

Theorem 4.2. *For every $H \subseteq \overline{1,m}$ and $y \in W_H$ the equality $\gamma^0(H;y) = J(H;y)$ is true. In particular, one has $\gamma^0 = J(\overline{1,m};x_0)$.*

Proof. Analogously to (7), define the value function $\gamma^0(H;y)$ for all $H \subseteq \overline{1,m}$, $y \in W_H$:

$$\gamma^0(H;y) \triangleq \min_R \sup_{x(\cdot) \in X(H;y;R)} I_H(x(\cdot)),$$

where

$$I_H(x(\cdot)) = \sum_{i=1}^{s} c_{j_i}(x_{i-1},x_i) + c_{m+1}(x_s,x_f),$$

$x_i \in M_{j_i}$, $i \in \overline{1,s}$; x_0 is y, $(j_1,\ldots,j_s)$ are any s distinct integers from H (so $s \leqslant |H|$); $X(H;y;R)$ is the bundle of all possible motions generated by strategy R from initial position (H,y). Surely, for $(H,y) = (\overline{1,m},x_0)$ one has $\gamma^0(\overline{1,m};x_0) = \gamma^0$ and $X(\overline{1,m};x_0;R) = X(R)$ (see Section 3). So, $\gamma^0(H;y)$ is the minimal guaranteed result in the subgame starting from initial position (H,y).

Analogously to Theorem 4.1, one obtains the inequality

$$\gamma^0(H;y) \leqslant J(H;y) \qquad (H \subseteq \overline{1,m},\ y \in W_H). \tag{11}$$

Let us prove that in fact

$$\gamma^0(H;y) = J(H;y). \tag{12}$$

In view of (11) it is sufficient to prove that

$$\gamma^0(H;y) \geqslant J(H;y). \tag{13}$$

For the proof of (13) we proceed by induction in $|H|$. For $H = \{i\}$ with $|H| = 1$ we have

$$\gamma^0(\{i\};y) = \min\Big\{c_{m+1}(y,x_f);\ \sup_{x\in M_i}[c_i(y,x)+c_{m+1}(x,x_f)]\Big\} = J(\{i\};y),$$

where $i \in \overline{1,m}$, so (13) holds. Now let (13) be true for all H such that $|H| \leqslant k_* < m$, $k_* \geqslant 1$, and prove (13) for all H with cardinality $|H| = k_*+1$. Fix any position (H,y) with $|H| = k_*+1$, and fix any strategy R. According to Definition 3.1, we can check the exact number $j_1 \in H \cup \{m+1\}$ of the set M_{j_1} which will be chosen *first* to visit by strategy R starting from initial position (H,y). Namely, $j_1 = R(H;y)$. Consider the guaranteed result under the application of strategy R (see Section 3). We have (x_0 is y):

$$\begin{aligned}
\varphi(H;y;R) &\triangleq \sup_{x(\cdot)\in X(H;y;R)}\Big\{\sum_{i=1}^{s} c_{j_i}(x_{i-1},x_i) + c_{m+1}(x_s,x_f)\Big\}\\
&= \sup_{x(\cdot)\in X(H;y;R)}\Big\{c_{j_1}(y,x_1) + \sum_{i=2}^{s} c_{j_i}(x_{i-1},x_i) + c_{m+1}(x_s,x_f)\Big\}\\
&= \sup_{x_1\in M_{j_1}}\ \sup_{x(\cdot)\in X(H\setminus\{j_1\};x_1;R)}\Big\{c_{j_1}(y,x_1)\\
&\qquad\qquad + \sum_{i=2}^{s} c_{j_i}(x_{i-1},x_i) + c_{m+1}(x_s,x_f)\Big\}\\
&= \sup_{x_1\in M_{j_1}}\Big\{c_{j_1}(y,x_1) + \varphi(H\setminus\{j_1\};x_1;R)\Big\}\\
&\geqslant \sup_{x_1\in M_{j_1}}\Big\{c_{j_1}(y,x_1) + \gamma^0(H\setminus\{j_1\};x_1)\Big\}\\
&\geqslant \sup_{x_1\in M_{j_1}}\Big\{c_{j_1}(y,x_1) + J(H\setminus\{j_1\};x_1)\Big\} \geqslant J(H;y).
\end{aligned}$$

The last inequality in this chain follows from (9), and the last but one inequality follows from the hypothesis of the inductive step. So, we have obtained the inequality for the guaranteed result $\varphi(H;y;R)$ under application of a fixed strategy R. Since R was chosen arbitrarily, then (13) is true. The inequalities (11), (13) imply the equality (12). $\square$

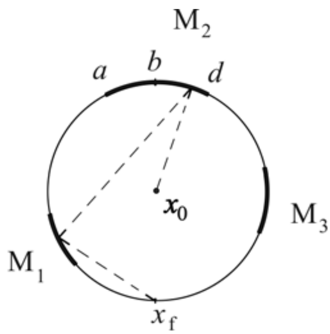


Figure 2: Feedback is better.

From Theorems 4.1 and 4.2 follows

Theorem 4.3. *The strategy R^0 is the solution to Problem 3.1 for the case of cost (1).*

Here is an example showing that, in general, feedback control is better than open-loop control. The sets M_1, M_2, M_3 are the arcs lying on a circumference $\mathcal{A}$ of radius p and with center at x_0 (see Figure 2).

Let the midpoint of M_2 be b, and let the length of M_2 be h. The terminal point x_f lies on $\mathcal{A}$ opposite to b. The dotted line indicates some possible path. Let all the costs $c_1(x_0, x)$, $x \in M_1$, and $c_3(x_0, x)$, $x \in M_3$, and also the cost $c_f(x_0, x_f)$, be unacceptably high. Put $c_2(x_0, x) = p$ for all $x \in M_2$. Let the costs $c_j(x, y)$ of transition from $x \in M_i$ to $y \in M_j$, and also the costs $c_4(x, x_f)$ of transition from $x \in M_i$, $i \in \overline{1,3}$, to $x_f = x_4$ be equal to the length of a lesser arc of $\mathcal{A}$ between x and y or between x and x_f respectively. Under these conditions, the only reasonable routes from x_0 to x_f are $r_*(2) = (0, 2, 1, 4)$ and $r_{**}(2) = (0, 2, 3, 4)$. For any of them the second player may choose at the first step of the game the farthest point in M_2 (d and a, respectively), and hence the first player in the open-loop game can guarantee himself only the value

$$\varepsilon^0 = p(1 + \pi) + h/2.$$

Now consider the feedback setting. If the second player chooses at the first step some point $x_1 \in M_{r_1} = M_2$ such that $x_1 \neq b$, then the first player goes along a lesser arc to x_f and takes the result $\gamma < p(1 + \pi)$. If the second player chooses $x_1 = b$, then the first player takes the result $\gamma = p(1 + \pi)$. Hence,

$$\gamma^0 = p(1 + \pi) < \varepsilon^0. \qquad \square$$

By construction, R^0 is a *universal* strategy (that is, it remains an optimal minimax strategy for any intermediate position (H, x) considered as an initial one), see [32,33,59,8,9,41]. If player 2 has selected an erroneous (i.e., nonworst) uncertainty, the strategy R^0 self-dependently guarantees (in an optimal way) the result strictly better than γ^0. So, there is no need to readjust R^0 into another strategy in order to seize the opportunity of unsuccessful action of the second player, and to improve the result. The universal strategy does not "forgive" any mistake of the opponent, and automatically takes the mistakes (if they occur) into account at each kth step. So, using the universal optimal strategy, one can obtain at the end of the game a much better result than the initial optimal guaranteed feedback result γ^0, especially in the case when the second player makes chaotic moves.

To realize the strategy R^0, one may store, along with values $J(H;x)$, the values $R^0(H;x)$ on the whole array of possible positions (H,x). But then a double memory space is needed. To avoid it, we suggest calculating the values $R^0(H;x)$ currently, starting from the initial position. Namely, we do not tabulate the values $R^0(H;x)$ during the "backward" DP procedure. Instead, we calculate them only for those current positions (H,x_k) which appear during the "forward" real game process, according to current realizations of x_k. Thus, we are to produce (in the worst case) $m-1$ additional local min sup operations to find R^0_k, $k \in \overline{2,m}$ (the value $R^0(\overline{1,m};x_0)$ we remember, and the last number is (in the worst case) $R^0_{m+1}(\varnothing,x_m) = \{m+1\}$). The additional amount of time to produce these operations is very small with respect to the running time of the whole algorithm (see Section 5).

5 Computational Aspects

In this section we calculate the time and space complexities of our algorithm for a sequential (one-processor) machine. For numerical computations replace the sets M_i by finite grids $\tilde{M}_i$, $i \in \overline{1,m}$. So, one should calculate the values $J(H;x)$ and $R^0(H;x)$ for nodes $x \in \bigcup_{i=1}^m \tilde{M}_i$ and sets $\tilde{M}_i$ instead of M_i, $i \in \overline{1,m}$. Denote the cardinality of $\tilde{M}_i$ by $N_i : N_i \triangleq |\tilde{M}_i|$, $i \in \overline{1,m}$.

The time complexity is evaluated in terms of a number of elementary steps which are the basic operations (addition, multiplication, comparison, transfer of a word from RAM to disk buffer and back, and executing subroutine calls to calculate $c_i(x,y)$, $i \in \overline{1,m}$, and $c_{m+1}(x,x_f)$). By p_i denote the number of basic operations required to calculate the value $c_i(x,y)$ for fixed x,y by ith subroutine. Without loss of generality, assume that N_i and p_i are constants: $N_i = N$, $p_i = p$ for any $i \in \overline{1,m}$; otherwise, putting $N \triangleq \max_i N_i$ and $p \triangleq \max_i p_i$, obtain the upper bounds for both complexities.

Theorem 5.1. *The time complexity of the optimal algorithm is*

$$O(N^2 m^2 2^m),$$

and its space complexity is

$$O(Nm2^m).$$

Proof. At first find the space complexity. For every $k \in \overline{1,m}$ the number of distinct sets H of size $|H| = k$ is C_m^k. For every $k \in \overline{1,m-1}$ the number of corresponding values $J(H;x)$ (with $|H| = k$) that have to be computed is $N(m-k)C_m^k$. Now, summing in k and using formulae

$$\sum_{k=1}^{m} kC_m^k = m2^{m-1}, \qquad \sum_{k=0}^{m} C_m^k = 2^m, \tag{14}$$

we obtain

$$\sum_{k=1}^{m-1} N(m-k)C_m^k = Nm2^{m-1} - Nm.$$

So, the (special) memory for storing the values $J(H;x)$ is

$$Q = Nm2^{m-1} - Nm + 1$$

(one cell added for $J(\overline{1,m};x_0)$). Thus, the space complexity is $O(Nm2^m)$.

Now find the time complexity. The number of basic operations to calculate and store a single value $J(H;x)$ (see (9)) with $|H| = k$ ($k \in \overline{1,m}$) is $T_k = (p+4)Nk + p + 2$. The time to calculate and store all values $J(H;x)$ with $|H| = 1, \dots, m$ is

$$T = T_m + \sum_{k=1}^{m-1} N(m-k)C_m^k T_k.$$

Using formulae (14) and

$$\sum_{k=1}^{m} k^2 C_m^k = m(m+1)2^{m-2},$$

we obtain

$$T = (p+4)N^2m(m-1)2^{m-2} + (p+2)Nm2^{m-1} - 2Nm + p + 2.$$

Further, as was said above, to avoid a double memory array, we calculate the values $R^0(H;x_k)$ during a process, along a real motion. The additional amount of time for this is no more than

$$T_{add} = \sum_{k=2}^{m-1} T_k = \frac{(p+4)}{2}N(m+1)(m-2) + (p+2)(m-2).$$

Finally, the total running time of the whole algorithm is $T_{tot} = T + T_{add}$. Hence, the time complexity is $O(N^2m^22^m)$. □

6 Heuristic Strategies

As follows from Theorem 4.2, the optimal feedback guaranteed result γ^0 is, in fact, a multiple minimax:

$$\gamma^0 = \min_{r_1} \sup_{x_1} \ldots \min_{r_k} \sup_{x_k} \ldots \min_{r_m} \sup_{x_m} \; I(r, (x_k)_{k \in \overline{1,m}}) \tag{15}$$

(where $r_k \in \overline{1, m+1} \setminus \{r_1, \ldots, r_{k-1}\}$, $x_k \in M_{r_k}$, $k \in \overline{1,m}$), and so at the kth step the local operation $\min_{r_k} \sup_{x_k}$ is calculated on the values of a specially designed optimal function, see (9). For any intermediate position (H, x) the optimal value $J(H; x)$ is a multiple minimax as well.

Since the time and space complexities of the exact (optimal) algorithm are large, a more pragmatic approach is to have approximate or heuristic solutions. We present here some heuristic strategies. In view of (15), it is reasonable, under designing a well-grounded feedback heuristic, to follow (as far as possible) this multiple minimax principle. Recall (see Section 3) that any feedback strategy is a rule which for every possible position (H, x) points out a number of the next set to reach or the number $\{m + 1\}$ to exit and end the game.

6.1 One Local Minimax Heuristic

For every current position (H, x) the first player goes to a set with a number $i^0(H; x)$ which provides a minimum in $i \in H \cup \{m + 1\}$ in

$$J^{(1)}(H; x) = \min \Big\{ c_{m+1}(x, x_f);\; \min_{i \in H} \sup_{y \in M_i} [c_i(x, y) + c_{m+1}(y, x_f)] \Big\}.$$

The total (worst-case) time complexity of the game under application of this heuristic is $O(Nm^2)$.

6.2 Two Local Minimax Heuristic

The next number $i^0(H; x)$ to go provides an outer minimum in $i \in H \cup \{m + 1\}$ in

$$J^{(2)}(H; x) = \min \Big\{ c_{m+1}(x, x_f);\; \min_{i \in H} \sup_{y \in M_i} \Big\{ c_i(x, y) + \min \;\Big\{ c_{m+1}(y, x_f);$$

$$\min_{j \in H \setminus \{i\}} \sup_{z \in M_j} \; [c_j(y, z) + c_{m+1}(z, x_f)] \Big\} \Big\} \Big\}. \tag{16}$$

The total time complexity of the game under this heuristic is no more than $O(N^2 m^3)$.

6.3 The l Local Minimax Heuristic

A number $i^0(H;x)$ of the next set to reach provides an outer minimum in $i \in H \cup \{m+1\}$ for the function $J^{(l)}(H;x)$, which is written down analogously to (16). Here $l \leqslant m$ (the case $l = m$ corresponds to the exhaustive search). The total (worst-case) time complexity of the game under this heuristic is $O(N^l m^{l+1})$.

Note that if the functions $c_i(x,y)$, $i \in \overline{1,m}$, and $c_{m+1}(y,x_f)$ have good properties in y (differentiability, convexity, concavity), then the last max operation in the above relations for $J^{(1)}$, $J^{(2)}$, $J^{(l)}$ can be implemented by numerical methods (since the functions c_i and c_{m+1} are given explicitly), and hence the time complexities are reduced by N: $O(m^2)$, $O(Nm^3)$, $O(N^{l-1}m^{l+1})$.

In these heuristics we do not calculate and tabulate the arrays of values of $J^{(l)}(H;x)$ or $i^0(H;x)$, but calculate only the current values $i^0(H;x)$ along a particular real motion. Further, these heuristics give predictions for l steps ahead and cut the last $|H|-l$ steps (if $|H|>l$) in a corresponding multiple minimax. In terms of graph theory, this is an exhaustive minimax (guaranteed) search on a game tree, when a tree is cut off at the lth level in depth.

One may use the sharper heuristics, to play the last $|H|-l$ steps of the search as well. Namely, we insert some function $h(H \setminus \{i_1,\dots,i_l\};x)$ which estimates a result of the search on the last (undone) $|H|-l$ steps of the exhaustive search. For example, for $l=1$ we have

$$J^{(1)}(H;x) = \min\Big\{c_{m+1}(x,x_f);\ \min_{i\in H}\sup_{y\in M_i}[c_i(x,y)+h(H\setminus\{i\};y)]\Big\}.$$

Quite in the spirit of artificial intelligence theory (see [13,29,40,43,45,51,65]) we call function h a *heuristic estimator*. If the calculation of its values is cheap enough, we can calculate them on-line, during a real motion. In view of the above principle of multiple minimax, h should be chosen of some minimax or mean character (e.g., h can be the program minimax $\varepsilon^0(H\setminus\{i_1,\dots,i_l\};y)$ on the last $|H|-l$ steps of the game, or the generalized analog of the minimum average arc length path [62]). The function h may be given tabularly, being calculated sometime before. Surely, if h is the full multiple minimax on the last $|H|-l$ steps of a search, then this (l,h)-strategy coincides with the optimal strategy R^0.

By these heuristic search procedures player 1 at each kth step of the game determines the next move he should make (i.e., the next number to go). As one can see, any special memory (for storing the values $J(H;x)$ or $i^0(H;x)$) is not needed at all. These are very cheap heuristic algorithms, and they avoid the main difficulty of the exact algorithm: the necessity to calculate and store a large amount of values $J(H;x)$.

The l-heuristics are easy and transparent for computer coding, since they consist solely of identical procedures of finding the minimal or maximal element in a (given) linear list (in i) or in an array (in y) respectively. In fact, the problem is even easier, and we do not need any sorting procedures. Going along a list H (or a grid M_i), we simply compare, node by node, the next calculated value with the previous one, and store the minimal (or, respectively, maximal) of them. So, at the last number i of H (or at the last node y of M_i) we obtain the minimal value on H (or the maximal value on M_i). On the one hand, l-heuristics are clever (since they imitate the optimal algorithm as far as your computer allows) and reflect the serious theory; on the other hand, they are simple to understand and can be easily programmed.

The time complexity of the optimal algorithm is $O(N^2m^22^m)$, and that of the l-heuristics defined above is $O(N^lm^{l+1})$. For l varying from 1 to m the complexity $O(N^lm^{l+1})$ varies from $O(Nm^2)$ to $O(N^mm^{m+1})$. Let us find the value l^* after which the l-heuristic becomes more expensive (in the sense of time complexity) than the optimal algorithm. Evidently, such $l^* = l^*(m, N)$ is the root of equation

$$N^lm^{l+1} = N^2m^22^m,$$

hence

$$l^* = 1 + \frac{\log N + m\log 2}{\log(Nm)}.$$

Let, for simplicity, N be fixed. The ratio $\theta(m) = l^*(m)/m$ is the ratio of lengths of "time intervals" $[1, l^*]$ and $[1, m]$. One has $\theta(m) \to 0$ as $m \to \infty$ (though $l^*(m) \to \infty$ as $m \to \infty$). In Table 1 the behavior of l^* and θ is presented under $m \to \infty$ (N equals 100).

Table 1: Comparison of time complexities.

m	15	50	100	500	10^4	10^{10}	10^{100}
$l^*(m)$	3	5	9	33	500	25×10^7	3×10^{97}
$\theta(m)$	0.2	0.1	0.09	0.065	0.05	0.025	0.003

7 Generalized Bottleneck SPP

Consider problem (3), (2) [56]. So, I is the maximum link value in a path $(x_0, x_1, \ldots, x_s, x_f)$ along a route $r(s)$. This problem may be thought of as an information security problem consisting of finding a route to transmit a signal in a network so as to minimize its detectability or decipherability (e.g., (2) is the maximal power of a signal on a whole path). The feedback Problem 3.1 for this case is stated and solved analogously to the methods in Sections 3–6.

The recursive equation here is

$$J(H;x) = \min\left\{c_{m+1}(x,x_f);\ \min_{i\in H}\ \sup_{y\in M_i} \max\left[c_i(x,y);\ J(H\setminus\{i\};y)\right]\right\}$$

($x \in M_i$ for $i \in \overline{1,m} \setminus H$) with the boundary condition (5). The corresponding statements repeat literally as Theorems 4.1–4.3, and are omitted.

Also note that problem (3), (2) is equivalent to the generalized version of the well-known *maximum capacity path* problem:

$$I(r(s),(x_i)_{i\in\overline{1,s}}) = \min\Big\{\min_{i\in\overline{1,s}} c_{r_i}(x_{i-1},x_i);\ c_{m+1}(x_s,x_f)\Big\},$$

$$I(r(s),(x_i)_{i\in\overline{1,s}}) \to \max_{r(s)\in\{r\}} \sup_{(x_i)_{i\in\overline{1,s}}\in\prod_{i=1}^{s} M_{r_i}},$$

where $c_i(x,y)$ is the capacity (throughput) of a link (x,y), and I is the capacity of a whole path along a route $r(s)$. This open-loop problem and corresponding $\max_R \inf_{x(\cdot)}$ feedback problem are solved analogously to problem (3), (2) and Problem 3.1 respectively, substituting everywhere all operations *min*, *inf*, and *max* by *max*, *sup*, and *min*, respectively (and supposing all functions c_i, $i \in \overline{1,m}$, are bounded below).

8 GGSPP with a Set of Terminal Nodes

The GSPP and GGSPP can be easily modified to the case when instead of a single terminal point x_f one has some multiport terminal set $X_f \subset \mathbb{R}^d$ of technical endpoints of a process, and it is allowed to finish a route at any (one) point (port) in X_f. The set X_f can be continuous or discrete.

The cost to minimize is either

$$I(r(s),(x_i)_{i\in\overline{1,s}}) = \sum_{i=1}^{s} c_{r_i}(x_{i-1},x_i) + \inf_{a\in X_f} c_{m+1}(x_s,a)$$

or

$$I(r(s),(x_i)_{i\in\overline{1,s}}) = \max\Big\{\max_{i\in\overline{1,s}} c_{r_i}(x_{i-1},x_i);\ \inf_{a\in X_f} c_{m+1}(x_s,a)\Big\}$$

instead of (1) or (2) respectively. Problem 3.1 is stated as before. The recursive equation is

$$J(H;x) = \min\Big\{\inf_{a\in X_f} c_{m+1}(x,a);\ \min_{i\in H}\sup_{y\in M_i}\ [c_i(x,y) + J(H\setminus\{i\};y)]\Big\}$$

or

$$J(H;x) = \min\Big\{\inf_{a\in X_f} c_{m+1}(x,a);\ \min_{i\in H}\sup_{y\in M_i} \max\ [c_i(x,y);\ J(H\setminus\{i\};y)]\Big\}$$

($x \in M_i$ for $i \in \overline{1,m} \setminus H$) respectively, with the boundary condition

$$J(\varnothing;x) = \inf_{a\in X_f} c_{m+1}(x,a) \qquad (x\in M_i,\ i\in\overline{1,m}).$$

The solving scheme and theorems are analogous to the previous ones, and are therefore omitted.

9 Conclusion

In [55] a new approach for solving dynamic combinatorial game problems was suggested, and it was applied to solve the GGTSP. This approach is a symbiosis of combinatorial optimization and dynamic games. The global minimax problem in the class of feedback strategies is reduced to a series of local min sup operations. In this chapter the second dynamic game problem of this type is considered, namely the GGSPP.

So, we have introduced and investigated the deterministic shortest path problem with set membership uncertainties. The problems similar to ours may be observed in the stochastic branch of the SPP, where the arcs (x, y) of a network have stochastic lengths ("travel times"), i.e., are random variables. In the *a priori* least-expected time path problem the entire route is selected before a process starts. No additional information during a motion is available, and no rerouting is permitted. In this setting the actual values of arcs do not become known until the arc is passed.

But in real problems it is of use to react on current information. For example, a driver on the road network should adapt to the changing situation, and reroute a path, if necessary, to get a better total travel time. In the *time-adaptive* route problem [10,21,35,36,47,48,61] a decision-maker can reevaluate at each node the remaining path on the basis of information obtained en-route, so at each node x the rerouting is possible during a motion. Namely, the actual lengths of all downstream arcs (x, y) become known to a decision-maker upon his arrival at x. Now he knows the revealed costs of all downstream arcs (x, y) deterministically. Besides, he knows the expected cost for each successor node y. Adding these two values (for each y) and then choosing the minimum (in y), a decision-maker obtains the next arc (x, y^*) to go. The initial ideas of this approach were originated by R. W. Hall [25].

The adaptive (i.e., feedback) strategies are formalized as *hyperpaths* [38,39] (a hyperpath is an acyclic subnetwork consisting of a set of paths between a given origin-destination pair). The adaptive setting gives, in general, a better result compared to the a priori one, and, clearly, is an attempt to fit the theory to real-life problems (for example, the problems of urban traffic congestion and the shortest-path protocols for data traffic in computer networks).

REFERENCES

[1] Aubin J.-P., *Nonlinear Analysis and Motivations from Economics*, Masson, Paris, 1984.

[2] Başar T. and Bernhard P., *H^∞-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*, 2nd ed., Birkhäuser, Boston, MA, 1995.

[3] Bellman R., On a routing problem, *Quart. Appl. Math.*, 16, 1, 87–90, 1958.

[4] Berkovitz L.D., A theory of differential games, in T. Başar and A. Haurie (eds.), *Annals of the ISDG*, Vol. 1, 3–22, Birkhäuser, Boston, MA, 1994.

[5] Cherkassky B.V., Goldberg A.V. and Radzik T., Shortest paths algorithms: Theory and experimental evaluation, *Math. Program.*, 73, 2, 129–174, 1996. Also in: *Proc. 5th SODA (Arlington, VA, 1994)*, 516–525, SIAM, Philadelphia, PA, 1994.

[6] Christofides N., *Graph Theory. An Algorithmic Approach*, Academic Press, New York, 1975.

[7] Christofides N., Badra H.O. and Sharaiha Y.M., Data structures for topological and geometric operations on networks, *Annals Oper. Res.*, 71, 259–289, 1997.

[8] Clarke F.H., Ledyaev Yu.S. and Subbotin A.I., The synthesis of universal feedback pursuit strategies in differential games, *SIAM J. Contr. Optim.*, 35, 2, 552–561, 1997.

[9] Clarke F.H., Ledyaev Yu.S. and Stern R.J., Proximal analysis and feedback construction, in *Issue in Memory of Academician A.I. Subbotin, Proc. Inst. Math. and Mech. of Russian Acad. Sci.*, Vol. 6(1–2), 91–109, IMM, Ekaterinburg, 2000. Also in *Proc. Steklov Inst. Math.*, Suppl. 1, S72–S89, 2000.

[10] Davies C. and Lingras P., Genetic algorithms for rerouting shortest paths in dynamic and stochastic networks, *Eur. J. Oper. Res.*, 144, 1, 27–38, 2003.

[11] Dendris N.D., Kirousis L.M. and Thilicos D.M., Fugitive-search games on graphs and related parameters, *Theoret. Comput. Sci.*, 172, 1–2, 233–254, 1997.

[12] Dial R., Glover F., Karney D. and Klingman D., A computational analysis of alternative algorithms and labeling techniques for finding shortest path trees, *Networks*, 9, 3, 215–248, 1979.

[13] Diderich C.G. and Gengler M., A survey on minimax trees and associated algorithms, in D.-Z. Du and P.M. Pardalos (eds.), *Minimax and Applications*, 25–54, Kluwer, Dordrecht, 1995.

[14] Dreyfus S.E., An appraisal of some shortest-path algorithms, *Oper. Res.*, 17, 3, 395–412, 1969.

[15] Dreyfus S.E. and Law A.M., *The Art and Theory of Dynamic Programming*, Ch. IV, Academic Press, New York, 1977.

[16] Fomin F.V., Note on a helicopter search problem on graphs, *Discrete Appl. Math.*, 95, 1–3, 241–249, 1999.

[17] Ford L.R., Jr., Network flow theory, RAND Corp. Report P-923, Santa Monica, CA, 1956.

[18] Frank H. and Frisch I.T., *Communication, Transmission and Transportation Networks*, Addison-Wesley, Reading, MA, 1971.

[19] Friedman A., *Differential Games*, Wiley Interscience, New York, 1971.

[20] Friedman A., Differential Games, in *CBMS Regional Conf. Ser. in Math.*, Vol. 18, Amer. Math. Soc., Providence, RI, 1974.

[21] Fu L., An adaptive routing algorithm for in-vehicle route guidance systems with real-time information, *Transport. Res. Part B*, 35, 8, 749–765, 2001.

[22] Gallo G. and Pallottino S., Shortest path methods: A unifying approach, *Math. Program. Study*, 26, 38–64, 1986.

[23] Gallo G. and Pallottino S., Shortest path algorithms, *Annals Oper. Res.*, 13, 1–4, 3–79, 1988.

[24] Hajek O., *Pursuit Games*, Academic Press, New York, 1975.

[25] Hall R., The fastest path through a network with random time-dependent travel times, *Transport. Sci.*, 20, 3, 182–188, 1986.

[26] Held M. and Karp R.M., The construction of discrete dynamic programming algorithms, *IBM Systems J.*, 4, 2, 136–147, 1965.

[27] Isaacs R., *Differential Games*, Wiley, New York, 1965; Krieger, New York, 1975.

[28] Johnson D.B., Efficient algorithms for shortest paths in sparse networks, *J. ACM*, 24, 1, 1–13, 1977.

[29] Korf R.E., Recent progress in the design and analysis of admissible heuristic functions, in *Proc. 4th Intern. SARA (Horseshoe Bay, TX, 2000), Lecture Notes Computer Sci.*, Vol. 1864, 45–56, Springer, Berlin, 2000.

[30] Krasovskii N.N., *Rendezvous Game Problems*, Nauka, Moscow, 1970. Engl. transl.: National Technical Inform. Serv., Springfield, VA, 1971.

[31] Krasovskii N.N. and Subbotin A.I., *Positional Differential Games*, Nauka, Moscow, 1974. French transl.: *Jeux Différentiels*, Éditions Mir, Moscow, 1977.

[32] Krasovskii N.N., Differential games. Approximate and formal models, *Math. USSR, Sborn.*, 35, 6, 795–822, 1979.

[33] Krasovskii N.N. and Subbotin A.I., *Game-Theoretical Control Problems*, Springer, New York, 1988.

[34] Lichtenstein D. and Sipser M., Go is polynomial-space hard, *J. ACM*, 27, 2, 393–401, 1980. Also in: *Proc. 19th FOCS Sympos. (Ann Arbor, MI, 1978)*, 48–54, IEEE, Long Beach, CA.

[35] Miller-Hooks E., Adaptive least-expected time paths in stochastic, time-varying transportation and data networks, *Networks*, 37, 1, 35–52, 2001.

[36] Miller-Hooks E. and Mahmassani H., Path comparisons for a priori and time-adaptive decisions in stochastic time-varying networks, *Eur. J. Oper. Res.*, 146, 1, 67–82, 2003. Erratum: *Ibid.*, 148, 2, 466, 2003.

[37] Moore E.F., The shortest path through a maze, in *Proc. Intern. Sympos. on Theory of Switching (Harvard Univ., 1957)*, Part 2, 285–292, Harvard Univ. Press, Cambridge, MA, 1959.

[38] Nguyen S. and Pallottino S., Hyperpaths and shortest hyperpaths, in *Combinatorial Optimization, Proc. 3rd C.I.M.E. Session (Como, Italy, 1986), Lecture Notes Math.*, Vol. 1403, 258–271, Springer, Berlin, 1989.

[39] Nguyen S. and Pallottino S., Equilibrium traffic assignment for large scale transit networks, *Eur. J. Oper. Res.*, 37, 2, 176–186, 1988.

[40] Nilsson N.J., *Principles of Artificial Intelligence*, Tioga Publ. Co., Palo Alto, CA, 1980.

[41] Nobakhtian S. and Stern R.J., Universal near-optimal feedbacks, *J. Optim. Theory Appl.*, 107, 1, 89–122, 2000.

[42] Parikh S.C. and Frisch I.T., Finding the most reliable routes in communication systems, *IEEE Trans. Commun. Syst.*, 11, 4, 402-406, 1963.

[43] Pearl J., *Heuristics: Intelligent Search Strategies for Computer Problem Solving*, Addison-Wesley, Reading, MA, 1984.

[44] Petrosjan L.A., *Differential Games of Pursuit*, World Scientific, Singapore-London, 1993.

[45] Pijls W. and de Bruin A., Game tree algorithms and solution trees, *Theoret. Comput. Sci.*, 252, 1–2, 197–215, 2001. Also in: *Proc. 1st Intern. Conf. "Computers and Games" (Tsukuba, 1998), Lecture Notes Computer Sci.*, Vol. 1558, 195–204, Springer, Berlin, 1999.

[46] Pontryagin L.S., Papers on differential games, in *Selected Scientific Works*, Vol. II, 270–573, Nauka, Moscow (in Russian), 1988.

[47] Pretolani D., A directed hypergraph model for random time dependent shortest paths, *Eur. J. Oper. Res.*, 123, 2, 315–324, 2000.

[48] Psaraftis H. and Tsitsiklis J.N., Dynamic shortest paths in acyclic networks with Markovian arc costs, *Oper. Res.*, 41, 1, 91–101, 1993.

[49] Pshenichnyi B.N. and Ostapenko V.V., *Differential Games*, Naukova Dumka, Kiev (in Russian), 1992.

[50] Pultr A. and Morris F.L., Prohibiting repetitions makes playing games substantially harder, *Intern. J. Game Theory*, 13, 1, 27–40, 1984.

[51] Rivest R.L., Game tree searching by min/max approximation, *Artif. Intell.*, 34, 1, 77–96, 1987.

[52] Robson J.M., The complexity of Go, in *Information Processing'83. Proc. 9th IFIP World Computer Congress (Paris, 1983)*, 413–417, North-Holland, Amsterdam, 1983.

[53] Robson J.M., Combinatorial games with exponential space complete decision problems, in *Proc. 11th MFOCS Sympos. (Prague, 1984), Lecture Notes Computer Sci.*, Vol. 176, 498–506, Springer, Berlin, 1984.

[54] Robson J.M., Alternation with restrictions on looping, *Information and Control*, 67, 1, 2–11, 1985.

[55] Serov V.P., Optimal feedback strategy in the game variant of generalized travelling salesman problem, in *Proc. 11th IFAC Workshop "Control Applications of Optimization,"* Vol. 2, 635–640, Pergamon, Oxford, 2000.

[56] Serov V.P., *Bellman-Ford-Moore algorithm for generalized shortest paths*, 2003, submitted.

[57] von Stengel B. and Werchner R., Complexity of searching an immobile hider in a graph, *Discrete Appl. Math.*, 78, 1–3, 235–249, 1997.

[58] Subbotin A.I. and Chentsov A.G., *Optimization of a Guarantee in Control Problems*, Nauka, Moscow (in Russian), 1981.

[59] Subbotina N.N., Universal optimal strategies in positional differential games, *Differ. Eq.*, 19, 11, 1377–1382, 1983.

[60] Tohyama H. and Adachi A., Complexity of path discovery game problems, *Theoret. Comput. Sci.*, 237, 1–2, 381–406, 2000.

[61] Waller S.T. and Ziliaskopoulos A.K., On the online shortest path problem with limited arc cost dependencies, *Networks*, 40, 4, 216–227, 2002.

[62] Wimer S., Koren I. and Cederbaum I., On paths with the shortest average arc length in weighted graphs, *Discrete Appl. Math.*, 45, 2, 169–179, 1993.

[63] Wolfe D., Go endgames are PSPACE-hard, in *More Games of No Chance. Proc. 2nd Combinatorial Games Theory MSRI Workshop (Berkeley, CA, 2000)*, 125–136, Math. Sci. Res. Inst. Publ., Vol. 42, Cambridge Univ. Press, Cambridge, 2002.

[64] Yen J.Y., *Shortest Path Network Problems*, Verlag Anton Hain, Meisenheim am Glan, 1975.

[65] Zhang W., Dechter R. and Korf R.E. (eds.), Heuristic Search in Artificial Intelligence, *Artif. Intell.*, 129, 1–2, 2001 (Special Issue).

New Approach to Improve the Accuracy in Delayed Information Pursuit-Evasion Games

Josef Shinar
Technion-Israel Institute of Technology
Faculty of Aerospace Engineering
Haifa, Israel
aer4301@aerodyne.technion.ac.il

Valery Y. Glizer
Technion-Israel Institute of Technology
Faculty of Aerospace Engineering
Haifa, Israel
valery@techunix.technion.ac.il

Abstract

This chapter presents a new approach to improve the homing performance of a pursuer with delayed information on the evader's lateral acceleration. This approach reduces the *uncertainty set* of the pursuer, created due to the estimation delay, by considering not only the current (pure feedback) measurements but also the available measurement history during the period of the estimation delay. The reduced *uncertainty set* is computed by solving two auxiliary optimization problems. By using the center of the new *uncertainty set's* convex hull as a new state variable, the original game is transformed to a nonlinear delayed dynamics game with perfect information for both players. The solution of this new game is obtained in pure strategies for the pursuer and mixed ones for the evader. The *value* of this game (the guaranteed miss distance) is substantially less than the one obtained in previous works by using only the current measurements.

Key words. Pursuit-evasion games, imperfect information, estimation delay, uncertainty set

1 Introduction

Interception end-game scenarios of maneuvering targets can be formulated as zero-sum pursuit-evasion games with bounded controls and prescribed duration. The interceptor is the *pursuer*, the maneuvering target is the *evader* and the *cost function* of the game is the miss distance. In most cases the use of a linearized kinematical model with first-order acceleration dynamics of both players is justified. Using such a model and assuming perfect information allows one to obtain a closed-form solution of the game [28]. The first step in the solution is to reduce the dimension of the original game (four state variables in a planar interception) by introducing a new (scalar for a planar interception) state variable, called the *zero effort miss distance*. This is the miss distance that would be obtained if both players used zero control until the end of the game. The *zero effort miss distance*, being based on the homogeneous solution of the original linear system, is the function of all the original state variables. The perfect information game solution, using the *time-to-go* as the independent variable, consists of the optimal interceptor guidance law (the optimal *pursuer* strategy), the "worst" target maneuver (the optimal *evader* strategy) and the guaranteed miss distance (the *value* of the game). This solution yields a decomposition of the reduced game space into two regions of different optimal strategies. In the *regular* region, the optimal strategies are unique and the *value* of the game is a function of the initial conditions. In the *singular* region, where almost all practical initial conditions are located, the optimal strategies are arbitrary and the *value* of the game is constant, depending on two nondimensional parameters, namely the *pursuer/evader* maneuver ratio (denoted by μ) and the ratio of the *evader/pursuer* (first-order) time constants (denoted by ε). If $\mu > 1$ and the product $\mu\varepsilon$ satisfies the inequality $\mu\varepsilon \geq 1$, the *value* of the game vanishes (a most desired *ideal* interception outcome).

Unfortunately, a realistic interception scenario is one of imperfect information. The *evader* has no information on the *pursuer*, while the *pursuer* has noise corrupted measurements on the relative position of the *evader*. Due to the partial and noisy measurements, guidance law implementation requires one to use an estimator. The estimated state variables are the relative position, relative velocity and the acceleration of the *evader*. The acceleration of the *pursuer* is assumed to be measured with high accuracy. Based on the linearized kinematical model, only the components normal to the line of sight are considered. Even if the accuracy and the convergence of the position estimate are satisfactory, the velocity estimate is less precise and it converges more slowly. The accuracy of the estimated acceleration is even worse and its convergence is the slowest. Implementing the optimal *pursuer* strategy of the perfect information game as the interceptor's guidance law in spite of the information delay yields very disappointing, but predictable, results [10]. A smart or lucky *evader* can execute a maneuver that maximizes the miss distance, even if the perfect information game solution guarantees a direct hit.

In simulation studies dealing with the interception of maneuvering targets in noise corrupted scenarios, the delay caused by the slowly converging target acceleration estimate was found to be the dominant source of miss distances. Based on these observations, in earlier investigations [29,32,34] it was assumed that the estimation process of the evader's acceleration can be roughtly approximated by a perfect information outcome delayed by the amount of Δt, while the other variables can be considered as accurate. This approximation leads to the formulation of the interception scenario with noisy measurements as a deterministic pursuit-evasion game with delayed information of the *pursuer*, allowing perfect information for the *evader* (the "worst case" situation for the pursuer).

The family of delayed information differential games is a part of the very wide class of dynamic problems with delay either in the information or in the state or in the control.

Differential equations with delay (with given inputs) have been studied more extensively than the other families of dynamic problems with delay. In the monographs [2,9,12,14,17], existence and uniqueness of solution, stability, approximate solution, and many other important questions were studied for delayed differential equations.

Control problems with delays were analyzed in many works in the open literature, see, for instance, the recent surveys [11,26,27]. Control problems with information delay were considered in [1,7,8], where the stability of feedback control was analyzed. H_∞-control problems with delays were studied in the monographs [3,37] as well as in numerous other works.

Various linear differential games with state delay were solved in [18,19,22,24]. In [23], a nonlinear differential game with state delay was analyzed. Zero-sum differential games with state delay were solved in [13,21]. The linear-quadratic game considered in [13] is a direct extension of the pursuit-evasion game solved in [15]. The solution of this game with state delay is based, similarly to [15], on its reduction to a game with much simpler dynamics which is independent of the state variable. The game considered in [21] is an extension of the linear-quadratic game solved in [13]. The cost functional in this game depends not only on the terminal state of the game but also on a number of the intermediate ones. Nevertheless, the approach to the solution is similar to those in [15,13].

Two-player differential games with a time delay in the information were analyzed in a number of works in the open literature. The works [6,25,35] consider the case when each of the players controls his own differential system and the first player gets the information on current values of all components of the state vector of the second player with the same constant delay. In [35], the original game is reduced to an equivalent perfect information game replacing the state vector of the second player by a new state variable, namely by its delayed state vector. An extension of the Hamilton–Jacobi theory and the main Isaacs's equation is presented in [6] under the assumption of separability of the integral part of the cost functional. In [25], the game is solved by using the *reachable set* concept under the assumption that the first player can ensure, by the final time of the game, the capture of the center of a minimal sphere containing the *reachable (uncertainty) set* associated with the second player's motion and the information delay (the Center Capture Assumption). The papers [4,5] consider the case when both players control the same linear differential time-independent system and the first player gets the information on current values of all components of the state vector with the same time-dependent delay. The objective of the first player is to transfer the state of the game from a given initial position to a given set in the game's state space in the shortest possible time against all admissible second player's controls. This game is reduced to an equivalent perfect information pursuit-evasion game by replacing the original state vector with a new one. This new state vector is the solution of the original equations of motion integrated during the delay assuming an arbitrary control of the first player and no control of the second player. Such a reduction is obtained assuming that the delay is a differentiable function of the time, and its derivative is less than unity. The dynamics of the new game, as well as the terminal set, become time dependent.

A more general case (from a theoretical viewpoint) that is also more reasonable (from a practical viewpoint) is the situation where the different components of the common state (or the state of the second player) are available to the first player with different values of time delay. In this case, in order to derive the optimal control, the first player can use not only the current (delayed) measurements, but also the available measurement history. This approach can lead to a considerable reduction of the information uncertainty. This general case has not yet been treated in the open literature.

A particular situation of this general case has been studied by the authors in several works (including this chapter). Namely, it has been assumed that the pursuer obtains only the current value of the evader's lateral acceleration with a time delay (either constant or time dependent), while the other components of the state vector are perfectly measured. In the earlier investigations [32,29–31], applying the concept of *reachable set*, suggested in [25] and then developed in [20] and some other works, the *zero effort miss distance* was replaced by the center of the convex hull of the *pursuer's uncertainty set* created by the delayed information. The rigorous solution of such a delayed information game [29,31] led to an interceptor guidance law that partially compensated for the estimation delay, substantially improving the guaranteed homing performance. However, due to the estimation delay, zero miss distance cannot be guaranteed even if the conditions $\mu > 1$ and $\mu\varepsilon \geq 1$ are satisfied. This (deterministic) improvement was confirmed by a set of Monte Carlo simulations with noisy measurements and a Kalman filter-type estimator in the guidance loop [33], but the residual reduced guaranteed miss distances were not sufficiently small.

The objective of this chapter is to outline a new approach aimed towards further improvement of the homing accuracy in interception scenarios against maneuvering targets with noisy measurements. This approach is aimed to reduce the *uncertainty set* of the *pursuer* due to its (constant) estimation delay, by considering in addition to the current (pure feedback) measurements also the available measurement history during the period of the estimation delay and taking into account the bounds on the *evader's* control. Using the additional measurements, the reduced *uncertainty set* of the *pursuer* is computed by solving two auxiliary optimization problems, which allows us to find a new center of the *uncertainty set's* convex hull, to serve as a new aim point. Based on this approach a newly formulated pursuit-evasion game of delayed information has to be solved, leading to a new game solution and to the synthesis of an improved interceptor guidance law. The solution methodology of the new game follows the steps of the earlier delayed information game solution [29,34]. However, there are substantial differences in the mathematical techniques due to the different assumptions on the available information in the game formulation.

The structure of the paper is the following. In the next section the problem of intercepting a maneuvering target is formulated, followed by brief outlines of the perfect information game solution [28] and the earlier solution of the delayed information game [29]. Section 3 is devoted to the description of the reduced *uncertainty set* of the pursuer and the solution of the respective delayed information game by transforming it to a perfect information game with delay in the evader's control. The problem of obtaining the optimal initial condition for the evader's control is formulated and solved in Section 4. In Section 5 the game space decomposition is presented. Proofs are given in the Appendices.

2 Problem Formulation

2.1 Interception Dynamics

The mathematical model of the interception end game is based on the geometry shown in Figure 1 and on the following assumptions:

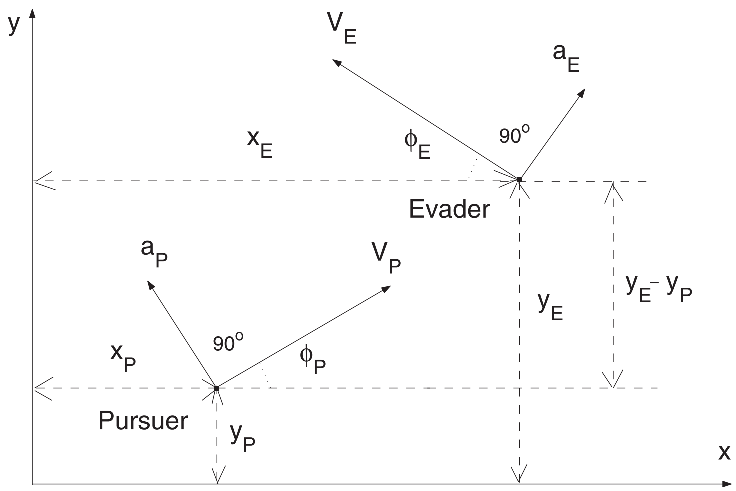


Figure 1: Interception geometry.

(A1) The interception end game starts when the interceptor missile (pursuer "P") is locked on the target (evader "E").
(A2) The pursuit-evasion is planar.
(A3) Both players have constant velocities V_i and bounded commanded lateral accelerations $|a_i| \leq A_i$, $(i = P, E)$, where A_i are given positive constants.
(A4) The dynamics of each player is given by a first-order transfer function with the time constant τ_i, $(i = P, E)$.
(A5) The angles ϕ_i, $(i = P, E)$ between the initial line of sight and the velocity vectors of the players are sufficiently small, yielding the approximations $\cos\phi_i \approx 1$, $\sin\phi_i \approx \phi_i$.

Using (A3) and (A5), one can calculate the duration of the game end for a known initial distance R_0 between the players

$$t_f = R_0/(V_P + V_E). \tag{1}$$

The state vector of the interception is

$$x \triangleq (x_1, x_2, x_3, x_4)^T, \tag{2}$$

where

$$x_1 \triangleq y_E - y_P \tag{3}$$

is the relative position, x_2 is the relative velocity, and x_3 and x_4 are the lateral accelerations of the pursuer and the evader, respectively. All the variables are normal to the initial line of sight.

The respective nondimensional controls of the pursuer and the evader are

$$u \triangleq a_P/A_P, \qquad |u| \leq 1, \tag{4}$$

$$v \triangleq a_E/A_E, \qquad |v| \leq 1. \tag{5}$$

Using these definitions and the assumptions (A1)–(A5), the mathematical model of the interception is

$$dx_1/dt = x_2, \tag{6}$$
$$dx_2/dt = x_4 - x_3, \tag{7}$$
$$dx_3/dt = (A_P u - x_3)/\tau_P, \tag{8}$$
$$dx_4/dt = (A_E v - x_4)/\tau_E. \tag{9}$$

The pursuit-evasion starts at $t = 0$. The cost function is the miss distance

$$J \triangleq |x_1(t_f)| \to \min_u \max_v . \tag{10}$$

2.2 Perfect Information Game

In [28], a zero-sum pursuit-evasion game, described by (4)–(10), was solved under the assumption that the players have perfect information on the parameters and state variables of the problem. This game is called the perfect information game (PIG). The PIG solution has been obtained by reducing the system (6)–(9) to a single equation using the following transformation:

$$z(t) = DF(t_f, t)x(t), \tag{11}$$

where $z(t)$ is the new state variable (*zero effort miss distance*), $D = (1, 0, 0, 0)$ and $F(t_f, t)$ is the transition matrix of the system (6)–(9).

In the sequel, nondimensional variables will be used: the normalized time-to-go θ

$$\theta \triangleq (t_f - t)/\tau_P, \tag{12}$$

and the normalized state variable $\bar{z}(\theta)$

$$\bar{z}(\theta) \triangleq z(t_f - \tau_P\theta)/(A_E\tau_P^2), \tag{13}$$

yielding [28]

$$\bar{z}(\theta) = \bar{z}^0(\theta) + \varepsilon^2\Psi(\theta/\varepsilon)\bar{x}_4(\theta), \tag{14}$$

where

$$\bar{z}^0(\theta) \triangleq \bar{x}_1(\theta) + \theta\bar{x}_2(\theta) - \mu\Psi(\theta)\bar{x}_3(\theta), \tag{15}$$

$$\bar{x}_1(\theta) \triangleq x_1(t_f - \tau_P\theta)/(A_E\tau_P^2), \qquad \bar{x}_2(\theta) \triangleq x_2(t_f - \tau_P\theta)/(A_E\tau_P), \tag{16}$$

$$\bar{x}_3(\theta) \triangleq x_3(t_f - \tau_P\theta)/A_P, \qquad \bar{x}_4(\theta) \triangleq x_4(t_f - \tau_P\theta)/A_E, \tag{17}$$

$$\Psi(\theta) \triangleq \exp(-\theta) + \theta - 1, \tag{18}$$

$$\varepsilon \triangleq \tau_E/\tau_P, \qquad \mu \triangleq A_P/A_E. \tag{19}$$

Note that $\Psi(\theta)$ is positive for all $\theta > 0$.

Using (11)–(13), the PIG dynamics (6)–(9) becomes

$$d\bar{z}(\theta)/d\theta = \mu\Psi(\theta)\bar{u}(\theta) - \varepsilon\Psi(\theta/\varepsilon)\bar{v}(\theta), \tag{20}$$

where

$$\bar{u}(\theta) \triangleq u(t_f - \tau_P\theta), \qquad |\bar{u}(\theta)| \le 1, \tag{21}$$

$$\bar{v}(\theta) \triangleq v(t_f - \tau_P\theta), \qquad |\bar{v}(\theta)| \le 1, \tag{22}$$

and the cost function (10) becomes

$$\bar{J} \triangleq |\bar{z}(0)| \to \min_{\bar{u}} \max_{\bar{v}}, \tag{23}$$

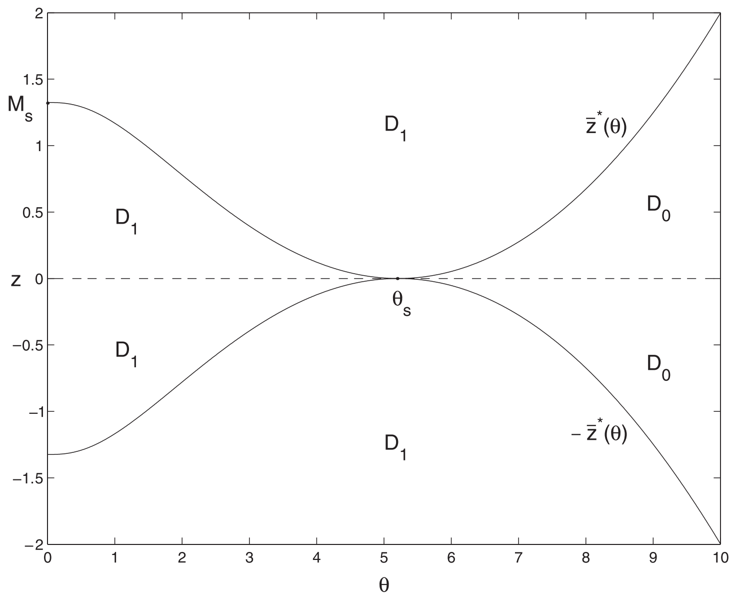


Figure 2: PIG space decomposition for $\mu > 1$, $\mu\varepsilon < 1$, ($\mu = 1.18$, $\varepsilon = 0.25$).

The game starts at $\theta_0 \triangleq t_f/\tau_P$.

The PIG solution yields [28] a decomposition of the reduced game space $(\theta, \bar{z})$ into two regions of different optimal strategies, as can be seen in Figure 2.

The pair of optimal trajectories $\bar{z}^*(\theta)$ and $-\bar{z}^*(\theta)$, which reach the θ-axis tangentially at $\theta = \theta_s$ (the single positive root of $\Gamma(\theta) \triangleq \mu\Psi(\theta) - \varepsilon\Psi(\theta/\varepsilon)$), are the boundary trajectories between the two regions. In the regular region

$$\mathcal{D}_1 \triangleq \{(\theta, \bar{z}) \in (\theta \leq \theta_s) \cup (|\bar{z}| \geq \bar{z}^*(\theta))\},$$

the optimal feedback strategies are

$$\bar{u}^*[\theta, \bar{z}(\theta)] = \bar{v}^*[\theta, \bar{z}(\theta)] = \text{sign}[\bar{z}(\theta)], \quad \bar{z}(\theta) \neq 0, \tag{24}$$

and the *value* of the game is a function of the initial conditions.

In the singular region

$$\mathcal{D}_0 \triangleq \{(\theta, \bar{z}) \in (\theta > \theta_s) \cap (|\bar{z}| < \bar{z}^*(\theta))\},$$

the optimal strategies are arbitrary, satisfying the constraints (21) and (22). Every trajectory starting in this region must go through the "throat" $(\theta, \bar{z}) = (\theta_s, 0)$, which is a *dispersal point* dominated by the evader. Consequently, the *value* of the game in this entire region is constant, denoted as M_s. Both M_s

and θ_s depend on the parameters μ and ε. These two values (M_s and θ_s) are of extreme importance, because almost all practical initial conditions are in $\mathcal{D}_0$. If $\mu > 1$ and $\mu\varepsilon \geq 1$, both θ_s and M_s vanish (a most desired *ideal* interception outcome). In the sequel, the paper concentrates on this case.

2.3 Imperfect Information Game

In [29], the case of imperfect information was considered. Namely, it was assumed that the pursuer has the information on the state variable x_4 with a constant time delay Δt, but the other state variables are perfectly observed. Thus, the information available to the pursuer at any instant t is described by the vector

$$w_s(t) = (x_1(t), x_2(t), x_3(t), x_4(t - \Delta t))^T. \tag{25}$$

The evader has perfect information on all the state variables of the game and also knows that $w_s(t)$ is the pursuer's information vector. This game is called the simple delayed information game (SDIG).

The SDIG solution was obtained in [29] by using the concept of *uncertainty domain* [25]. A current value of the nondimensional state variable $\bar{x}_4$ was replaced by the center of its reachable (*uncertainty*) set, constructed by using the information vector $w_s(t)$. For the nondimensional variables $(\theta, \bar{x}_4)$, this set is an interval $[\bar{x}_{4s}^-(\theta), \bar{x}_{4s}^+(\theta)]$, the bounds of which are obtained as the outcome of the following auxiliary dual optimal control problem:

$$\bar{x}_4(\sigma = \theta) \to \min_{\bar{v}} \quad (\max_{\bar{v}}) \tag{26}$$

along trajectories of the equation

$$\frac{d\bar{x}_4(\sigma)}{d\sigma} = \frac{[\bar{x}_4(\sigma) - \bar{v}(\sigma)]}{\varepsilon}, \quad \bar{x}_4(\sigma)|_{\sigma=\theta+\Delta\theta} = \bar{x}_4(\theta + \Delta\theta), \quad \Delta\theta \triangleq \frac{\Delta t}{\tau_P} \tag{27}$$

subject to the evader's control constraint (22). The center $[\bar{x}_{4s}(\theta)]_c$ of this simple *uncertainty set* is given by

$$[\bar{x}_{4s}(\theta)]_c = \bar{x}_4(\theta + \Delta\theta) \exp(-\Delta\theta/\varepsilon). \tag{28}$$

The solution of the SDIG was obtained by transforming it to a perfect information game using a new state variable

$$\bar{z}_s^c(\theta) = \bar{z}^0(\theta) + \varepsilon^2 \Psi(\theta/\varepsilon)[\bar{x}_{4s}(\theta)]_c. \tag{29}$$

As a consequence, the dynamics of the SDIG becomes

$$\begin{aligned} d\bar{z}_s^c(\theta)/d\theta = {} & \mu\Psi(\theta)\bar{u}(\theta) - \varepsilon\Psi(\theta/\varepsilon) \exp(-\Delta\theta/\varepsilon)\bar{v}(\theta + \Delta\theta) \\ & + (\theta/\varepsilon) \int_{\Delta\theta}^{0} \exp(-\sigma/\varepsilon)\bar{v}(\theta + \sigma)d\sigma, \end{aligned} \tag{30}$$

indicating a delay in the evader's control and the perfect information nature of the new game. In this game, in order to avoid ambiguity, the value of the evader's control $\bar{v}(\theta)$ for $\theta > \theta_0$ is assumed to be constant, equal to $\bar{v}(\theta_0)$.

The decomposition structure of the SDIG space $(\theta, \bar{z}_s^c)$ is qualitatively similar to the PIG solution in the $(\theta, \bar{z})$ space shown in Figure 2. In the singular region $\mathcal{D}_0^c$, the pursuer's optimal strategy is arbitrary, satisfying (21). The evader's optimal strategy is arbitrary subject to (22) for $\theta > \theta_s^c + \Delta\theta$, where $\theta = \theta_s^c$ is the coordinate of the "throat" in the SDIG. At $\theta = \theta_s^c + \Delta\theta$ the evader must choose (with the probability 0.5) either $\bar{v}(\theta_s^c + \Delta\theta) = 1$ or $\bar{v}(\theta_s^c + \Delta\theta) = -1$ and keeps $\bar{v}(\theta) = \bar{v}(\theta_s^c + \Delta\theta)$ for all $\theta \in (0, \theta_s^c + \Delta\theta)$.

Due to the information delay, the value of θ_s^c and, hence, M_s^c (the *value* of SDIG in $\mathcal{D}_0^c$) is never zero (even if $\mu > 1$, $\mu\varepsilon \geq 1$). The resulting guaranteed normalized miss distances are, however, much smaller than those predicted for the case where the pursuer uses the optimal PIG strategy [10]. The performance degradation of the pursuer in the SDIG, compared to the perfect information game, is a monotonically increasing function of the normalized delay $\delta \stackrel{\triangle}{=} \Delta t/\tau_E = \Delta\theta/\varepsilon$. If $\delta \to +\infty$, this function tends to the value of PIG in the case where $\tau_E = 0$ ($x_4(t) \equiv v(t)$).

3 Improved Delayed Information Game (IDIG)

In this chapter, a new approach to the solution of the delayed information game is proposed assuming that in addition to the current information vector $w_s(t)$, the history of the "undelayed" state variables of the game during the period of the delay Δt is also available to the pursuer.

The pursuer has the following information at any instant t:

$$w_{im} = \{x_i(\sigma), \quad \sigma \in [t - \Delta t, t], \quad (i = 1, 2, 3), \quad x_4(t - \Delta t)\}. \tag{31}$$

Thus, the IDIG consists of the dynamics (6)–(9), the control constraints (4), (5), the cost function (10) and the pursuer's information set (31). The evader has perfect information.

3.1 Improved (Reduced) Uncertainty Set

Using the information (31), one can obtain an additional control constraint for the dual optimization problem defined by (22), (26), (27).

Integrating the nondimensional form of Equations (9), (7) and (6) consecutively on the interval $[\theta + \Delta\theta, \theta]$ yields

$$\begin{aligned} \bar{x}_1(\theta) = \bar{x}_1(\theta + \Delta\theta) + \Delta\theta \bar{x}_2(\theta + \Delta\theta) - \mu \int_\theta^{\theta+\Delta\theta} (\sigma - \theta)\bar{x}_3(\sigma)d\sigma \\ + \varepsilon^2 \Psi(\delta)\bar{x}_4(\theta + \Delta\theta) + \varepsilon \int_\theta^{\theta+\Delta\theta} \Psi((\sigma - \theta)/\varepsilon)\bar{v}(\sigma)d\sigma, \end{aligned} \tag{32}$$

leading to the following additional integral control constraint

$$\int_{\theta}^{\theta+\Delta\theta} \Psi((\sigma-\theta)/\varepsilon)\bar{v}(\sigma)d\sigma = \bar{q}(\theta), \tag{33}$$

where

$$\bar{q}(\theta) \triangleq \left[\bar{x}_1(\theta) - \bar{x}_1(\theta+\Delta\theta) - \Delta\theta\bar{x}_2(\theta+\Delta\theta) + \mu\int_{\theta}^{\theta+\Delta\theta}(\sigma-\theta)\bar{x}_3(\sigma)d\sigma\right] \Big/ \varepsilon$$
$$- \varepsilon\Psi(\delta)\bar{x}_4(\theta+\Delta\theta). \tag{34}$$

Denote by $\bar{v}^-_{im}(\sigma)$ and $\bar{x}^-_{4im}(\theta)$ the solution (the optimal control and the optimal cost) of the auxiliary problem (22), (26), (27), (33) minimizing the cost function. The solution of this problem maximizing the cost function is denoted by $\bar{v}^+_{im}(\sigma)$ and $\bar{x}^+_{4im}(\theta)$. Denote also

$$\bar{q}^- \triangleq \int_{\theta}^{\theta+\Delta\theta} \Psi((\sigma-\theta)/\varepsilon)\bar{v}(\sigma)d\sigma|_{\bar{v}(\sigma)\equiv -1} = \varepsilon[\Psi(\delta)-\delta^2/2], \tag{35}$$

$$\bar{q}^+ \triangleq \int_{\theta}^{\theta+\Delta\theta} \Psi((\sigma-\theta)/\varepsilon)\bar{v}(\sigma)d\sigma|_{\bar{v}(\sigma)\equiv 1} = -\bar{q}^-. \tag{36}$$

Since $\Psi(\delta) - \delta^2/2 < 0 \ \forall\ \delta > 0$, one has $\bar{q}^- < 0 < \bar{q}^+$.

Lemma 3.1. *If the pair of inequalities*

$$\bar{q}^- \le \bar{q}(\theta) \le \bar{q}^+ \tag{37}$$

is not satisfied, the auxiliary optimal control problem (22), (26), (27), (33) has no solution. If (37) is satisfied, there is a unique solution in the following form:

$$\bar{v}^-_{im}(\sigma) = \begin{cases} 1, & \sigma \in [\theta+\Delta\theta, \sigma^-), \\ -1, & \sigma \in (\sigma^-, \theta], \end{cases} \tag{38}$$

$$\bar{x}^-_{4im}(\theta) = \bar{x}_4(\theta+\Delta\theta)\exp(-\delta) + [2\exp((\theta-\sigma^-)/\varepsilon) - \exp(-\delta) - 1], \tag{39}$$

$$\bar{v}^+_{im}(\sigma) = \begin{cases} -1, & \sigma \in [\theta+\Delta\theta, \sigma^+), \\ 1, & \sigma \in (\sigma^+, \theta], \end{cases} \tag{40}$$

$$\bar{x}^+_{4im}(\theta) = \bar{x}_4(\theta+\Delta\theta)\exp(-\delta) - [2\exp((\theta-\sigma^+)/\varepsilon) - \exp(-\delta) - 1], \tag{41}$$

where the switch points σ^- and σ^+ of the optimal controls $\bar{v}^-_{im}(\sigma)$ and $\bar{v}^+_{im}(\sigma)$, respectively, are given by

$$\sigma^j = \theta + \varepsilon B(\alpha^j), \qquad (j = -, +), \tag{42}$$

$$\alpha^- = \bar{q}(\theta)/(2\varepsilon) + \beta, \qquad \alpha^+ = -\bar{q}(\theta)/(2\varepsilon) + \beta, \tag{43}$$

$$\beta = [\exp(-\delta) - \delta^2/2 + \delta + 1]/2, \tag{44}$$

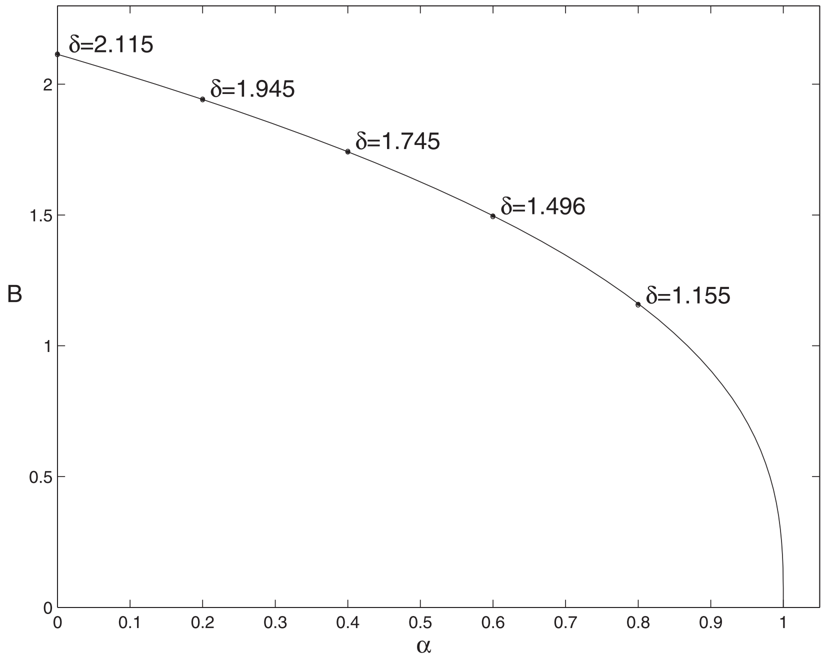


Figure 3: Graph of the function $B(\alpha)$.

and $B = B(\alpha)$ is an implicit function defined by the equation

$$\exp(-B) - 0.5B^2 + B = \alpha. \tag{45}$$

The function $B(\alpha)$ exists and is continuous for $\alpha \in [\alpha_0, 1]$, where $\alpha_0 \triangleq \exp(-\delta) - \delta^2/2 + \delta$. Its values vary from δ to 0 as α varies from α_0 to 1. Moreover, $B(\alpha)$ is differentiable on the interval $[\alpha_0, 1)$.

The proof of the lemma is presented in Appendix A.

In Figure 3, the graph of the function $B(\alpha)$ is depicted. The values of $\delta = 2.115, 1.945, \dots, 0$ correspond to the values of $\alpha_0 = 0, 0.2, \dots, 1$.

Based on Equations (39), (41) and (42) and using Equation (28), one directly obtains the center of the improved *uncertainty set* as

$$[\bar{x}_{4im}(\theta)]_c = [\bar{x}_{4s}(\theta)]_c + \exp(-B(\alpha^-)) - \exp(-B(\alpha^+)). \tag{46}$$

Corollary 3.1. *If the evader uses the control $\bar{v} = 1$ (-1) almost everywhere on any interval $[\theta + \Delta\theta, \theta]$, then $[\bar{x}_{4im}(\theta)]_c = \bar{x}_4(\theta)$.*

The proof of the corollary is presented in Appendix B.

In Figure 4, time histories of $\bar{x}_4(\theta)$, $[\bar{x}_{4s}(\theta)]_c$ and $[\bar{x}_{4im}(\theta)]_c$, corresponding

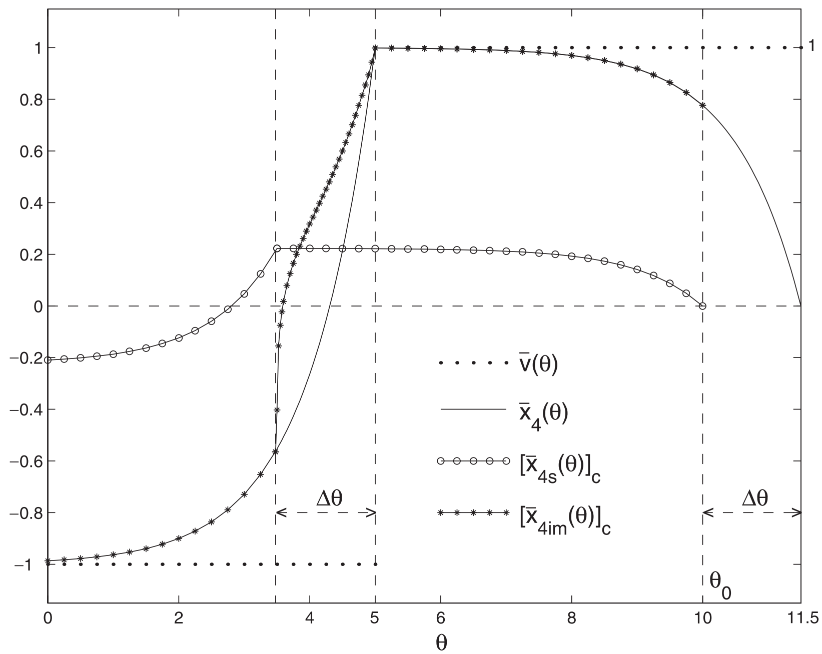


Figure 4: Actual normalized evader's acceleration and centers of respective uncertainty sets ($\mu = 1.5$, $\varepsilon = 1.0$, $\Delta\theta = 1.5$).

to an example where the evader's control is switched from 1 to -1 at $\theta = 5$, are depicted. The parameters of the game are $\mu = 1.5$, $\varepsilon = 1.0$, $\Delta\theta = 1.5$, $\theta_0 = 10$, thus the evader starts its control at $\theta_0 + \Delta\theta = 11.5$. Moreover, $\bar{x}_4(11.5) = 0$.

As seen from this figure, the center of the improved *uncertainty set* $[\bar{x}_{4im}(\theta)]_c$ coincides with the evader's true normalized lateral acceleration $\bar{x}_4(\theta)$ in the interval $\theta \in [10, 5]$, where $\bar{v}(\theta) = 1$. During the period of $\Delta\theta$ after the switch point, $[\bar{x}_{4im}(\theta)]_c$ lags behind the variations of $\bar{x}_4(\theta)$, but once the delay is over, $[\bar{x}_{4im}(\theta)]_c$ becomes equal again to $\bar{x}_4(\theta)$. The center of the simple *uncertainty set*, used in [29], $[\bar{x}_{4s}(\theta)]_c$ is quite different than $\bar{x}_4(\theta)$ on the whole interval $\theta \in [0, 10]$.

Remark 3.1. By using the uncertainty set $[\bar{x}^-_{4im}(\theta), \bar{x}^+_{4im}(\theta)]$ for $\bar{x}_4(\theta)$, one can construct the uncertainty set $[\bar{z}^-_{im}(\theta), \bar{z}^+_{im}(\theta)]$ for the normalized zero effort miss distance $\bar{z}(\theta)$. The boundaries $\bar{z}^j_{im}(\theta)$, $(j = -, +)$ of this uncertainty set are obtained by substituting $\bar{x}_4(\theta) = \bar{x}^j_{4im}(\theta)$ into Equation (14). It can be directly calculated that $\bar{z}^-_{im}(0) = \bar{z}^+_{im}(0) = \bar{x}_1(0)$, i.e., the uncertainty set for $\bar{z}(\theta)$ vanishes at the final time $\theta = 0$, becoming a point $([\bar{z}^-_{im}(0), \bar{z}^+_{im}(0)] = \bar{x}_1(0))$. This implies that the Center Capture Assumption is not satisfied for the considered game. Otherwise, the pursuer could provide zero game value in spite of the information delay, which is impossible.

3.2 Transformation of the IDIG

By introducing a new state variable, the center of the uncertainty set for the normalized zero-effort miss distance,

$$\bar{z}^c_{im}(\theta) = \bar{z}^0(\theta) + \varepsilon^2\Psi(\theta/\varepsilon)[\bar{x}_{4im}(\theta)]_c \tag{47}$$

and by using (29) and (46) one obtains

$$\bar{z}^c_{im}(\theta) = \bar{z}^c_s(\theta) + \varepsilon^2\Psi(\theta/\varepsilon)[\exp(-B(\alpha^-)) - \exp(-B(\alpha^+))]. \tag{48}$$

Differentiating (48) with respect to θ and using Equation (29) yield the dynamics

$$\begin{aligned} d\bar{z}^c_{im}(\theta)/d\theta = {} & \mu\Psi(\theta)\bar{u}(\theta) - \varepsilon\Psi(\theta/\varepsilon)\exp(-\delta)\bar{v}(\theta+\Delta\theta) \\ & + (\theta/\varepsilon)\int_{\Delta\theta}^{0}\exp(-\sigma/\varepsilon)\bar{v}(\theta+\sigma)d\sigma \\ & + \varepsilon[1-\exp(-\theta/\varepsilon)]Q(\Phi^-,\Phi^+) + E(\theta,\bar{v}(\cdot)), \end{aligned} \tag{49}$$

where

$$\begin{aligned} E(\theta,\bar{v}(\cdot)) \triangleq {} & -(\varepsilon/2)\Psi(\theta/\varepsilon)R(\Phi^-,\Phi^+)[\Psi(\delta)\bar{v}(\theta+\Delta\theta) \\ & + (1/\varepsilon)\int_{\Delta\theta}^{0}(1-\exp(-\sigma/\varepsilon))\bar{v}(\theta+\sigma)d\sigma], \end{aligned} \tag{50}$$

$$Q(\alpha_1,\alpha_2) \triangleq \exp(-B(\alpha_1)) - \exp(-B(\alpha_2)), \tag{51}$$

$$R(\alpha_1,\alpha_2) \triangleq -[\exp(-B(\alpha_1))/\Psi(B(\alpha_1)) + \exp(-B(\alpha_2))/\Psi(B(\alpha_2))], \tag{52}$$

$$\Phi^- \triangleq -(1/2\varepsilon)\int_{\Delta\theta}^{0}\Psi(\sigma/\varepsilon)\bar{v}(\theta+\sigma)d\sigma + \beta, \tag{53}$$

$$\Phi^+ \triangleq (1/2\varepsilon)\int_{\Delta\theta}^{0}\Psi(\sigma/\varepsilon)\bar{v}(\theta+\sigma)d\sigma + \beta. \tag{54}$$

Based on (47), the cost function (10) becomes

$$\bar{J}_{im} \triangleq |\bar{z}^c_{im}(0)| \rightarrow \min_{\bar{u}}\max_{\bar{v}}. \tag{55}$$

Equation (49), describing the dynamics of the IDIG in the game space $(\theta, \bar{z}^c_{im})$, is linear with respect to $\bar{u}$ but nonlinear with respect to $\bar{v}$. Moreover, the evader's control $\bar{v}$ is delayed. However, the information on the current value of the state variable $\bar{z}^c_{im}$ is perfectly available to the pursuer.

Remark 3.2. Calculating Φ^- with $\bar{v}(\sigma) = 1$ and Φ^+ with $\bar{v}(\sigma) = -1$ for almost all $\sigma \in [\theta, \theta + \Delta\theta]$, one obtains

$$\Phi^- = \Phi^+ = 1. \tag{56}$$

For any other form of $\bar{v}(\sigma)$, $\sigma \in [\theta, \theta + \Delta\theta]$, satisfying (22), the values of the functionals Φ^- and Φ^+ satisfy the inequality

$$\alpha_0 \le \Phi^j < 1, \qquad (j = -, +). \tag{57}$$

Since $B(1) = 0$ and $\Psi(0) = 0$, Equations (52) and (56) imply that the right-hand part of Equation (50) becomes infinity if either $\bar{v}(\sigma) = 1$ or $\bar{v}(\sigma) = -1$ almost everywhere on any interval $[\theta, \theta+\Delta\theta]$. In order to avoid it, one has to use in the dynamics (49) a slightly more narrow constraint for $\bar{v}$ than (22), namely,

$$|\bar{v}| \le 1 - \nu, \tag{58}$$

where $\nu > 0$ is a small parameter ($\nu \ll 1$).

Thus, based on (47), the IDIG is transformed into a new perfect information game with the dynamics (49), the control constraints (21) and (58) and the cost function (55). This game is called the nonlinear delayed dynamics game (NLDDG). Since the evader's control $\bar{v}$ appears in the game dynamics (49) with a delay, this control has to be specified on the interval $[\theta_0 + \Delta\theta, \theta_0)$, as is usually done in the case of a delayed control (see, for instance, [36] and the list of references therein). One can set

$$\bar{v}(\theta) = \bar{v}_0(\theta), \qquad \theta \in [\theta_0 + \Delta\theta, \theta_0), \tag{59}$$

where $\bar{v}_0(\theta)$ is some given piecewise differentiable function satisfying the inequality (58).

3.3 Necessary Conditions of Optimality

Proposition 3.1. *The candidate optimal controls in the NLDDG are*

$$\bar{u}^*(\theta) = -\text{sign}[\varphi(\theta)], \tag{60}$$

$$\bar{v}^*(\theta) = -(1 - \nu)\text{sign}[\lambda(\theta, 0)]. \tag{61}$$

The function $\varphi(\theta)$, defined on the interval $\theta \in [0, \theta_0]$, and the function $\lambda(\theta, \sigma)$, defined in the domain $(\theta, \sigma) \in \Omega \triangleq [0, \theta_0] \times [0, \Delta\theta]$, satisfy the following equations:

$$d\varphi(\theta)/d\theta = 0, \tag{62}$$

$$\partial\lambda(\theta, \sigma)/\partial\theta - \partial\lambda(\theta, \sigma)/\partial\sigma = \varphi(\theta)G(\theta, \sigma, \bar{v}^*(\cdot)), \tag{63}$$

where

$$
\begin{aligned}
G(\theta,\sigma,\bar{v}^*(\cdot)) \triangleq\ & (\theta/\varepsilon)\exp(-\sigma/\varepsilon) \\
& + 0.5R[(\Phi_\xi^-)^*, (\Phi_\xi^+)^*] \\
& \times [(1-\exp(-\theta/\varepsilon))\Psi(\sigma/\varepsilon) - (1-\exp(-\sigma/\varepsilon))\Psi(\theta/\varepsilon)] \\
& + 0.25\Psi(\theta/\varepsilon)\Psi(\sigma/\varepsilon)S[(\Phi_\xi^-)^*, (\Phi_\xi^+)^*][\Psi(\delta)\xi^*(\theta,\Delta\theta) \\
& + (1/\varepsilon)\int_{\Delta\theta}^{0}(1-\exp(-\sigma/\varepsilon))\xi^*(\theta,\sigma)d\sigma],
\end{aligned}
\tag{64}
$$

$$S(\alpha_1,\alpha_2) \triangleq \partial R(\alpha_1,\alpha_2)/\partial\alpha_1 - \partial R(\alpha_1,\alpha_2)/\partial\alpha_2, \tag{65}$$

$$(\Phi_\xi^-)^* \triangleq -(1/2\varepsilon)\int_{\Delta\theta}^{0}\Psi(\sigma/\varepsilon)\xi^*(\theta,\sigma)d\sigma + \beta, \tag{66}$$

$$(\Phi_\xi^+)^* \triangleq (1/2\varepsilon)\int_{\Delta\theta}^{0}\Psi(\sigma/\varepsilon)\xi^*(\theta,\sigma)d\sigma + \beta, \tag{67}$$

$$\xi^*(\theta,\sigma) = \begin{cases} \bar{v}^*(\theta+\sigma), & 0 \le \theta+\sigma \le \theta_0, \\ \bar{v}_0(\theta+\sigma), & \theta+\sigma > \theta_0. \end{cases} \tag{68}$$

The transversality conditions have the form

$$\varphi(0) = -\mathrm{sign}[\bar{z}_{im}^c(0)], \tag{69}$$

$$\lambda(\theta,\Delta\theta) = \varepsilon\Psi(\theta/\varepsilon)[\exp(-\delta) + 0.5R[(\Phi_\xi^-)^*, (\Phi_\xi^+)^*]\Psi(\delta)]\varphi(\theta), \tag{70}$$

$$\lambda(0,\sigma) = 0. \tag{71}$$

The proof of the proposition is presented in Appendix C.

3.4 Solution of Equations (62)–(71)

Solving Equation (62) with the initial condition (69) yields

$$\varphi(\theta) = -\mathrm{sign}[\bar{z}_{im}^c(0)]. \tag{72}$$

Substituting (72) into (63) and (70), one obtains

$$\partial\lambda(\theta,\sigma)/\partial\theta - \partial\lambda(\theta,\sigma)/\partial\sigma = -\mathrm{sign}[\bar{z}_{im}^c(0)]G(\theta,\sigma,\bar{v}^*(\cdot)), \tag{73}$$

$$\lambda(\theta,\Delta\theta) = -\mathrm{sign}[\bar{z}_{im}^c(0)]\varepsilon\Psi(\theta/\varepsilon)[\exp(-\delta) + 0.5R[(\Phi_\xi^-)^*, (\Phi_\xi^+)^*]\Psi(\delta)]. \tag{74}$$

Proposition 3.2. *For any given $\bar{v}^*(\theta)$, $\theta \in [0,\theta_0]$, satisfying the constraint (58), the function*

$$\lambda(\theta,\sigma) = -\mathrm{sign}[\bar{z}_{im}^c(0)]\varepsilon\Psi(\theta/\varepsilon)[\exp(-\sigma/\varepsilon) + 0.5R[(\Phi_\xi^-)^*, (\Phi_\xi^+)^*]\Psi(\sigma/\varepsilon)], \tag{75}$$

is a unique solution of the partial differential equation (73) with the boundary conditions (71) and (74) in the domain Ω (introduced in Proposition 3.1).

Proof. The proposition can be verified by direct substitution. □

Substituting (72) and (75) into (60) and (61), respectively, yields the following unique pair of candidate optimal controls in the NLDDG:

$$\bar{u}^*(\theta) = \mathrm{sign}[\bar{z}^c_{im}(0)], \quad \theta \in [\theta_0, 0], \tag{76}$$
$$\bar{v}^*(\theta) = (1-\nu)\mathrm{sign}[\bar{z}^c_{im}(0)] \quad \theta \in [\theta_0, 0], \tag{77}$$

indicating that these unique candidate optimal controls are independent of the initial function $\bar{v}_0(\theta)$, $\theta \in [\theta_0 + \Delta\theta, \theta_0)$ selected for the evader's control. However, the cost function (55) may depend on this initial function, leading to the following optimization problem: finding the initial function $\bar{v}_0(\theta)$ that maximizes the cost function (55) under the condition that the players apply their candidate optimal controls (76), (77). This problem will be formulated precisely and solved in the next section.

4 Optimal Initial Function for the Evader's Control

Substituting (77) into (68), one has

$$\xi^*(\theta,\sigma) = \begin{cases} (1-\nu)\mathrm{sign}[\bar{z}^c_{im}(0)], & 0 \le \theta+\sigma \le \theta_0, \\ \bar{v}_0(\theta+\sigma), & \theta+\sigma > \theta_0. \end{cases} \tag{78}$$

Two cases can be distinguished: (i) $\theta_0 \ge \Delta\theta$; (ii) $\theta_0 < \Delta\theta$. In this chapter only the first case is considered in detail, because of its practical importance.

4.1 Formulation of an Extremal Problem for the Initial Function

For $\theta_0 \ge \Delta\theta$, Equation (78) yields

$$\xi^*(\theta,\sigma) = \begin{cases} (1-\nu)\mathrm{sign}[\bar{z}^c_{im}(0)], & \theta \in [\theta_0, \theta_0 - \Delta\theta), \quad \sigma \in [0, \theta_0 - \theta], \\ \bar{v}_0(\theta+\sigma), & \theta \in [\theta_0, \theta_0 - \Delta\theta), \quad \sigma \in (\theta_0 - \theta, \Delta\theta], \end{cases} \tag{79}$$

$$\xi^*(\theta,\sigma) = (1-\nu)\mathrm{sign}[\bar{z}^c_{im}(0)], \qquad \theta \in [\theta_0 - \Delta\theta, 0], \quad \sigma \in [0, \Delta\theta]. \tag{80}$$

By substituting (76) and (79) into Equation (49) for $\bar{u}$ and $\bar{v}$, respectively, this equation becomes on the interval $\theta \in [\theta_0, \theta_0 - \Delta\theta)$

$$\begin{aligned} d\bar{z}^c_{im}(\theta)/d\theta =& \mu\Psi(\theta)\mathrm{sign}[\bar{z}^c_{im}(0)] - \varepsilon\Psi(\theta/\varepsilon)\exp(-\delta)\bar{v}_0(\theta+\Delta\theta) \\ &+ (\theta/\varepsilon)\Bigg\{\int_{\Delta\theta}^{\theta_0-\theta} \exp(-\sigma/\varepsilon)\bar{v}_0(\theta+\sigma)d\sigma \\ &\qquad - \varepsilon(1-\nu)\mathrm{sign}[\bar{z}^c_{im}(0)][1-\exp(-(\theta_0-\theta)/\varepsilon)]\Bigg\} \\ &+ \varepsilon[1-\exp(-\theta/\varepsilon)]Q(\Phi^-_\nu, \Phi^+_\nu) + E_\nu(\theta, \bar{v}_0(\cdot)), \end{aligned} \tag{81}$$

where

$$E_\nu(\theta, \bar{v}_0(\cdot)) \triangleq -(\varepsilon/2)\Psi(\theta/\varepsilon)R(\Phi_\nu^-, \Phi_\nu^+) \\ \times \Bigg[\Psi(\delta)\bar{v}_0(\theta + \Delta\theta) + (1/\varepsilon)\int_{\Delta\theta}^{\theta_0-\theta}(1 - \exp(-\sigma/\varepsilon))\bar{v}_0(\theta+\sigma)d\sigma \\ - (1-\nu)\mathrm{sign}(\bar{z}_{im}^c(0))\Psi((\theta_0 - \theta)/\varepsilon)\Bigg], \tag{82}$$

$$\Phi_\nu^- \triangleq -(1/2\varepsilon)\Bigg\{\int_{\Delta\theta}^{\theta_0-\theta}\Psi(\sigma/\varepsilon)\bar{v}_0(\theta+\sigma)d\sigma \\ + (1-\nu)\mathrm{sign}[\bar{z}_{im}^c(0)]\int_{\theta_0-\theta}^{0}\Psi(\sigma/\varepsilon)d\sigma\Bigg\} + \beta, \tag{83}$$

$$\Phi_\nu^+ \triangleq (1/2\varepsilon)\Bigg\{\int_{\Delta\theta}^{\theta_0-\theta}\Psi(\sigma/\varepsilon)\bar{v}_0(\theta+\sigma)d\sigma \\ + (1-\nu)\mathrm{sign}[\bar{z}_{im}^c(0)]\int_{\theta_0-\theta}^{0}\Psi(\sigma/\varepsilon)d\sigma\Bigg\} + \beta, \tag{84}$$

and $R(\cdot,\cdot)$ is defined by (52).

Similarly, by using (76) and (80), Equation (49) becomes on the interval $\theta \in [\theta_0 - \Delta\theta, 0]$

$$d\bar{z}_{im}^c(\theta)/d\theta = \Gamma_{im}(\theta,\nu)\mathrm{sign}[\bar{z}_{im}^c(0)], \tag{85}$$

where

$$\Gamma_{im}(\theta,\nu) \triangleq \mu\Psi(\theta) - (1-\nu)[\varepsilon\Psi(\theta/\varepsilon)\exp(-\delta) - \theta(1-\exp(-\delta))] \\ + \varepsilon[1 - \exp(-\theta/\varepsilon)]Q_\nu, \tag{86}$$

$$Q_\nu \triangleq Q(\Phi^-, \Phi^+)|_{\bar{v}\equiv 1-\nu}, \tag{87}$$

and $Q(\cdot,\cdot)$ is defined by (51).

Remark 4.1. It can be verified directly that

$$\Gamma_{im}(\theta,\nu) \to \Gamma(\theta) = \mu\Psi(\theta) - \varepsilon\Psi(\theta/\varepsilon) \quad \text{as} \quad \nu \to +0. \tag{88}$$

From Equations (85)–(87), one can see that by using the candidate optimal controls of the players the dynamics of the NLDDG is independent of $\bar{v}_0(\cdot)$ on the interval $\theta \in [\theta_0 - \Delta\theta, 0]$. However, due to Equations (81)–(84), this dynamics depends on $\bar{v}_0(\cdot)$ on the interval $\theta \in [\theta_0, \theta_0 - \Delta\theta)$. Thus, in order to maximize $|\bar{z}_{im}^c(0)|$ by choosing $\bar{v}_0(\cdot)$, it is necessary and sufficient for the evader to maximize $|\bar{z}_{im}^c(\theta_0 - \Delta\theta)|$. Therefore, one can consider the NLDDG only on the interval $\theta \in [\theta_0, \theta_0 - \Delta\theta]$ via Equation (81), where $\bar{z}^c_{im}$ is the state variable and $\bar{v}_0$ is the control.

The control function $\bar{v}_0(\cdot)$ appears in Equation (81) with the argument $\theta+\Delta\theta$ and with the argument $\theta+\sigma$, $\sigma \in [\theta_0-\theta, \Delta\theta]$. When θ varies from θ_0 to $\theta_0-\Delta\theta$, the argument $\theta + \Delta\theta$ varies from $\theta_0 + \Delta\theta$ to θ_0. Thus, $\theta + \Delta\theta$ must be taken as the current argument of $\bar{v}_0(\cdot)$. Moreover, one has that $\theta + \sigma < \theta + \Delta\theta$ for all $\sigma \in [\theta_0 - \theta, \Delta\theta)$, i.e., the argument $\theta + \sigma$ of the control $\bar{v}_0(\cdot)$ appearing in the integral is smaller than the current argument $\theta+\Delta\theta$ of this function. Since both arguments of $\bar{v}_0(\cdot)$ in Equation (81) refer to the time-to-go, one can directly conclude that $\theta + \sigma$ represents an advance with respect to $\theta + \Delta\theta$. In order to transform this advance to a delay, one has to vary θ in the opposite direction, i.e., from $\theta_0 - \Delta\theta$ to θ_0, and instead of maximizing $|\bar{z}^c_{im}(\theta_0 - \Delta\theta)|$, to solve an equivalent problem of minimizing, for a given $\bar{z}^c_{im}(\theta_0 - \Delta\theta)$, $\mathrm{sign}[\bar{z}^c_{im}(\theta_0 - \Delta\theta)]\bar{z}^c_{im}(\theta_0)$. Thus, one obtains an optimal control problem with a dynamics described by Equation (81), $\theta \in (\theta_0 - \Delta\theta, \theta_0]$, the control constraint

$$|\bar{v}_0(\theta + \Delta\theta)| \leq 1 - \nu, \tag{89}$$

and the cost function

$$\bar{J}_{in} \triangleq \mathrm{sign}[\bar{z}^c_{im}(\theta_0 - \Delta\theta)]\bar{z}^c_{im}(\theta_0) \to \min_{\bar{v}_0(\theta+\Delta\theta)} . \tag{90}$$

We shall call this problem the evader's optimal control problem (EOCP). Since the upper bound of the integral in Equation (81) depends on θ, the delay in the EOCP dynamics is not constant.

4.2 Necessary Conditions of Optimality

Proposition 4.1. *The candidate optimal control in the EOCP satisfies the equation for $\theta \in (\theta_0 - \Delta\theta, \theta_0]$*

$$\begin{aligned}&\bar{v}^*_0(\theta + \Delta\theta)\\ &\quad = (1-\nu)\mathrm{sign}\{\chi(\theta, \Delta\theta) - \varepsilon\kappa(\theta)\Psi(\theta/\varepsilon)[\exp(-\delta) + 0.5R((\Phi^-_\nu)^*, (\Phi^+_\nu)^*)\Psi(\delta)]\},\end{aligned} \tag{91}$$

where $(\Phi^-_\nu)^*$ and $(\Phi^+_\nu)^*$ are obtained from Φ^-_ν and Φ^+_ν, respectively, by replacing $\bar{v}_0(\cdot)$ with $\bar{v}^*_0(\cdot)$.

The functions $\kappa(\theta)$ and $\chi(\theta, \sigma)$ satisfy the following equations:

$$d\kappa(\theta)/d\theta = 0, \qquad \theta \in (\theta_0 - \Delta\theta, \theta_0], \tag{92}$$

$$\begin{aligned}&\partial\chi(\theta, \sigma)/\partial\theta - \partial\chi(\theta, \sigma)/\partial\sigma = \kappa(\theta)G_1(\theta, \sigma, \bar{v}^*_0(\cdot)),\\ &\qquad\qquad \theta \in (\theta_0 - \Delta\theta, \theta_0], \quad \sigma \in (\theta_0 - \theta, \Delta\theta],\end{aligned} \tag{93}$$

where

$$
\begin{aligned}
&G_1(\theta,\sigma,\bar{v}_0^*(\cdot))\\
&\quad\triangleq (\theta/\varepsilon)\exp(-\sigma/\varepsilon) + 0.5R((\Phi_\nu^-)^*,(\Phi_\nu^+)^*)\\
&\qquad\times[(1-\exp(-\theta/\varepsilon))\Psi(\sigma/\varepsilon) - (1-\exp(-\sigma/\varepsilon))\Psi(\theta/\varepsilon)]\\
&\qquad+0.25\Psi(\theta/\varepsilon)\Psi(\sigma/\varepsilon)S((\Phi_\nu^-)^*,(\Phi_\nu^+)^*)\\
&\qquad\times\Bigg[\Psi(\delta)\bar{v}_0^*(\theta+\Delta\theta) + (1/\varepsilon)\int_{\Delta\theta}^{\theta_0-\theta}(1-\exp(-\sigma/\varepsilon))\bar{v}_0^*(\theta+\sigma)d\sigma\\
&\qquad\quad-(1-\nu)\mathrm{sign}(\bar{z}_{im}^c(0))\Psi((\theta_0-\theta)/\varepsilon)\Bigg]. \qquad (94)
\end{aligned}
$$

The transversality conditions have the form

$$\kappa(\theta_0) = -\mathrm{sign}[\bar{z}_{im}^c(\theta_0-\Delta\theta)], \qquad (95)$$

$$\chi(\theta_0,\sigma)=0, \qquad \sigma\in(0,\Delta\theta]. \qquad (96)$$

Proof. The proposition is proved similarly to Proposition 3.1 with some technical differences arising due to the time-varying character of the delay in the EOCP dynamics. These technical differences can be found in [31] for a much more general case of variable delay. □

4.3 Solution of Equations (92)–(96)

First, note that due to Equation (85), Remark 4.1 and results of [28], one has for a candidate optimal trajectory of the NLDDG

$$\mathrm{sign}[\bar{z}_{im}^c(0)] = \mathrm{sign}[\bar{z}_{im}^c(\theta_0-\Delta\theta)]. \qquad (97)$$

Solving Equation (92) with the condition (95) and using (97) yields

$$\kappa(\theta) = -\mathrm{sign}[\bar{z}_{im}^c(0)]. \qquad (98)$$

Substituting (98) into Equation (93) leads to

$$
\begin{aligned}
\partial\chi(\theta,\sigma)/\partial\theta - \partial\chi(\theta,\sigma)/\partial\sigma &= -\mathrm{sign}[\bar{z}_{im}^c(0)]G_1(\theta,\sigma,\bar{v}_0^*(\cdot)),\\
&\theta\in(\theta_0-\Delta\theta,\theta_0],\quad \sigma\in(\theta_0-\theta,\Delta\theta]. \quad (99)
\end{aligned}
$$

Proposition 4.2. *For any $\bar{v}_0^*(\theta+\Delta\theta)$, $\theta\in(\theta_0-\Delta\theta,\theta_0]$, satisfying the constraint (89), the unique solution of the problem (99), (96) is given by the function*

$$\begin{aligned}\chi(\theta,\sigma) = \operatorname{sign}[\bar{z}_{im}^{c}(0)][(\varepsilon-\theta)\exp(-\sigma/\varepsilon) - (\varepsilon-\theta_0)\exp(-(\theta+\sigma-\theta_0)/\varepsilon)\\ - 0.5\varepsilon R((\Phi_{\nu}^{-})^{*},(\Phi_{\nu}^{+})^{*})\Psi(\theta/\varepsilon)\Psi(\sigma/\varepsilon)\\ + 0.5\varepsilon R_{\nu 0}^{*}\Psi(\theta_0/\varepsilon)\Psi((\theta+\sigma-\theta_0)/\varepsilon)],\end{aligned} \tag{100}$$

where

$$R_{\nu 0}^{*} \stackrel{\triangle}{=} R((\Phi_{\nu}^{-})^{*},(\Phi_{\nu}^{+})^{*})|_{\theta=\theta_0}. \tag{101}$$

Proof. The proposition is proved by direct substitution of (100) into the problem (99), (96). □

Substituting (98) and (100) into (91) yields after some rearrangement the equation for the candidate optimal control in the EOCP

$$\begin{aligned}\bar{v}_0^{*}(\theta+\Delta\theta) &= (1-\nu)\operatorname{sign}[\bar{z}_{im}^{c}(0)]\\ &\quad\times \operatorname{sign}\{\varepsilon\Psi(\theta_0/\varepsilon)[\exp(-(\theta-\theta_0)/\varepsilon)\exp(-\delta)\\ &\qquad + 0.5R_{\nu 0}^{*}\Psi((\theta-\theta_0+\Delta\theta)/\varepsilon)]\}\\ &= (1-\nu)\operatorname{sign}[\bar{z}_{im}^{c}(0)]\operatorname{sign}[C(\theta,\bar{v}_0^{*}(\cdot))],\end{aligned} \tag{102}$$

where

$$C(\theta,\bar{v}_0^{*}(\cdot)) \stackrel{\triangle}{=} \exp(-(\theta-\theta_0)/\varepsilon)\exp(-\delta) + 0.5R_{\nu 0}^{*}\Psi((\theta-\theta_0+\Delta\theta)/\varepsilon). \tag{103}$$

From Equation (102), one can see that the candidate optimal control $\bar{v}_0^{*}(\theta+\Delta\theta)$ in the EOCP has a bang-bang structure. Moreover, the switch function $C(\theta,\bar{v}_0^{*}(\cdot))$ depends on $\bar{v}_0^{*}(\cdot)$, i.e., Equation (102) is a functional equation with respect to $\bar{v}_0^{*}(\theta+\Delta\theta)$. In the next subsection, we shall solve this equation.

4.4 Solution of Equation (102)

Proposition 4.3. *Equation (102) has the unique solution*

$$\bar{v}_0^{*}(\theta+\Delta\theta) = (1-\nu)\operatorname{sign}[\bar{z}_{im}^{c}(0)]\left\{\begin{array}{rl} 1, & \theta\in(\theta_0-\Delta\theta,\theta_{sw}),\\ -1, & \theta\in(\theta_{sw},\theta_0]\end{array}\right\}, \tag{104}$$

where the switch point θ_{sw} is

$$\theta_{sw} = \theta_0 - \Delta\theta + \sigma_{sw}, \tag{105}$$

and σ_{sw} is a unique positive solution of the equation

$$\exp(-\sigma_{sw}/\varepsilon) + 0.5R_{\nu 0}^{*}\Psi(\sigma_{sw}/\varepsilon) = 0. \tag{106}$$

The proof of the proposition is presented in Appendix D, yielding the candidate initial optimal control function of the evader in the NLDDG.

Remark 4.2. Note that σ_{sw} is independent of θ_0 as well as of $\operatorname{sign}[\bar{z}_{im}^{c}(0)]$. Moreover, due to (104), (105), σ_{sw} satisfies the inequality $\sigma_{sw} < \Delta\theta$.

In Figure 5, the values of σ_{sw} normalized with respect to ε, i.e., $\bar{\sigma}_{sw} \stackrel{\triangle}{=} \sigma_{sw}/\varepsilon$, are depicted as a function of $\delta = \Delta\theta/\varepsilon$.

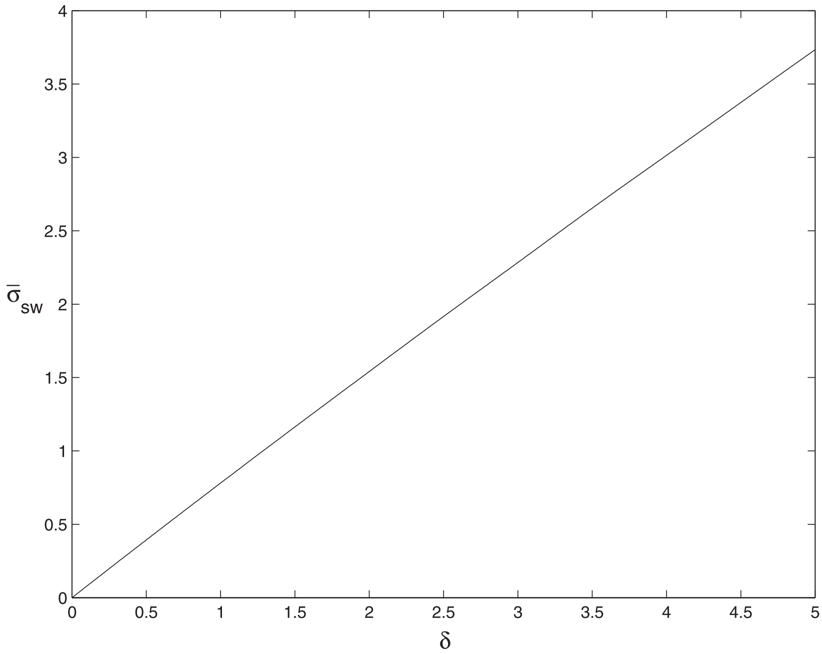


Figure 5: $\bar{\sigma}_{sw}$ as a function of δ ($\nu = 0.01$).

5 NLDDG space decomposition

Based on the candidate optimal controls of the players and on the candidate optimal initial function for the evader's control, the game space decomposition can be constructed.

Similarly to the previous section, it is assumed that $\theta_0 \geq \Delta\theta$. In this case, a candidate optimal trajectory of the NLDDG is described by Equation (81) (with $\bar{v}_0(\cdot) = \bar{v}_0^*(\cdot)$) on the interval $\theta \in [\theta_0, \theta_0 - \Delta\theta)$ and by Equation (85) on the interval $\theta \in [\theta_0 - \Delta\theta, 0]$.

The value of $\bar{z}_{im}^c(\theta_0)$ can be expressed by $\bar{z}_{im}^c(0)$. Integrating Equation (85) from $\theta = 0$ to $\theta = \theta_0 - \Delta\theta$ yields

$$\bar{z}_{im}^c(\theta_0 - \Delta\theta) = \bar{z}_{im}^c(0) + \int_0^{\theta_0 - \Delta\theta} \Gamma_{im}(\theta, \nu) d\theta. \tag{107}$$

Similarly, integrating Equation (81) from $\theta = \theta_0 - \Delta\theta$ to $\theta = \theta_0$, one has

$$\bar{z}_{im}^c(\theta_0) = \bar{z}_{im}^c(\theta_0 - \Delta) + \int_{\theta_0 - \Delta\theta}^{\theta_0} L(\theta, \bar{v}_0^*(\cdot)) d\theta, \tag{108}$$

where $L(\theta, \bar{v}_0(\cdot))$ is the right-hand part of Equation (81).

Substituting (107) into (108) yields the expression of $\bar{z}^c_{im}(\theta_0)$ for a given value of $\bar{z}^c_{im}(0)$:

$$\bar{z}^c_{im}(\theta_0) = \bar{z}^c_{im}(0) + \int_0^{\theta_0-\Delta\theta} \Gamma_{im}(\theta,\nu)d\theta + \int_{\theta_0-\Delta\theta}^{\theta_0} L(\theta,\bar{v}^*_0(\cdot))d\theta. \qquad (109)$$

When θ_0 varies on the interval $[\Delta\theta, +\infty)$, Equation (109) describes a curve in the plane $(\theta, \bar{z}^c_{im})$. This curve, denoted in the sequel as the Locus, is the set of all initial points $(\theta_0, \bar{z}^c_{im}(\theta_0))$ of candidate optimal trajectories in the NLDDG which arrive at the same final point $(0, \bar{z}^c_{im}(0))$. Thus, depending on $\bar{z}^c_{im}(0)$, the differential equation describing the family of all Loci can be obtained as follows.

Let $(\theta, \bar{z}_L(\theta))$, $\theta \in [\Delta\theta, +\infty)$ be an arbitrary point of the Locus. Denote

$$\eta(\theta,s) \triangleq (1-\nu)\left\{ \begin{array}{ll} 1, & s \in (\theta, s_{sw}), \\ -1, & s \in (s_{sw}, \theta+\Delta\theta] \end{array} \right\}, \quad \theta \in [\Delta\theta, +\infty), \qquad (110)$$

where the switch point s_{sw} has the form

$$s_{sw} = s_{sw}(\theta) = \theta + \sigma_{sw}. \qquad (111)$$

Proposition 5.1. *The family of Loci is described by the following differential equation:*

$$d\bar{z}_L(\theta)/d\theta = \mathrm{sign}[\bar{z}^c_{im}(0)]\Gamma_L(\theta), \qquad \theta \in [\Delta\theta, +\infty), \qquad (112)$$

where

$$\begin{aligned}\Gamma_L(\theta) \triangleq\ & \mu\Psi(\theta) - \varepsilon\Psi(\theta/\varepsilon)\exp(-\delta)\eta(\theta,\theta+\Delta\theta) \\ & + (\theta/\varepsilon)\int_{\Delta\theta}^0 \exp(-\sigma/\varepsilon)\eta(\theta,\theta+\sigma)d\sigma \\ & + \varepsilon[1-\exp(-\theta/\varepsilon)]Q(\Phi^-_\eta, \Phi^+_\eta) + E_\eta(\theta,\eta(\cdot)), \qquad (113)\end{aligned}$$

$$\Phi^-_\eta \triangleq -(1/2\varepsilon)\int_{\Delta\theta}^0 \Psi(\sigma/\varepsilon)\eta(\theta,\theta+\sigma)d\sigma + \beta, \qquad (114)$$

$$\Phi^+_\eta \triangleq (1/2\varepsilon)\int_{\Delta\theta}^0 \Psi(\sigma/\varepsilon)\eta(\theta,\theta+\sigma)d\sigma + \beta, \qquad (115)$$

and $E_\eta(\theta,\eta(\cdot))$ is obtained from Equation (50) replacing there $\bar{v}(\theta+\sigma)$ by $\eta(\theta,\theta+\sigma)$, $\sigma \in [0,\Delta\theta]$, Φ^- by Φ^-_η and Φ^+ by Φ^+_η.

Proof. The proposition is proved by differentiating (109) with respect to θ_0, and then replacing θ_0 by θ. □

In Figure 6, one of the Loci is depicted. In this figure, the continuation of the Locus for $\theta \in [0, \Delta\theta)$ has been obtained by numerical calculation of the candidate optimal initial function $\bar{v}^*_0(\theta)$ in the case when the initial value of θ is

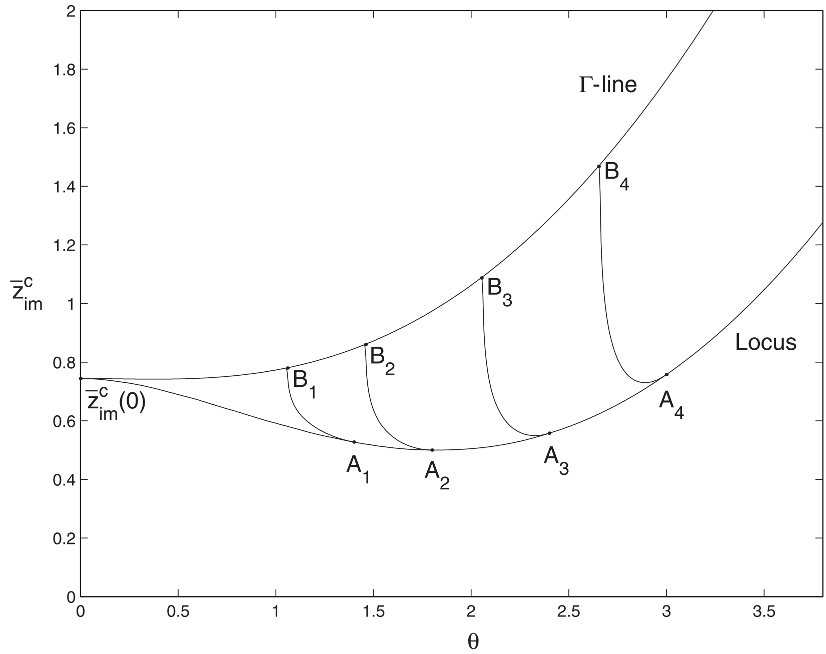


Figure 6: $(\text{Locus})_+$ and $(\Gamma\text{-line})_+$, $(\mu = 1.5, \varepsilon = 1.0,\ \delta = 1.5,\ \nu = 0.01)$.

less than $\Delta\theta$. The line, denoted as the Γ-line, is the trajectory of Equation (85) for $\theta \in [0, +\infty)$ beginning at the point $(0, \bar{z}^c_{im}(0))$. The points A_i $(i = 1, \ldots, 4)$ are various positions of the initial point $(\theta_0, \bar{z}^c_{im}(\theta_0))$ on the Locus. Each of the points B_i $(i = 1, \ldots, 4)$ is the one where the candidate optimal trajectory starting at A_i arrives to the Γ-line. From B_i to the end point $(0, \bar{z}^c_{im}(0))$, the candidate optimal trajectory coincides with the Γ-line.

Numerical calculations show that for $\mu > 1, \mu\varepsilon \geq 1$ the function $\Gamma_L(\theta)$ has a single root $\theta = \theta_L$ on the interval $\theta \in [\Delta\theta, +\infty)$. Thus, there exist two symmetric Loci with respect to the θ-axis, which reach this axis tangentially at $\theta = \theta_L$. These lines, denoted in Figure 7 as $(\text{Locus})^*_+$ and $(\text{Locus})^*_-$, decompose the $(\theta, \bar{z}^c_{im})$-plane into two regions. In this figure, the continuation of these Loci for $\theta \in [0, \Delta\theta)$ has been obtained similarly to the method used in Figure 6. The lines $(\Gamma\text{-line})^*_+$ and $(\Gamma\text{-line})^*_-$ correspond to $(\text{Locus})^*_+$ and $(\text{Locus})^*_-$, respectively.

In the regular region $(\mathcal{D}^c_1)_{im}$, the candidate optimal feedback pursuer's strategy is

$$\bar{u}^*[\theta, \bar{z}^c_{im}(\theta)] = \text{sign}[\bar{z}^c_{im}(\theta)], \qquad \bar{z}^c_{im}(\theta) \neq 0, \tag{116}$$

while the candidate optimal feedback evader's strategy is

$$\bar{v}^*[\theta, \bar{z}^c_{im}(\theta)] = (1-\nu)\text{sign}[\bar{z}^c_{im}(\theta)], \qquad \bar{z}^c_{im}(\theta) \neq 0, \tag{117}$$

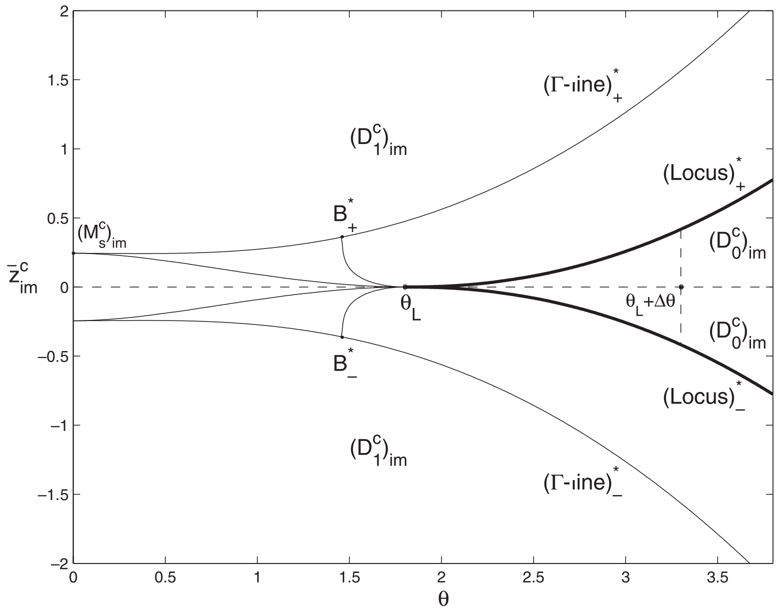


Figure 7: NLDDG space decomposition for $\mu > 1$, $\mu\varepsilon \geq 1$ ($\mu = 1.5$, $\varepsilon = 1$, $\delta = 1.5$, $\nu = 0.01$).

and the evader's initial function is

$$\bar{v}^*_0[s, \bar{z}^c_{im}(\theta_0)] = \text{sign}[\bar{z}^c_{im}(\theta_0)]\eta(\theta_0, s), \quad s \in [\theta_0 + \Delta\theta, \theta_0), \tag{118}$$

where $\eta(\cdot, \cdot)$ is given by (110).

The value of the game is a function of the initial position of the game $(\theta_0, \bar{z}^c_{im}(\theta_0))$.

Note also that the regions between the (Γ-line)$^*_+$ and the (Locus)$^*_+$ (including the (Locus)$^*_+$), as well as between the (Γ-line)$^*_-$ and the (Locus)$^*_-$ (including the (Locus)$^*_-$), belong to $(\mathcal{D}^c_1)_{im}$.

In the singular region $(\mathcal{D}^c_0)_{im}$, the pursuer's candidate optimal strategy is arbitrary, satisfying (21). The evader's candidate optimal strategy is the following: on the interval $\theta \in [\theta_0, \theta_L + \Delta\theta)$, this strategy is arbitrary satisfying (58) and the condition

$$\Bigg|\mu\Psi(\theta)\text{sign}(\bar{z}^c_{im}) - \varepsilon\Psi(\theta/\varepsilon)\exp(-\delta)\bar{v}(\theta + \Delta\theta) + (\theta/\varepsilon)\int_{\Delta\theta}^{0} \exp(-\sigma/\varepsilon)\bar{v}(\theta + \sigma)d\sigma$$
$$+ \varepsilon[1 - \exp(-\theta/\varepsilon)]Q(\Phi^-, \Phi^+) + E(\theta, \bar{v}(\cdot))\Bigg|_{\bar{v}(\cdot)=\bar{v}^*(\cdot)}$$
$$\geq \Gamma_L(\theta), \quad (\theta, \bar{z}^c_{im}) \in \partial(\mathcal{D}^c_0)_{im}, \tag{119}$$

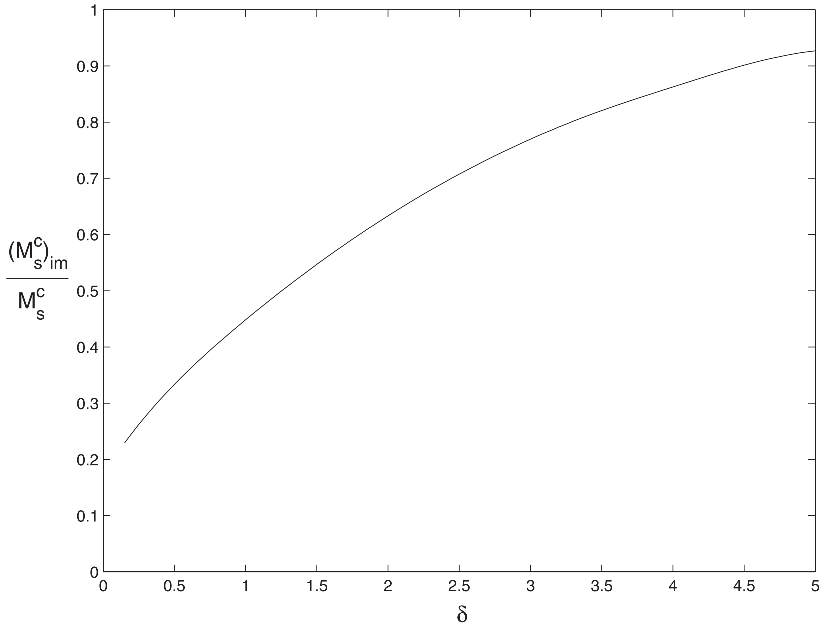


Figure 8: Guaranteed miss distance ratio ($\mu = 1.5$, $\varepsilon = 1.0$, $\delta = 1.5$).

where $\Gamma_L(\theta)$ is given by (113) and $\partial(\mathcal{D}_0^c)_{im}$ is the boundary of the singular region.

Using such a strategy, the evader can keep the game's trajectory in the closure of the singular region, because, due to (49) and (112), the slope of the game trajectory at points of $\partial(\mathcal{D}_0^c)_{im}$ (the Loci depicted by the bold line in Figure 7) is not less than the slope of these Loci.

At $\theta = \theta_L + \Delta\theta$ the evader must choose (with the probability 0.5) either the control $\bar{v}^*(\theta) = \eta(\theta_L, \theta)$ or $\bar{v}^*(\theta) = -\eta(\theta_L, \theta)$ on the interval $\theta \in [\theta_L + \Delta\theta, \theta_L]$, where $\eta(\cdot, \cdot)$ is given by (110). Recall that $\eta(\theta_L, \theta)$ has a bang-bang form with the switch at $\theta = \theta_L + \sigma_{sw}$, where σ_{sw} is a unique positive solution of Equation (106) and $\sigma_{sw} < \Delta\theta$.

Starting in the singular region, a candidate optimal trajectory of the NLDDG arrives to either the (Γ-line)$_+^*$ or the (Γ-line)$_-^*$ at the points B_+^* and B_-^*, respectively, and then it moves along the corresponding Γ-line until the game ends. The value of the game in the singular region is constant, denoted as $(M_s^c)_{im}$.

The guaranteed normalized miss distance $(M_s^c)_{im}$ obtained by the candidate optimal strategies for initial conditions in $(\mathcal{D}_0^c)_{im}$ is much smaller than M_s^c (the one obtained in the SDIG with the same delay), as can be seen in Figure 8. The improvement due to the new approach is particularly rewarding for small values of the normalized delay δ.

Remark 5.1. Since the candidate optimal strategies of the players generate trajectories that fill the entire $(\theta, \bar{z}^c_{im})$ game space, they are indeed optimal.

6 Conclusions

In this chapter important additional progress has been made towards the synthesis of a new guidance law that guarantees satisfactory interceptions of highly maneuvering targets in a noise corrupted environment. The engagement is formulated as a deterministic linear pursuit-evasion game with bounded controls. The indispensable estimator of the pursuer's guidance system is modeled by a constant delay in observing the evader's lateral acceleration, while the other state variables are assumed to be perfectly measured. By taking into account the measurement time history of these variables during the delay, the *uncertainty set* of the pursuer created by the information delay is minimized. This approach allows transformation of the original problem into a nonlinear differential game with delayed evader control. The solution of this new game is obtained in pure strategies for the pursuer and in mixed strategies for the evader. The guidance law (the realization of the optimal pursuer strategy) provides a substantial improvement in the guaranteed homing performance compared to earlier results. The improvement is particularly significant for small normalized estimation delays.

Appendix A: Proof of Lemma 3.1

The lemma will be proved for the problem (22), (26), (27), (33) with the minimization of the cost function. The case of the maximization is proved similarly. The proof starts with the case when the inequality (37) is satisfied.

Based on the results of [16], the problem has a solution. In order to obtain the optimal control and to show it uniqueness, first, the constraint (33) is rewritten in an equivalent form by introducing an auxiliary state variable $\bar{x}_5(\sigma), \sigma \in [\theta + \Delta\theta, \theta]$ satisfying the following boundary-value problem:

$$d\bar{x}_5(\sigma)/d\sigma = -\Psi((\sigma - \theta)/\varepsilon)\bar{v}(\sigma), \tag{A1}$$

$$\bar{x}_5(\sigma)|_{\sigma=\theta+\Delta\theta} = 0, \qquad \bar{x}_5(\sigma)|_{\sigma=\theta} = \bar{q}(\theta). \tag{A2}$$

It is clear that the boundary-value problem (A1), (A2) is equivalent to the constraint (33). Thus, in the sequel the optimal control problem (22), (26), (27), (A1), (A2), called the auxiliary optimal control problem (AOCP), is considered instead of the equivalent one (22), (26), (27), (33).

The variational Hamiltonian for the AOCP has the form

$$\bar{H} = \bar{\psi}_4(\sigma)[\bar{x}_4(\sigma) - \bar{v}(\sigma)]/\varepsilon - \bar{\psi}_5(\sigma)\Psi((\sigma - \theta)/\varepsilon)\bar{v}(\sigma), \tag{A3}$$

where the costate variables $\bar{\psi}_4(\sigma)$ and $\bar{\psi}_5(\sigma)$ satisfy the equations

$$d\bar{\psi}_4(\sigma)/d\sigma = -\partial\bar{H}/\partial\bar{x}_4 = -\bar{\psi}_4(\sigma)/\varepsilon, \tag{A4}$$

$$d\bar{\psi}_5(\sigma)/d\sigma = -\partial\bar{H}/\partial\bar{x}_5 = 0. \tag{A5}$$

The transversality condition is

$$\bar{\psi}_4(\theta) = 1. \tag{A6}$$

Due to Pontryagin's maximum principle, the candidate optimal control in the AOCP becomes

$$\bar{v}(\sigma) = \operatorname{sign}[\bar{S}(\sigma)], \tag{A7}$$

where

$$\bar{S}(\sigma) \stackrel{\triangle}{=} -\bar{\psi}_4(\sigma)/\varepsilon - \bar{\psi}_5(\sigma)\Psi((\sigma-\theta)/\varepsilon). \tag{A8}$$

Solving Equation (A4) with condition (A6) yields

$$\bar{\psi}_4(\sigma) = \exp((\theta-\sigma)/\varepsilon), \qquad \theta \le \sigma \le \theta + \Delta\theta. \tag{A9}$$

Similarly, solving Equation (A5) leads to

$$\bar{\psi}_5(\sigma) = c, \qquad \theta \le \sigma \le \theta + \Delta\theta, \tag{A10}$$

where c is an arbitrary value independent of σ.

Substituting (A9) and (A10) into (A8), one has the switch function $\bar{S}(\sigma)$ as follows:

$$\bar{S}(\sigma) = -\exp((\theta-\sigma)/\varepsilon)/\varepsilon - c\Psi((\sigma-\theta)/\varepsilon), \quad \theta \le \sigma \le \theta + \Delta\theta, \tag{A11}$$

yielding

$$\bar{S}(\sigma)|_{\sigma=\theta} = -1. \tag{A12}$$

Based on Equation (A12), two different cases can be distinguished: 1) $\bar{S}(\sigma) < 0 \;\; \forall\sigma \in [\theta, \theta+\Delta\theta)$; 2) there exists $\sigma^- \in (\theta, \theta+\Delta\theta)$, such that $\bar{S}(\sigma^-) = 0$.

Case 1. In this case, one has a single candidate optimal control, namely, $\bar{v}(\sigma) = -1$. Due to the existence of the AOCP solution, this control is optimal. Hence, Equations (A1), (A2) and (35) yield $\bar{x}_5(\sigma)|_{\sigma=\theta} = \bar{q}^-$, which is valid iff $\bar{q}(\theta) = \bar{q}^-$. Thus, the first case occurs iff the left-hand inequality in (37) becomes an equality. In this case, one can directly obtain from Equations (42)–(45) that $\sigma^- = \theta + \Delta\theta$, which proves the second statement of the lemma (when the inequality (37) is satisfied).

Case 2. Substituting $\sigma = \sigma^-$ into (A11) and solving the equation with respect to c yields

$$c = -\exp((\theta - \sigma^-)/\varepsilon)/[\varepsilon\Psi((\sigma^- - \theta)/\varepsilon)] < 0. \qquad (A13)$$

Differentiating (A11) with respect to σ and using (A13), one obtains

$$\begin{aligned} d\bar{S}(\sigma)/d\sigma &= \exp((\theta - \sigma)/\varepsilon)/\varepsilon^2 - (c/\varepsilon)[1 - \exp((\theta - \sigma)/\varepsilon)] \\ &> 0, \qquad \sigma \in [\theta, \theta + \Delta\theta]. \end{aligned} \qquad (A14)$$

The inequality (A14) implies that $\bar{S}(\sigma)$ monotonically increases on the interval $\sigma \in [\theta, \theta + \Delta\theta]$. Hence, due to (A12), one has

$$\bar{S}(\sigma) < 0, \quad \sigma \in [\theta, \sigma^-); \qquad \bar{S}(\sigma) > 0, \quad \sigma \in (\sigma^-, \theta + \Delta\theta], \qquad (A15)$$

which yields by using (A7) the expression for the candidate optimal control coinciding with the right-hand part of Equation (38). By substituting this expression into Equation (A1) and solving this equation with the boundary conditions (A2), one obtains the following algebraic equation with respect to σ^-:

$$\exp[-(\sigma^- - \theta)/\varepsilon] - 0.5(\sigma^- - \theta)^2/\varepsilon^2 + (\sigma^- - \theta)/\varepsilon = \alpha^-, \qquad (A16)$$

where α^- is given by Equations (43), (44).

Equation (A16) directly yields expression (42) for the switch point σ^-. By the implicit function theorem one can show the existence and the smoothness of the function $B(\alpha)$ claimed in the lemma. Thus, the necessary optimality conditions yield a single candidate optimal control of the AOCP given by Equation (38). Since the AOCP has a solution, this control is optimal, which proves the main statement of the lemma. The other statement of the lemma (when the inequality (37) is not satisfied) is obvious.

Appendix B: Proof of Corollary 3.1

The proof of the corollary concentrates on the case where $\bar{v} = 1$ almost everywhere on an interval $[\theta + \Delta\theta, \theta]$. The case $\bar{v} = -1$ is analyzed similarly.

Using (33), (35) and (36), one has in the first case

$$q(\theta) = -\varepsilon[\Psi(\delta) - \delta^2/2]. \qquad (B1)$$

Substituting (B1) into (43) and using (44), one obtains

$$\alpha^- = 1, \qquad \alpha^+ = \alpha_0. \qquad (B2)$$

Using (45) and (B2) yields directly

$$B(\alpha^-) = 0, \qquad B(\alpha^+) = \delta. \qquad (B3)$$

Substituting (28) and (B3) into (46), one obtains

$$[\bar{x}_{4im}(\theta)]_c = \bar{x}_4(\theta + \Delta\theta)\exp(-\delta) + 1 - \exp(-\delta), \qquad (B4)$$

which coincides with the solution $\bar{x}_4(\theta)$ of Equation (27), the nondimensional form of Equation (9), with $\bar{v}(\sigma) = 1$ for almost all $\sigma \in [\theta + \Delta\theta, \theta]$. Thus, the corollary is proved.

Appendix C: Proof of Proposition 3.1

The proposition is proved by adopting the idea proposed in [36] for obtaining the optimality conditions in a linear-quadratic optimal control problem with delays in state and control variables. The proof consists of four stages. At the first stage, Equation (49) is transformed into a set of two equations, which no longer contain the delay. At the second stage, applying a discretization, the set of two equations is approximated by a set of $K+1$ ordinary differential equations of the first order ($K+1$ is the number of the collocation points) describing a dynamics of an auxiliary game. At the third stage, necessary optimality conditions for this auxiliary game are derived. At the fourth stage, a transformation of these conditions is carried out, and then the limit (as $K \to +\infty$) of these conditions is calculated.

C1. Transformation of Equation (49). Consider the following function of two variables $\xi(\theta, \sigma)$ in the domain Ω:

$$\xi(\theta,\sigma) \triangleq \begin{cases} \bar{v}(\theta+\sigma), & 0 \le \theta+\sigma \le \theta_0, \\ \bar{v}_0(\theta+\sigma), & \theta+\sigma > \theta_0. \end{cases} \qquad (C1)$$

Under the reasonable assumption that $\bar{v}(s)$, $s \in [0, \theta_0]$ is piecewise differentiable, $\xi(\theta, \sigma)$ satisfies the equation

$$\partial\xi(\theta,\sigma)/\partial\theta = \partial\xi(\theta,\sigma)/\partial\sigma \qquad (C2)$$

almost everywhere in Ω, and the condition

$$\xi(\theta, 0) = \bar{v}(\theta), \qquad \theta \in [0, \theta_0]. \qquad (C3)$$

Substituting (C1) into (49) yields the equation

$$\begin{aligned} d\bar{z}_{imc}(\theta)/d\theta =& \mu\Psi(\theta)\bar{u}(\theta) - \varepsilon\Psi(\theta/\varepsilon)\exp(-\delta)\xi(\theta,\Delta\theta) \\ &+ (\theta/\varepsilon)\int_{\Delta\theta}^{0} \exp(-\sigma/\varepsilon)\xi(\theta,\sigma)d\sigma \\ &+ \varepsilon[1-\exp(-\theta/\varepsilon)]Q(\Phi_\xi^-, \Phi_\xi^+) + E_\xi(\theta, \xi(\theta,\cdot)), \qquad (C4) \end{aligned}$$

where Φ_ξ^- and Φ_ξ^+ are obtained from Equations (53) and (54), respectively, replacing there $\bar{v}(\theta+\sigma)$ by $\xi(\theta,\sigma)$; $E_\xi(\theta, \xi(\theta,\cdot))$ is obtained from Equation (50), replacing there $\bar{v}(\theta+\sigma)$ by $\xi(\theta,\sigma)$, $\sigma \in [0, \Delta\theta]$, and Φ^- and Φ^+ by Φ_ξ^- and Φ_ξ^+, respectively.

C2. Approximation of the NLDDG. Let us approximate the continuous argument σ, varying on the entire interval $[0, \Delta\theta]$, by the discrete variable σ_i given by

$$\sigma_0 = 0, \quad \sigma_i = \sigma_{i-1} + \Delta\theta/K, \quad i = 1, \dots, K, \tag{C5}$$

where $K > 0$ is an integer. Using (C5), one can approximate $\partial\xi(\theta,\sigma)/\partial\sigma$ and all the integrals in Equation (C4) containing $\xi(\theta,\sigma)$ as follows:

$$\partial\xi(\theta,\sigma)/\partial\sigma|_{\sigma=\sigma_i} \approx [\xi(\theta,\sigma_i) - \xi(\theta,\sigma_{i-1}]/(\Delta\theta/K), \quad i = 1, \dots, K \tag{C6}$$

for almost all $\theta \in [0, \theta_0]$, and

$$\begin{aligned}&\int_{\Delta\theta}^{0} \exp(-\sigma/\varepsilon)\xi(\theta,\sigma)d\sigma \\ &\qquad \approx -(\Delta\theta/K)\sum_{i=1}^{K} \exp(-\sigma_i/\varepsilon)\xi(\theta,\sigma_i) \qquad \forall\ \theta \in [0,\theta_0],\end{aligned} \tag{C7}$$

$$\begin{aligned}&\int_{\Delta\theta}^{0} [1-\exp(-\sigma/\varepsilon)]\xi(\theta,\sigma)d\sigma \\ &\qquad \approx -(\Delta\theta/K)\sum_{i=1}^{K} [1-\exp(-\sigma_i/\varepsilon)]\xi(\theta,\sigma_i), \qquad \forall\ \theta \in [0,\theta_0],\end{aligned} \tag{C8}$$

$$\begin{aligned}&\int_{\Delta\theta}^{0} \Psi(\sigma/\varepsilon)\xi(\theta,\sigma) \\ &\qquad \approx -(\Delta\theta/K)\sum_{i=1}^{K} \Psi(\sigma_i/\varepsilon)\xi(\theta,\sigma_i) \qquad \forall\ \theta \in [0,\theta_0].\end{aligned} \tag{C9}$$

By introducing new state variables

$$\xi_i(\theta) \triangleq \xi(\theta,\sigma_i), \quad i = 1, \dots, K, \tag{C10}$$

and applying (C7)–(C10), one can approximate (C4) by

$$\begin{aligned} d\bar{z}_{imc}(\theta)/d\theta = &\mu\Psi(\theta)\bar{u}(\theta) - (\delta/K)\theta\sum_{i=1}^{K-1}\exp(-\sigma_i/\varepsilon)\xi_i(\theta) \\ &- \exp(-\delta)[\varepsilon\Psi(\theta/\varepsilon) + (\delta/K)\theta]\xi_K(\theta) \\ &+ \varepsilon[1-\exp(-\theta/\varepsilon)]Q(\Phi_{K\xi}^{-}, \Phi_{K\xi}^{+}) \\ &+ E_{K\xi}(\theta, \xi_1(\theta), \dots, \xi_K(\theta)),\end{aligned} \tag{C11}$$

where

$$\Phi_{K\xi}^{-} \triangleq 0.5(\delta/K)\sum_{i=1}^{K}\Psi(\sigma_i/\varepsilon)\xi_i(\theta) + \beta, \tag{C12}$$

$$\Phi^{+}_{K\xi} \triangleq -0.5(\delta/K)\sum_{i=1}^{K}\Psi(\sigma_i/\varepsilon)\xi_i(\theta) + \beta, \qquad (C13)$$

$$\begin{aligned} E_{K\xi}(\theta, \xi_1(\theta), \dots, \xi_K(\theta)) &\triangleq -(\varepsilon/2)\Psi(\theta/\varepsilon)R(\Phi^{-}_{K\xi}, \Phi^{+}_{K\xi}) \\ &\quad \times [\Psi(\delta)\xi_K(\theta) - (\delta/K)\sum_{i=1}^{K}(1 - \exp(-\sigma_i/\varepsilon))\xi_i(\theta)]. \end{aligned} \qquad (C14)$$

Similarly, using (C6) and (C10), one can approximate (C2) as follows:

$$d\xi_1(\theta)/d\theta = (K/\Delta\theta)[\xi_1(\theta) - \bar{v}(\theta)], \qquad (C15)$$

$$d\xi_i(\theta)/d\theta = (K/\Delta\theta)[\xi_i(\theta) - \xi_{i-1}(\theta)], \qquad i = 2, \dots, K. \qquad (C16)$$

Due to (58), (C1) and (C10), the state variables $\xi_i(\theta)$, $(i = 1, \dots, K)$ must satisfy the constraints

$$|\xi_i(\theta)| \le 1 - \nu, \qquad i = 1, \dots, K, \qquad (C17)$$

and the initial conditions

$$\xi_i(\theta_0) = \bar{v}_0(\theta_0 + \sigma_i), \qquad i = 1, \dots, K. \qquad (C18)$$

It can be shown similarly to [30] that the solution of Equations (C15), (C16) with the initial conditions (C18) satisfies (C17) as $\bar{v}(\theta)$ satisfies (58). Hence, (C17) can be omitted in the sequel. Equations (C11), (C15), (C16) with the control constraints (21), (58) and the cost function (55) describe an auxiliary game approximating the NLDDG.

C3. Solution of the auxiliary game. The variational Hamiltonian of the auxiliary game has the form

$$\begin{aligned} H = \psi_0(\theta)\Big\{ &\mu\Psi(\theta)\bar{u}(\theta) - (\delta/K)\theta\sum_{i=1}^{K-1}\exp(-\sigma_i/\varepsilon)\xi_i(\theta) \\ &- \exp(-\delta)[\varepsilon\Psi(\theta/\varepsilon) + (\delta/K)\theta]\xi_K(\theta) \\ &+ \varepsilon[1 - \exp(-\theta/\varepsilon)]Q(\Phi^{-}_{K\xi}, \Phi^{+}_{K\xi}) + E_{K\xi}(\theta, \xi_1(\theta), \dots, \xi_K(\theta))\Big\} \\ &+ \psi_1(\theta)(K/\Delta\theta)[\xi_1(\theta) - \bar{v}(\theta)] + \dots + \psi_i(\theta)(K/\Delta\theta)[\xi_i(\theta) - \xi_{i-1}(\theta)] \\ &+ \dots + \psi_K(\theta)(K/\Delta\theta)[\xi_K(\theta) - \xi_{K-1}(\theta)], \end{aligned} \qquad (C19)$$

where ψ_i, $i = 0, 1, \dots, K$, are costate variables. The necessary conditions of the optimality for this game are

$$\bar{u}^*(\theta) = \arg\min_{\bar{u}} H = -\mathrm{sign}[\mu\psi_0(\theta)\Psi(\theta)] = -\mathrm{sign}[\psi_0(\theta)], \qquad (C20)$$

$$\bar{v}^*(\theta) = \arg\max_{\bar{v}} H = -(1-\nu)\mathrm{sign}[(K/\Delta\theta)\psi_1(\theta)]. \tag{C21}$$

The costate variables satisfy the adjoint equations

$$d\psi_0(\theta)/d\theta = -\partial H/\partial \bar{z}_{imc} = 0, \tag{C22}$$

$$\begin{aligned}
d\psi_i(\theta)/d\theta &= -\partial H/\partial \xi_i \\
&= \Bigg\{ (\delta/K)\theta \exp(-\sigma_i/\varepsilon) \\
&\quad + 0.5(\Delta\theta/K)[1-\exp(-\theta/\varepsilon)]\Psi(\sigma_i/\varepsilon)R(\Phi^-_{K\xi}, \Phi^+_{K\xi}) \\
&\quad + 0.25(\Delta\theta/K)\Psi(\theta/\varepsilon)\Psi(\sigma_i/\varepsilon)S(\Phi^-_{K\xi}, \Phi^+_{K\xi}) \\
&\quad \times \left[\Psi(\delta)\xi_K(\theta) - (\delta/K)\sum_{i=1}^{K}(1-\exp(-\sigma_i/\varepsilon))\xi_i(\theta)\right] \\
&\quad - 0.5(\Delta\theta/K)[1-\exp(-\sigma_i/\varepsilon)]\Psi(\theta/\varepsilon)R(\Phi^-_{K\xi}, \Phi^+_{K\xi}) \Bigg\} \psi_0(\theta) \\
&\quad - (K/\Delta\theta)[\psi_i(\theta) - \psi_{i+1}(\theta)], \qquad i = 1, \ldots, K-1,
\end{aligned} \tag{C23}$$

$$\begin{aligned}
d\psi_K(\theta)/d\theta &= -\partial H/\partial \xi_K \\
&= \Bigg\{ \exp(-\delta)[\varepsilon\Psi(\theta/\varepsilon) + (\delta/K)\theta] \\
&\quad + 0.5(\Delta\theta/K)[1-\exp(-\theta/\varepsilon)]\Psi(\delta)R(\Phi^-_{K\xi}, \Phi^+_{K\xi}) \\
&\quad + 0.25(\Delta\theta/K)\Psi(\theta/\varepsilon)\Psi(\delta)S(\Phi^-_{K\xi}, \Phi^+_{K\xi}) \\
&\quad \times \left[\Psi(\delta)\xi_K(\theta) - (\delta/K)\sum_{i=1}^{K}(1-\exp(-\sigma_i/\varepsilon))\xi_i(\theta)\right] \\
&\quad + 0.5\varepsilon\Psi(\theta/\varepsilon)R(\Phi^-_{K\xi}, \Phi^+_{K\xi})[\Psi(\delta) - (\delta/K)(1-\exp(-\delta))] \Bigg\} \psi_0(\theta) \\
&\quad - (K/\Delta\theta)\psi_K(\theta),
\end{aligned} \tag{C24}$$

and transversality conditions

$$\psi_0(0) = -\partial \bar{J}_{im}/\partial \bar{z}_{imc}(0) = -\mathrm{sign}[\bar{z}_{imc}(0)], \tag{C25}$$

$$\psi_i(0) = -\partial \bar{J}_{im}/\partial \xi_i(0) = 0, \qquad i = 1, \ldots, K. \tag{C26}$$

C4. Transformation of Equations (C20)–(C26) and limit process as $K \to +\infty$. Let the variables $\psi_i(\theta)$ in (C20)–(C26) be transformed into the variables $\varphi_i(\theta)$, $(i = 0, \ldots, K)$ as follows:

$$\varphi_0(\theta) = \psi_0(\theta); \quad \varphi_i(\theta) = (K/\Delta\theta)\psi_i(\theta), \qquad i = 1, \ldots, K. \tag{C27}$$

Due to this transformation, Equations (C20)–(C26) become

$$\bar{u}^*(\theta) = -\mathrm{sign}[\varphi_0(\theta)], \tag{C28}$$

$$\bar{v}^*(\theta) = -(1-\nu)\mathrm{sign}[\varphi_1(\theta)], \tag{C29}$$

$$d\varphi_0(\theta)/d\theta = 0, \tag{C30}$$

$$\begin{aligned} d\varphi_i(\theta)/d\theta = &\Bigg\{(\theta/\varepsilon)\exp(-\sigma_i/\varepsilon) \\ &+ 0.5[1-\exp(-\theta/\varepsilon)]\Psi(\sigma_i/\varepsilon)R(\Phi^-_{K\xi},\Phi^+_{K\xi}) \\ &+ 0.25\Psi(\theta/\varepsilon)\Psi(\sigma_i/\varepsilon)S(\Phi^-_{K\xi},\Phi^+_{K\xi}) \\ &\times\left[\Psi(\delta)\xi_K(\theta) - (\delta/K)\sum_{i=1}^{K}(1-\exp(-\sigma_i/\varepsilon))\xi_i(\theta)\right] \\ &- 0.5\Psi(\theta/\varepsilon)R(\Phi^-_{K\xi},\Phi^+_{K\xi})[1-\exp(-\sigma_i/\varepsilon)]\Bigg\}\varphi_0(\theta) \\ &- [\varphi_i(\theta)-\varphi_{i+1}(\theta)]/(\Delta\theta/K), \qquad i=1,\dots,K-1, \end{aligned} \tag{C31}$$

$$\begin{aligned} (\Delta\theta/K)d\varphi_K(\theta)/d\theta = &\Bigg\{\exp(-\delta)[\varepsilon\Psi(\theta/\varepsilon)+(\delta/K)\theta] \\ &+ 0.5(\Delta\theta/K)[1-\exp(-\theta/\varepsilon)]\Psi(\delta)R(\Phi^-_{K\xi},\Phi^+_{K\xi}) \\ &+ 0.25(\Delta\theta/K)\Psi(\theta/\varepsilon)\Psi(\delta)S(\Phi^-_{K\xi},\Phi^+_{K\xi}) \\ &\times\left[\Psi(\delta)\xi_K(\theta) - (\delta/K)\sum_{i=1}^{K}(1-\exp(-\sigma_i/\varepsilon))\xi_i(\theta)\right] \\ &+ 0.5\varepsilon\Psi(\theta/\varepsilon)R(\Phi^-_{K\xi},\Phi^+_{K\xi}) \\ &\times\Big[\Psi(\delta)-(\delta/K)(1-\exp(-\delta))\Big]\Bigg\}\varphi_0(\theta) - \varphi_K(\theta), \end{aligned} \tag{C32}$$

$$\varphi_0(0) = -\mathrm{sign}[\bar{z}_{imc}(0)], \tag{C33}$$

$$\varphi_i(0) = 0, \qquad i=1,\dots,K. \tag{C34}$$

By introducing a continuous function of two variables $\lambda(\theta,\sigma)$ having first-order partial derivatives and satisfying

$$\lambda(\theta,\sigma_i) = \varphi_i(\theta), \qquad i=1,\dots,K, \tag{C35}$$

and substituting (C35) into Equations (C29), (C31), (C32) and (C34), one has, respectively,

$$\bar{v}^*(\theta) = -(1-\nu)\mathrm{sign}[\lambda(\theta,\Delta\theta/K)], \tag{C36}$$

$$
\begin{aligned}
&\partial\lambda(\theta,\sigma)/\partial\theta - [\lambda(\theta,\sigma_{i+1}) - \lambda(\theta,\sigma_i)]/(\Delta\theta/K) \\
&\quad = \Bigg\{ (\theta/\varepsilon)\exp(-\sigma_i/\varepsilon) + 0.5R(\Phi^-_{K\xi},\Phi^+_{K\xi}) \\
&\qquad \times [(1-\exp(-\theta/\varepsilon))\Psi(\sigma_i/\varepsilon) - (1-\exp(-\sigma_i/\varepsilon))\Psi(\theta/\varepsilon)] \\
&\qquad + 0.25\Psi(\theta/\varepsilon)\Psi(\sigma_i/\varepsilon)S(\Phi^-_{K\xi},\Phi^+_{K\xi}) \\
&\qquad \times \left[\Psi(\delta)\xi_K(\theta) - (\delta/K)\sum_{i=1}^{K}(1-\exp(-\sigma_i/\varepsilon))\xi_i(\theta)\right]\Bigg\}\varphi_0(\theta), \qquad (C37)
\end{aligned}
$$

$$
\begin{aligned}
&(\Delta\theta/K)\partial\lambda(\theta,\Delta\theta)/\partial\theta \\
&\quad = \Bigg\{ \exp(-\delta)[\varepsilon\Psi(\theta/\varepsilon) + (\delta/K)\theta] \\
&\qquad + 0.5(\Delta\theta/K)[1-\exp(-\theta/\varepsilon)]\Psi(\delta)R(\Phi^-_{K\xi},\Phi^+_{K\xi}) \\
&\qquad + 0.25(\Delta\theta/K)\Psi(\theta/\varepsilon)\Psi(\delta)S(\Phi^-_{K\xi},\Phi^+_{k\xi}) \\
&\qquad \times \left[\Psi(\delta)\xi_K(\theta) - (\delta/K)\sum_{i=1}^{K}(1-\exp(-\sigma_i/\varepsilon))\xi_i(\theta)\right] \\
&\qquad + 0.5\varepsilon\Psi(\theta/\varepsilon)R(\Phi^-_{K\xi},\Phi^+_{K\xi}) \\
&\qquad \times [\Psi(\delta) - (\delta/K)(1-\exp(-\delta))]\Bigg\}\varphi_0(\theta) - \lambda(\theta,\Delta\theta), \qquad (C38)
\end{aligned}
$$

$$
\lambda(0,\sigma_i) = 0, \qquad i = 1,\dots,K. \qquad (C39)
$$

Next, the limit of Equations (C28), (C30), (C33) and (C36)–(C39) as $K \to +\infty$ is calculated. The first three equations coincide with (60), (62) and (69), respectively. Calculating the limit of (C37)–(C39) using (C10) and the fact that the discrete variable σ_i tends to the continuous variable σ as $K \to +\infty$ yields directly (63), (70), (71), respectively. Calculating the limit of (C36) yields (61). Thus, the proposition is proved.

Appendix D: Proof of Proposition 4.3

First, note that $\bar{v}_0^*(\theta + \Delta\theta)$ satisfying Equation (102) must be of a bang-bang structure. It has to be shown that this function has at least one switch point on the interval $\theta \in (\theta_0 - \Delta\theta, \theta_0)$. This is proved by contradiction. Assume that $\bar{v}_0^*(\theta + \Delta\theta)$ keeps its sign on this interval. In this case one obtains directly from (83) and (84) that one of the values, either $\Phi^-_\nu|_{\theta=\theta_0}$ or $\Phi^+_\nu|_{\theta=\theta_0}$, tends to 1 as $\nu \to +0$, which yields together with Equations (52), (101)

$$
\lim_{\nu \to +0} R^*_{\nu 0} = -\infty. \qquad (D1)
$$

Using Equations (103) and (D1), one has by direct calculation

$$C(\theta, \bar{v}_0^*(\cdot))|_{\theta=\theta_0-\Delta\theta} > 0, \qquad C(\theta, \bar{v}_0^*(\cdot))|_{\theta=\theta_0} < 0. \tag{D2}$$

The inequalities in (D2) imply that $\bar{v}_0^*(\theta+\Delta\theta)$, satisfying (102), has at least one switch point on the interval $\theta \in (\theta_0 - \Delta\theta, \theta_0)$, which contradicts the assumption made above. Thus, this function has indeed at least one switch point, and, due to the first inequality in (D2), it equals $(1-\nu)\mathrm{sign}[\bar{z}_{imc}(0)]$ in a right-hand neighborhood of $\theta = \theta_0 - \Delta\theta$.

Next, it has to be shown that there is no more than one switch point of the function $\bar{v}_0^*(\theta + \Delta\theta)$ satisfying (102). For this purpose it is sufficient to show that $C(\theta, \bar{v}_0^*(\cdot))$ is monotonic with respect to θ.

Since $C(\theta, \bar{v}_0^*(\cdot))$ has at least one root with respect to θ on the interval $(\theta_0 - \Delta\theta, \theta_0)$, Equation (103) yields that

$$R_{\nu 0}^* < 0. \tag{D3}$$

Now, calculating the first-order derivative of $C(\theta, \bar{v}_0^*(\cdot))$ with respect to θ and using (D3), one directly obtains

$$dC(\theta, \bar{v}_0^*(\cdot))/d\theta < 0, \tag{D4}$$

yielding a monotonic decrease of $C(\theta, \bar{v}_0^*(\cdot))$ with respect to θ on the interval $(\theta_0 - \Delta\theta, \theta_0)$. Thus, the switch point is unique. By denoting this point θ_{sw}, and setting $\sigma_{sw} \triangleq \theta_{sw} - \theta_0 + \Delta\theta$, (105) is obtained for θ_{sw}. Substituting (105) into (103), one directly obtains (106) for σ_{sw}. Thus, the proposition is proved.

REFERENCES

[1] Barman J.F., Callier F.M. and Desoer C.A., L^2-stability and L^2-instability of linear time-invariant distributed feedback system perturbed by a small delay in the loop, *IEEE Transactions on Automatic Control*, 18, 479–484, 1973.

[2] Bellman R. and Cooke K.L., *Differential-Difference Equations*, Academic Press, New York, 1963.

[3] Bensoussan A., Da Prato G., Delfour M.C. and Mitter S.K., *Representation and Control of Infinite Dimensional Systems*, Vol. 2, Birkhäuser, Boston, 1992.

[4] Chikrii G.Ts., An approach to the solution of linear differential games with variable information delay, *Journal of Automation and Information Sciences*, 27, 163–171, 1995.

[5] Chikrii G.Ts., Using the impact of information delay for solution of game problems of pursuit, *Dopovidi Natsional'noi Akagemii Nauk Ukraini*, 12, 107–111, 1999.

[6] Ciletti M.D., New results in the theory of differential games with information time lag, *Journal of Optimization Theory and Applications*, 8, 287–315, 1971.

[7] Datko R., Two examples of ill-posedness with respect to time delays revisited, *IEEE Transactions on Automatic Control*, 42, 511–515, 1997.

[8] Datko R., Lagnese J. and Polis M.P., An example of the effect of time delays in boundary feedback stabilization of wave equations, *SIAM Journal of Control and Optimization*, 24, 152–156, 1986.

[9] El'sgol'ts L.E. and Norkin S.B., *Introduction to the Theory and Applications of Differential Equations with Deviating Arguments*, Academic Press, London, 1973.

[10] Glizer V.Y. and Shinar J., Optimal evasion from a pursuer with delayed information, *Journal of Optimization Theory and Applications*, 111, 7–38, 2001.

[11] Gu K. and Niculescu S.-I., Survey on recent results in the stability and control of time delay systems, *Journal of Dynamic Systems, Measurements and Control*, 125, 158–165, 2003.

[12] Halanay A., *Differential Equations: Stability, Oscillations, Time Lags*, Academic Press, New York, 1966.

[13] Halanay A., Differential games with delay, *SIAM Journal of Control*, 6, 579–593, 1968.

[14] Hale J., *Theory of Functional Differential Equations*, Springer-Verlag, New York, 1977.

[15] Ho J.C., Bryson A.E. and Baron S., Differential games and optimal pursuit-evasion strategies, *IEEE Transactions on Automatic Control*, 10, 385–389, 1965.

[16] Ioffe A. and Tikhomirov V.M., *Theory of Extremal Problems*, North-Holland, Amsterdam, 1979.

[17] Kolmanovskii V.B. and Nosov V.R., *Stability of Functional Differential Equations*, Academic Press, London, 1986.

[18] Krasovskii N.N. and Osipov Ju.S., Linear differential-difference games, *Soviet Mathematics Doklady*, 12, 554–558, 1971.

[19] Kryazhimskii A.V. and Osipov Ju.S., Differential-difference approach

games with a functional target set, *Journal of Applied Mathematics and Mechanics*, 37, 3–13, 1973.

[20] Kumkov S.I. and Patsko V.S., Optimal strategies in a pursuit problem with incomplete information, *Journal of Applied Mathematics and Mechanics*, 59, 75–85, 1995.

[21] Lukoyanov N.Yu. and Reshetova T.N., Problems of conflict control of high dimensionality functional systems, *Journal of Applied Mathematics and Mechanics*, 62, 545–554, 1998.

[22] Nikol'skii M.S., Linear differential games of pursuit with delays, *Soviet Mathematics Doklady*, 12, 614–618, 1971.

[23] Osipov Ju.S., Differential games of systems with aftereffect, *Soviet Mathematics Doklady*, 12, 262–266, 1971.

[24] Osipov Ju.S., Alternative in a differential-difference game, *Soviet Mathematics Doklady*, 12, 619–624, 1971.

[25] Petrosjan L.A., *Differential Pursuit Games*, Leningrad University Publishing, Leningrad, 1977 (in Russian). [English translation: *Differential Games of Pursuit*, World Scientific Publishing, Singapore, 1993.]

[26] Richard J.-P., Time-delay systems: an overview of some recent advances and open problems, *Automatica*, 39, 1667–1694, 2003.

[27] Richard J.-P., Dambrine M., Gouaisbaut F. and Perruquetti W., Systems with delays: an overview of some recent advances, *Stability and Control: Theory and Applications*, 3, 3–23, 2000.

[28] Shinar J., Solution for realistic pursuit-evasion games, in C.T. Leondes (ed.), *Advances in Control and Dynamic Systems*, Vol. 17, 63–124, Academic Press, New York, 1981.

[29] Shinar J. and Glizer V.Y., Solution of a delayed information linear pursuit-evasion game with bounded controls, *International Game Theory Review*, 1, 197–218, 1999.

[30] Shinar J. and Glizer V.Y., Complete solution of a delayed information linear pursuit-evasion game with bounded controls, Research Report TAE 841, Technion-Israel Institute of Technology, 1999.

[31] Shinar J. and Glizer V.Y., A linear pursuit-evasion game with time varying information delay, Research Report TAE 889, Technion-Israel Institute of Technology, 2002.

[32] Shinar J. and Shima T., Robust missile guidance law against highly maneuvering targets, in *Proceedings of the 7th Mediterranean Conference on Control and Automation*, 1548–1572, 1999.

[33] Shinar J. and Shima T., Nonorthodox guidance law development approach for intercepting maneuvering targets, *Journal of Guidance, Control, and Dynamics*, 25, 658–666, 2002.

[34] Shinar J., Shima T. and Glizer V.Y., On the compensation of imperfect information in dynamic games, *Intenational Game Theory Review*, 2, 229–248, 2000.

[35] Sokolov B.N. and Chernous'ko F.L., Differential games with information lag, *Journal of Applied Mathematics and Mechanics*, 34, 812–819, 1970.

[36] Soliman M.A. and Ray W.H., Optimal feedback control for linear-quadratic systems having time delays, *International Journal of Control*, 15, 609–627, 1972.

[37] Van Keulen B., *H_∞-Control for Distributed Parameter Systems: A State-Space Approach*, Birkhäuser, Boston, 1993.

Game Problems for Systems with Fractional Derivatives of Arbitrary Order

Arkadii A. Chikrii
Institute of Cybernetics NAS
Kiev, Ukraine
chik@insyg.ua

Abstract

Conflict-controlled processes for systems with Riemann–Liouville derivatives of arbitrary order are studied here. A solution of such a system is presented in the form of a Cauchy formula analog. Using the resolving functions method, sufficient conditions for termination of the game are obtained. These conditions are based on the modified Pontryagin condition, expressed in terms of the generalized matrix functions of Mittag–Leffler. To find the latter, the interpolating polynomial of Lagrange–Sylvester is used. An illustrative example is given.

Key words. Fractional derivative, dynamic game, Mittag–Leffler function.

1 Introduction

A system described by equations with fractional derivatives [12] is one example of systems which are not dynamic according to Birkhoff because of their failure to meet the semigroup property. This feature is an essential obstacle to the development of optimality conditions. However, such processes can be studied on the basis of the principle of guaranteed result [1]. It seems likely that the paper of Eidelman–Chikrii–Rurenko [7] was among the first in which game problems for systems with fractional derivatives were analyzed. The more thorough studies of Chikrii–Eidelman [6] contain sufficient conditions for solvability of the pursuit problem for systems with fractional derivatives of arbitrary order α, $0 < \alpha < 1$. In so doing, both the systems with fractional derivatives of Riemann–Liouville and those with Dzhrbashyan–Nersesyan regularized derivatives are treated. Using Dzhrbashyan asymptotic formulas for functions of Mittag–Leffler [5], sufficient conditions for the game termination are obtained in the case of a simple matrix system. In Chikrii–Eidelman [3] these results are developed in the case of arbitrary order α of Riemann–Liouville derivatives. The generalized matrix functions of Mittag–Leffler, first introduced in Chikrii–Eidelman [4], play a key role in this investigation.

2 Problem Statement and Auxiliary Results

Let us consider the following conflict-controlled process with two players as participants. The evolution of each of them is described, respectively, by the following equations with fractional derivatives:

$$D^{\alpha}x = Ax + u, \quad x \in \mathbb{R}^{n_1},\ u \in U,\ k-1 < \alpha \leq k,\ k \geq 1, \tag{1}$$

$$D^{\beta}y = By + v, \quad y \in \mathbb{R}^{n_2},\ v \in V,\ m-1 < \beta \leq m,\ m \geq 1. \tag{2}$$

Here

$$D^{\alpha}x = \frac{d^{\alpha}}{dt^{\alpha}}x = \frac{1}{\Gamma(1-\{\alpha\})}\left(\frac{d}{dt}\right)^{[\alpha]+1}\int\limits_0^t \frac{x(s)}{(t-s)^{\{\alpha\}}}ds,$$

$[\alpha]$, $\{\alpha\}$ are, respectively, the integer and the fractional parts of number α, $\alpha > 0$, D^{α} is the α-order left-side Riemann–Liouville derivative [12], and $\Gamma(\alpha)$ is the γ-function. The fractional derivative D^{β} is defined in a similar manner. Evidently, there exist positive integers k and m, furnishing the inequalities (1), (2) for α and β. The matrices A and B are square matrices of order n_1 and n_2, and the players' control domains U and V are compact sets in finite-dimensional Euclidean spaces $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$, respectively, that is $U \in K(\mathbb{R}^{n_1})$, $V \in K(\mathbb{R}^{n_2})$.

The initial conditions for systems (1), (2) are given in the form

$$\begin{aligned} \left.\frac{d^{\alpha-i}}{dt^{\alpha-i}}x(t)\right|_{t=0} &= x_i^0,\ i = 1, \dots, k, \\ \left.\frac{d^{\beta-j}}{dt^{\beta-j}}x(t)\right|_{t=0} &= y_j^0,\ j = 1, \dots, m. \end{aligned} \tag{3}$$

Set $x^0 = (x_1^0, \dots, x_k^0)$, $y^0 = (y_1^0, \dots, y_m^0)$.

For simplicity we assume that the terminal set is defined by the first s components of x and y, $s \leq \min\{n_1, n_2\}$. Let us introduce orthoprojectors $\pi_1 : \mathbb{R}^{n_1} \to \mathbb{R}^s$, $\pi_2 : \mathbb{R}^{n_2} \to \mathbb{R}^s$, extracting first the s components from vectors x and y. Then the terminal set may be described by the inequality

$$\|\pi_1 x - \pi_2 y\| \leq \varepsilon. \tag{4}$$

We will analyze the following game scheme. The first player, exerting control over the process (1), strives in the shortest possible time to achieve fulfillment of the inequality (4) while the second player, controlling the process (2), strives to maximally delay this moment.

Standing on the pursuer's side and choosing its control in the form

$$u(t) = u\left(x^0, y^0, v_t(\cdot)\right), \quad v_t(\cdot) = \{v(s) : s \in [0, t]\}, \tag{5}$$

we will orient ourselves to an arbitrary measurable function with values from the set V as the second player's control choice.

The purpose of this paper consists in the development of sufficient conditions for the game termination in favor of the first player in some guaranteed time.

Consider the Cauchy problem for the system of equations (1), (2) under the initial conditions (3). Following Chikrii–Eidelman [4], we introduce the generalized matrix function of Mittag–Leffler

$$E_\rho(C;\mu) = \sum_{k=0}^{\infty} \frac{C^k}{\Gamma(k\rho^{-1}+\mu)}$$

for arbitrary positive numbers ρ and μ and arbitrary quadratic matrix C of finite order.

Lemma 2.1. *Let the system of differential equations (1), (2) under the initial conditions (3) be given. Then, under the chosen Lebesgue measurable functions as players' controls, the solution of the Cauchy problem is unique, and it has a form*

$$\begin{aligned} x(t) &= \sum_{i=1}^{k} t^{\alpha-i} E_{1/\alpha}\left(At^\alpha; 1+\alpha-i\right) x_i^0 \\ &\quad + \int_0^t (t-\tau)^{\alpha-1} E_{1/\alpha}\left(A(t-\tau)^\alpha;\alpha\right) u(\tau)d\tau, \\ y(t) &= \sum_{j=1}^{m} t^{\beta-j} E_{1/\beta}\left(Bt^\beta; 1+\beta-j\right) y_j^0 \\ &\quad + \int_0^t (t-\tau)^{\beta-1} E_{1/\beta}\left(B(t-\tau)^\beta;\beta\right) v(\tau)d\tau. \end{aligned}$$

3 Method of Problem Solution

Following the resolving function's method scheme of Chikrii [1,2] let us study the following set-valued mappings:

$$W(t,v) = t^{\alpha-1}\pi_1 E_{1/\alpha}\left(At^\alpha;\alpha\right) U - t^{\beta-1}\pi_2 E_{1/\beta}\left(Bt^\beta;\beta\right) C(t)V,$$

$$W(t) = t^{\alpha-1}\pi_1 E_{1/\alpha}\left(At^\alpha;\alpha\right) U \overset{*}{-} t^{\beta-1}\pi_2 E_{1/\beta}\left(Bt^\beta;\beta\right) C(t)V, \tag{6}$$

$$M(t) = \varepsilon S \overset{*}{-} \int_0^t \tau^{\beta-1}\pi_2 E_{1/\beta}\left(B\tau^\beta;\beta\right)\left(E - C(\tau)\right) V d\tau. \tag{7}$$

In the preceding formulas, $C(t)$ is some bounded measurable matrix function, $\overset{*}{-}$ is the operation of geometric subtraction of sets [11], and E is a unit sphere from space $\mathbb{R}^s$ centered at the origin. If there exists a function $C(t)$, such that both mappings $W(t)$ and $M(t)$ take nonempty values for all $t \geq 0$, we will say that Pontryagin condition holds. In the case $C(t) \equiv E$, the condition $M(t) \neq \emptyset$ is readily satisfied, and the condition $W(t) \neq \emptyset$ appears as the regular Pontryagin condition [11].

Let us pick a measurable selection $\gamma(t)$ [10] and set

$$\xi\left(t, x^0, \gamma(\cdot)\right) = \sum_{i=1}^{k} t^{\alpha-i}\pi_1 E_{1/\alpha}\left(At^\alpha; 1+\alpha-i\right) x_i^0 - \sum_{j=1}^{m} t^{\beta-j}\pi_2 E_{1/\beta}\left(Bt^\beta; 1+\beta-j\right) y_j^0 + \int_0^t \gamma(\tau)d\tau.$$

Then, the resolving function [1] for the problem (1)–(4) is

$$\alpha(t,\tau,v) = \sup\left\{\rho \geq 0 : \left[W(t-\tau, v) - \gamma(t-\tau)\right] \cap \rho\left[M(t) - \xi\left(t, x^0, y^0, \gamma(\cdot)\right)\right] \neq \emptyset\right\}. \tag{8}$$

In its terms we define the set

$$T\left(x^0, y^0, \gamma(\cdot)\right) = \left\{t \geq 0 : \int_0^t \inf_{v\in V} \alpha(t,\tau,v)d\tau \geq 1\right\}. \tag{9}$$

Theorem 3.1. *Assume that, for the conflict-controlled process (1)–(4), there exist a bounded measurable matrix function $C(t)$, $t \geq 0$, such that the modified Pontryagin condition holds and a measurable selection $\gamma(t)$, $\gamma(t) \in W(t)$ such that $T\left(x^0, y^0, \gamma(\cdot)\right) \neq \emptyset$, and $T \in T(x^0, y^0, \gamma(\cdot))$, $T < +\infty$.*

Then the game (1)–(5) may be terminated at the instant T.

The proof can be constructed based on Chikrii–Eidelman [4] and Chikrii [1]. Let us illustrate the main theoretical statements on a model example. To find the generalized Mittag–Leffler matrix functions $E_{1/\alpha}\left(At^\alpha; \alpha\right)$, $E_{1/\beta}\left(Bt^\beta; \beta\right)$ we will use the technique of Lagrange–Sylvester interpolating polynomials. This makes it feasible to find the above-mentioned functions through the matrices A and B.

4 Example

Let us consider the problem "Boy and Crocodile" [9,11] generalized to the case of fractional derivatives. Then the motions of the players are described by equations

$$\frac{d^{k\alpha}}{dt^{k\alpha}}x = u,\ \|u\| \leq 1,\ x \in \mathbb{R}^s,\ 0 < \alpha \leq 1, \tag{10}$$

and

$$\frac{d^{m\beta}}{dt^{m\beta}}y = v,\ \|v\| \leq 1,\ y \in \mathbb{R}^s,\ 0 < \beta \leq 1, \tag{11}$$

respectively, where k and m are arbitrary natural numbers.

The initial conditions for systems (10), (11) have the form

$$\left.\frac{d^{i\alpha-1}}{dt^{i\alpha-1}}x(t)\right|_{t=0} = x_i^0, \quad i = 1, \dots, k, \tag{12}$$

$$\left.\frac{d^{j\beta-1}}{dt^{j\beta-1}}y(t)\right|_{t=0} = y_j^0, \quad j = 1, \dots, m. \tag{13}$$

The game (10)–(13) is assumed to be terminated when

$$\| x - y \| \leq \varepsilon, \quad \varepsilon \geq 0. \tag{14}$$

Let us reduce systems (10), (11) to systems of orders α and β, respectively. For this purpose we set

$$x = x_1, \quad \frac{d^\alpha}{dt^\alpha}x_1 = x_2, \quad \dots, \quad \frac{d^\alpha}{dt^\alpha}x_{k-1} = x_k.$$

Then

$$\frac{d^\alpha}{dt^\alpha}x_k = u.$$

Denoting $\bar{x} = (x_1, \dots, x_k)$ we rewrite equation (10) in the equivalent form

$$D^\alpha \bar{x} = A\bar{x} + \bar{u},$$

where $A = A' \otimes E$,

$$A' = \left(\underbrace{\begin{matrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{matrix}}_{k}\right), \qquad \bar{u} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ u \end{pmatrix},$$

and $\otimes$ is the Kronecker matrix product [8]. Here A' is a quadratic matrix of order k, E is a unit matrix of order s, and $\bar{u}$ is a $k \times s$-dimensional vector.

Performing a similar substitution of variables in system (11) and setting $\bar{y} = (y_1, \dots, y_m)$, we obtain

$$D^\beta \bar{y} = B\bar{y} + \bar{v},$$

where $B = B' \otimes E$, B' is a matrix analogous to A' but of order m, and $\bar{v}$ is an ms-dimensional vector.

The orthoprojectors π_1 and π_2 extract components x_1, y_1 from vectors $\bar{x}$ and $\bar{y}$, respectively. Note that

$$E_{1/\alpha}\left(At^\alpha;\alpha\right) = E_{1/\alpha}\left(A't^\alpha;\alpha\right) \otimes E, \quad E_{1/\beta}\left(Bt^\beta;\beta\right) = E_{1/\beta}\left(B't^\beta;\beta\right) \otimes E.$$

Let us find the matrix functions $E_{1/\alpha}\left(A't^\alpha;\alpha\right)$ and $E_{1/\beta}\left(B't^\beta;\beta\right)$. For this purpose we study the matrices A' and B' in detail. The minimal polynomial of matrix A' is λ^k and its k-fold root is a k-fold characteristic number of the matrix A' [8]. Therefore, the values of the function $f(\lambda) = E_{1/\alpha}\left(\lambda t^\alpha;\alpha\right)$ on the spectrum of matrix A' are numbers $f(0)$, $f^{(1)}(0), \dots, f^{(k-1)}(0)$ and the interpolating polynomial of Lagrange–Sylvester [8], has the form

$$r(\lambda) = f(0) + \frac{f'(0)}{1!}\lambda + \dots + \frac{f^{k-1}(0)}{(k-1)!}\lambda^{k-1}.$$

Hence

$$\begin{aligned} r(A') &= f(A') = E_{1/\alpha}\left(A't^\alpha;\alpha\right) \\ &= f(0)E + \frac{f^{(1)}(0)}{1!}A' + \dots + \frac{f^{k-1}(0)}{(k-1)!}(A')^{k-1} \\ &= \begin{pmatrix} 1/\Gamma(\alpha) & t^\alpha/\Gamma(2\alpha) & t^{2\alpha}/\Gamma(3\alpha) & \cdots & t^{(k-1)\alpha}/\Gamma(k\alpha) \\ 0 & 1/\Gamma(\alpha) & t^\alpha/\Gamma(2\alpha) & \cdots & t^{(k-2)\alpha}/\Gamma((k-1)\alpha) \\ 0 & 0 & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & \ddots & \\ 0 & 0 & 0 & \cdots & t^\alpha/\Gamma(2\alpha) \\ 0 & 0 & 0 & \cdots & 1/\Gamma(\alpha) \end{pmatrix}. \end{aligned}$$

A similar reasoning leads to a similar presentation for $E_{1/\beta}(B't^\beta;\beta)$. Performing the necessary calculations we have

$$W(t) = \frac{t^{k\alpha-1}}{\Gamma(k\alpha)}S \overset{*}{-} \frac{t^{m\beta-1}}{\Gamma(m\beta)}C(t)S.$$

Evidently, $W(t) \neq \emptyset$, $C(t) \equiv E$ if and only if $k\alpha = m\beta$. Thus we introduce the function $C_1(t) = c(t)E$, $t \geq 0$, where the scalar function $c(t)$ has the form

$$c(t) = \begin{cases} 1, & t^\delta \geq \frac{\Gamma(k\alpha)}{\Gamma(m\beta)} \\ \frac{\Gamma(m\beta)}{\Gamma(k\alpha)} t^\delta, & t^\delta < \frac{\Gamma(k\alpha)}{\Gamma(m\beta)}, \end{cases}$$

and $\delta = k\alpha - m\beta$.
Therefore

$$W(t) = \left\{ \begin{array}{ll} \left[\frac{t^{k\alpha-1}}{\Gamma(k\alpha)} - \frac{t^{m\beta-1}}{\Gamma(m\beta)}\right] S, & t^\delta \geq \frac{\Gamma(k\alpha)}{\Gamma(m\beta)} \\ 0, & t^\delta < \frac{\Gamma(k\alpha)}{\Gamma(m\beta)} \end{array} \right\} \neq \emptyset \qquad \forall\, t \geq 0.$$

Then

$$M(t) = \left[\varepsilon - \int_0^t \frac{\tau^{m\beta-1}}{\Gamma(m\beta)} \left(1 - c(\tau)\right) d\tau\right] S.$$

Hence, $M(t) \neq \emptyset$ for all $t \geq 0$, if

$$\varepsilon \geq \int_0^t \frac{\tau^{m\beta-1}}{\Gamma(m\beta)} \left(1 - c(\tau)\right) d\tau, \quad t \geq 0. \tag{15}$$

It can be shown that for $\delta > 0$ the inequality (15) takes the form

$$\varepsilon \geq \frac{\left[\frac{\Gamma(k\alpha)}{\Gamma(m\beta)}\right]^{m\beta/\delta}}{\Gamma(m\beta+1)} - \frac{\left[\frac{\Gamma(k\alpha)}{\Gamma(m\beta)}\right]^{k\alpha/\delta}}{\Gamma(k\alpha+1)} = \sigma(\delta). \tag{16}$$

Let us analyze the case when $\delta < 0$. Set

$$\Delta(t) = \frac{t^{m\beta}}{\Gamma(m\beta+1)} - \frac{t^{k\alpha}}{\Gamma(k\alpha+1)} - \sigma(-|\delta|)$$

and denote

$$t^* = \max\left\{ t > \left[\frac{\Gamma(m\beta)}{\Gamma(k\alpha)}\right]^{1/|\delta|} : \quad \forall \tau \in \left(\left[\frac{\Gamma(m\beta)}{\Gamma(k\alpha)}\right]^{1/|\delta|}, t\right] \varepsilon - \Delta(\tau) \geq 0 \right\}.$$

Then

$$M(t) = \left\{ \begin{array}{ll} \varepsilon S, & t \leq \left[\frac{\Gamma(m\beta)}{\Gamma(k\alpha)}\right]^{1/|\delta|} \\ \left[\varepsilon - \Delta(t)\right] S, & t > \left[\frac{\Gamma(m\beta)}{\Gamma(k\alpha)}\right]^{1/|\delta|} \end{array} \right\} \neq \emptyset \quad \forall\, t \in [0, t^*]$$

and therefore for $\delta < 0$ the modified Pontryagin condition holds only on the interval $[0, t^*]$.

Since $0 \in W(t)$, $t \geq 0$, we set $\gamma_1(t) \equiv 0$. Then

$$\xi\left(t, x^0, y^0, 0\right) = \sum_{i=1}^{k} \frac{t^{i\alpha-1}}{\Gamma(i\alpha)} x_i^0 - \sum_{j=1}^{m} \frac{t^{j\beta-1}}{\Gamma(j\beta)} y_j^0.$$

Upon completion of the required calculation, the resolving function (8) takes the form

$$\alpha(t,\tau,v) = \sup\left\{\rho \geq 0 : \left\|\rho\xi\left(t, x^0, y^0, 0\right) - \frac{(t-\tau)^{m\beta-1}}{\Gamma(m\beta)} c(t-\tau)v\right\| = \frac{(t-\tau)^{k\alpha-1}}{\Gamma(k\alpha)} + \rho m(t)\right\},$$

and

$$\min_{\|v\|\leq 1} \alpha\left(t,\tau,v\right) = \frac{\dfrac{(t-\tau)^{k\alpha-1}}{\Gamma(k\alpha)} - \dfrac{(t-\tau)^{m\beta-1}}{\Gamma(m\beta)} c(t-\tau)}{\|\xi\left(t, x^0, y^0, 0\right)\| - m(t)},$$

where

$$m(t) = \begin{cases} \frac{t^{k\alpha}}{\Gamma(k\alpha+1)} - \frac{t^{m\beta}}{\Gamma(m\beta+1)} + \varepsilon, & \delta > 0, \quad t < \left[\frac{\Gamma(k\alpha)}{\Gamma(m\beta)}\right]^{1/\delta} \\ -\sigma(\delta) + \varepsilon, & \delta > 0, \quad t \geq \left[\frac{\Gamma(k\alpha)}{\Gamma(m\beta)}\right]^{1/\delta} \\ \varepsilon, & \left(\delta < 0, \quad t \leq \left[\frac{\Gamma(m\beta)}{\Gamma(k\alpha)}\right]^{1/|\delta|}\right) \cup \left(\delta = 0, t \geq 0\right) \\ \varepsilon - \Delta(t), & \delta < 0, \quad \left[\frac{\Gamma(m\beta)}{\Gamma(k\alpha)}\right]^{1/|\delta|} \leq t \leq t_*\ . \end{cases}$$

The analysis of all possible cases shows that the least element of set (9) or, what is the same, the moment of the game termination is to be sought as a solution of the equation

$$\frac{t^{k\alpha}}{\Gamma(k\alpha+1)} - \frac{t^{m\beta}}{\Gamma(m\beta+1)} + \varepsilon = \left\|\sum_{i=1}^{k} \frac{t^{i\alpha-1}}{\Gamma(i\alpha)} x_i^0 - \sum_{j=1}^{m} \frac{t^{j\beta-1}}{\Gamma(j\beta)} y_j^0\right\|. \qquad (17)$$

If at $t = 0$ the right side of (17) is greater than the left one, then this equation has a solution due to the left part's greater growth rate as $t \to +\infty$.

Note that from relations (16), (17) for $k = 2$, $\alpha = \beta = m = 1$ the classic result of L. S. Pontryagin [11] immediately follows for the regular problem "Boy and Crocodile." It consists in the fact that the corresponding equation (17) always has a solution when $\varepsilon \geq 1/2$.

REFERENCES

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

[2] Chikrii A.A., Quasilinear controlled processes under conflict, *J. Math. Sci., Dynamical Systems—2*, 80, 1, 1489–1518, 1996.

[3] Chikrii A.A. and Eidelman S.D., Game problems of control for quasi-linear systems with fractional derivatives of Riemann–Liouville, *Kibernetica i Sistemnyi Analiz*, 6, 66–99, 2001.

[4] Chikrii A.A. and Eidelman S.D., Generalized matrix functions of Mittag–Leffler in game problems for evolutionary equations of fractional order, *Kibernetica i Sistemnyi Analiz*, 3, 3–32, 2000.

[5] Dzhrbashyan M.M., *Integral Transformations and Function Representations in Complex Plane*, Nauka, Moscow, 1966.

[6] Eidelman S.D. and Chikrii A.A., Game problems for systems with Volterra evolution. Fractal games, *Game Theory and Appl.*, 6, Nova Science Publishers, Inc., Huntington, N.Y., USA, 2001, 9–44.

[7] Eidelman S.D., Chikrii A.A. and Rurenko A.G., Game problems for fractional systems, *Dokl. Akad. Nauk Ukrainy*, 1, 92–96, 1999.

[8] Gantmakher F.R., *Matrix Theory*, Nauka, Moscow, 1967.

[9] Isaacs R., *Differential Games*, John Wiley, New York, 1965.

[10] Joffe A.D. and Tikhomirov V.M., *Theory of Extremal Problems*, Nauka, Moscow, 1974.

[11] Pontryagin L.S., *Selected Scientific Works*, Vol. 2, Nauka, Moscow, 1988.

[12] Samko S.G., Kilbas A.A. and Marychev O.I., *Fractional Order Integrals and Derivatives, and Some Applications*, Minsk, Tekhnika, 1987.

On Two Problems of Group Pursuit*

N. N. Petrov
Udmurt State University
Izevsk, Russia
npetrov@udmnet.ru

D. A. Vagin
Udmurt State University
Izevsk, Russia
davagin@udm.ru

Abstract

Necessary conditions are obtained for the capture of several evaders in a group pursuit problem, where all evaders use the same control. Necessary conditions for capture in such a group pursuit problem are also obtained for the special case of "soft" capture.

Key words. Differential game, group pursuit, capture, evasion.

1 Introduction

In B. N. Pshenichnij [1] necessary and sufficient conditions were obtained for the capture of one evader in a group pursuit problem

with common dynamics (law of motion). The problem is similar to the one treated by L. S. Pontrjagin [2] but with many participants and with an identical law of motion and initial resources. In Ref. [3], the capture condition is given in terms of phase coordinates. In this chapter we consider two problems. In the first problem sufficient conditions for the capture of at least one evader in a group pursuit problem à la Pontrjagin are obtained, with the assumption that all evaders use the same control. In the second problem the conditions of "soft" capture for the group pursuit problem à la Pontrjagin with a group of pursuers and one evader are also obtained. The article is related to Refs. [4–15].

2 The Problem of Group Pursuit of Hard-United Evaders

In the space R^k, we consider an $n+m$-person differential game with n pursuers $P_1, \dots, P_n$ and m evaders $E_1, \dots, E_m$. The law of motion of each pursuer P_i is

*Supported by RFFI (project 03-01-00014) and by the Competition Centre of Fundamental Natural Science (project E02-1.0-100).

given by

$$x_i^{(l)} + a_1 x_i^{(l-1)} + \cdots + a_l x_i = u_i, \quad \|u_i\| \leqslant 1, \quad x_i \in R^k. \tag{1}$$

The law of motion of each evader E_j is given by

$$y_j^{(l)} + a_1 y_j^{(l-1)} + \cdots + a_l y_j = v, \quad \|v\| \leqslant 1, \quad y_j \in R^k. \tag{2}$$

Here x_i, y_j, u_i, $v \in R^k$, $a_1, \ldots, a_l \in R^1$.

The initial conditions at $t = 0$ are

$$x_i^{(\alpha)}(0) = x_{i,\alpha}^0, \quad y_j^{(\alpha)}(0) = y_{j,\alpha}^0, \quad \alpha = 0, \ldots, l-1. \tag{3}$$

Here $x_{i,0}^0 - y_{j,0}^0 \neq 0$ for all i and j.

Instead of systems (1)—(3) let's consider the system

$$z_{i,j}^{(l)} + a_1 z_{i,j}^{(l-1)} + \cdots + a_l z_{i,j} = u_i - v, \quad \|u_i\| \leq 1, \ \|v\| \leq 1, \tag{4}$$

$$z_{i,j}(0) = z_{i,j,0}^0 = x_{i,0}^0 - y_{j,0}^0, \ldots, z_{i,j}^{(l-1)}(0) = z_{i,j,l-1}^0 = x_{i,l-1}^0 - y_{j,l-1}^0. \tag{5}$$

Definition 2.1. In the game Γ there is a capture, if there exists a time $T > 0$ and measurable functions $u_i(t) = u_i(t, x_{i\alpha}^0, y_{j\alpha}^0, v_t(\cdot))$, $\|u_i(t)\| \leq 1$, such that for any measurable function $v(\cdot)$, $\|v(t)\| \leq 1$, $t \in [0, T]$ there exist $\tau \in [0, T]$ and numbers $q \in \{1, 2, \ldots, n\}$, $r \in \{1, \ldots, m\}$, such that

$$x_q(\tau) = y_r(\tau).$$

Let $\varphi_q(t)$, $q = 0, 1, \ldots, l-1$ be solutions of the equation

$$\omega^{(l)} + a_1 \omega^{(l-1)} + \cdots + a_l \omega = 0 \tag{6}$$

with initial conditions

$$\omega(0) = 0, \ldots, \omega^{(q-1)}(0) = 0, \omega^{(q)}(0) = 1, \omega^{(q+1)}(0) = 0, \ldots, \omega^{(l-1)}(0) = 0.$$

Assumption 2.1. All roots of the characteristic equation

$$\lambda^l + a_1 \lambda^{l-1} + \cdots + a_l = 0 \tag{7}$$

have non-positive real parts.

Assumption 2.2. $\varphi_{l-1}(t) \geqslant 0$ for all $t \geqslant 0$.

Note that Assumption 2.2 holds if equation (7) has real roots only. Let's designate by $\lambda_1, \ldots, \lambda_s$ ($\lambda_1 < \lambda_2 < \cdots < \lambda_s$) the real roots and by $\mu_1 \pm$

$i\nu_1, \dots, \mu_p \pm i\nu_p$ ($\mu_1 \leqslant \mu_2 \leqslant \dots \leqslant \mu_p$) the complex roots of equation (7). Let k_s be the order of the root λ_s.

Further denote

$$\zeta_i(T,t) = \varphi_0(T)x_i(t) + \dots + \varphi_{l-1}(T)x_i^{(l-1)}(t),$$

$$\eta_j(T,t) = \varphi_0(T)y_j(t) + \dots + \varphi_{l-1}(T)y_j^{(l-1)}(t),$$

$$\xi_{i,j}(T,t) = \varphi_0(T)z_{i,j}(t) + \dots + \varphi_{l-1}(T)z_{i,j}^{(l-1)}(t).$$

Since

$$\varphi_q(t) = \sum_{\beta=1}^{s} e^{\lambda_\beta t} P_{q,\beta}(t) + \sum_{\alpha=1}^{p} e^{\mu_\alpha t}(Q_{q,\alpha}(t)\cos(\nu_\alpha t) + R_{q,\alpha}(t)\sin(\nu_\alpha t)),$$

$\zeta_i(T,0)$, $\eta_j(T,0)$, $\xi_{i,j}(T,0)$ can be represented by

$$\zeta_i(T,0) = \sum_{\beta=1}^{s} e^{\lambda_\beta T} P^1_{i,\beta}(T) + \sum_{\alpha=1}^{p} e^{\mu_\alpha T}(Q^1_{i,\alpha}(T)\cos(\nu_\alpha T) + R^1_{i,\alpha}(T)\sin(\nu_\alpha T)),$$

and

$$\eta_j(T,0) = \sum_{\beta=1}^{s} e^{\lambda_\beta T} P^2_{j,\beta}(T) + \sum_{\alpha=1}^{p} e^{\mu_\alpha T}(Q^2_{j,\alpha}(T)\cos(\nu_\alpha T) + R^2_{j,\alpha}(T)\sin(\nu_\alpha T)),$$

$$\xi_{i,j}(T,0) = \sum_{\beta=1}^{s} e^{\lambda_\beta T} P^3_{i,j,\beta}(T) + \sum_{\alpha=1}^{p} e^{\mu_\alpha T}(Q^3_{i,j,\alpha}(T)\cos(\nu_\alpha T) + R^3_{i,j,\alpha}(T)\sin(\nu_\alpha T)).$$

We assume that $\xi_{i,j}(T,0) \neq 0$ for all i, j and $T > 0$, because if $\xi_{p,q}(T,0) = 0$ for some p, q and T, then pursuer P_p catches the evader E_q at time T. Also we assume that $P^3_{i,j,s}(t)$ is not equal to 0 for all i and j.

Let's designate by $\gamma_{i,j}$ the order of polynomial $P^3_{i,j,s}(t)$, by γ the order of polynomial $P_{l-1,s}$. We assume that $\gamma_{i,j} = \gamma$ for all i and j.

Let

$$X_i^0 = \lim_{t\to\infty} \frac{P^1_{i,s}}{t^\gamma}, \quad Y_j^0 = \lim_{t\to\infty} \frac{P^2_{j,s}}{t^\gamma}, \quad Z^0_{i,j} = \lim_{t\to\infty} \frac{P^3_{i,j,s}}{t^\gamma},$$

$$C_{\alpha,\beta}(T,t) = \eta_\alpha(T,t) - \eta_\beta(T,t) = C_{\alpha,\beta}(T+t,0).$$

We first state the following obvious result.

Lemma 2.1. *Let $Y(t,0)$ be the fundamental matrix of Equation (6), such that $Y(0,0) = E$. Then*

$$Y(t,0)Y(T,0) = Y(t+T,0). \tag{8}$$

Now we can establish the following.

Lemma 2.2. *For roots $\varphi_0(t),\dots,\varphi_{l-1}(t)$ of Equation (6), we have*

$$\varphi_0(T)\varphi_\alpha(t)+\dots+\varphi_{l-1}(T)\varphi_\alpha^{(l-1)}(t)=\varphi_\alpha(T+t),\quad \alpha\in\{0,\dots,l-1\}.$$

Proof. Let

$$Y(T,0)=\begin{pmatrix}\varphi_0(T) & \dots & \varphi_{l-1}(T)\\ \dots & \dots & \dots\\ \varphi_0^{(l-1)}(T) & \dots & \varphi_{l-1}^{(l-1)}(T)\end{pmatrix}$$

be the fundamental matrix of Equation (6), normalized at the point 0. Then

$$Y(T,0)\begin{pmatrix}\varphi_\alpha(t)\\ \dots\\ \varphi_\alpha^{(l-1)}(t)\end{pmatrix}=Y(T,0)Y(t,0)\begin{pmatrix}\varphi_\alpha(0)\\ \dots\\ \varphi_\alpha^{(l-1)}(0)\end{pmatrix}.$$

If we use Lemma 2.1 and replace the terms $Y(T,0)Y(t,0)$ by $Y(T+t,0)$, then

$$Y(T+t,0)\begin{pmatrix}\varphi_\alpha(0)\\ \dots\\ \varphi_\alpha^{(l-1)}(0)\end{pmatrix}=\begin{pmatrix}\varphi_\alpha(T+t)\\ \dots\\ \varphi_\alpha^{(l-1)}(T+t)\end{pmatrix}=Y(T,0)\begin{pmatrix}\varphi_\alpha(t)\\ \dots\\ \varphi_\alpha^{(l-1)}(t)\end{pmatrix}.$$

The lemma follows from this last equality. □

Lemma 2.3. *The function*

$$\chi(T,t)=\varphi_0(T)\nu(t)+\dots+\varphi_{l-1}(T)\nu^{(l-1)}(t),$$

where

$$\nu(t)=\chi(t,0)+\int_0^t\varphi_{l-1}(t-\tau)f(\tau)d\tau$$

is the root of the equation

$$\nu^{(l)}(t)+a_1\nu^{(l-1)}(t)+\dots+a_l\nu(t)=f(t),$$

can be represented by

$$\chi(T,t)=\chi(T+t,0)+\int_0^t\varphi_{l-1}(T+t-\tau)f(\tau)d\tau.$$

Proof. Let's note that for all $s=1,\dots,l-1$ the function ν satisfies the condition

$$\nu^{(s)}(t)=\chi^{(s)}(t,0)+\int_0^t\varphi_{l-1}^{(s)}(t-\tau)f(\tau)d\tau.$$

Therefore, the function $\chi(T,t)$ can be represented as

$$\begin{aligned}\chi(T,t) &= \varphi_0(T)(\chi(t,0) + \int\limits_0^t \varphi_{l-1}(t-\tau)f(\tau)d\tau) + \varphi_1(T)(\dot\chi(t,0)\\ &\quad + \int\limits_0^t \dot\varphi_{l-1}(t-\tau)f(\tau)d\tau) + \cdots + \varphi_{l-1}(T)(\chi^{(l-1)}(t,0)\\ &\quad + \int\limits_0^t \varphi_{l-1}^{(l-1)}(t-\tau)f(\tau)d\tau)\\ &= \varphi_0(T)\chi(t,0) + \varphi_1(T)\dot\chi(t,0) + \cdots + \varphi_{l-1}(T)\chi^{(l-1)}(t,0)\\ &\quad + \int\limits_0^t (\varphi_0(T)\varphi_{l-1}(t-\tau) + \cdots + \varphi_{l-1}(T)\varphi_{l-1}^{(l-1)}(t-\tau))f(\tau)d\tau.\end{aligned}$$

By Lemma 2.2

$$\begin{aligned}&\varphi_{l-1}(T+t-\tau)\\ &\quad = \varphi_0(T)\varphi_{l-1}(t-\tau) + \varphi_1(T)\dot\varphi_{l-1}(t-\tau) + \cdots + \varphi_{l-1}(T)\varphi_{l-1}^{(l-1)}(t-\tau).\end{aligned}$$

Since

$$\chi(t,0) = \varphi_0(t)\chi(0) + \varphi_1(t)\dot\chi(0) + \cdots + \varphi_{l-1}(t)\chi^{(l-1)}(0),$$

we have

$$\chi^{(s)}(t,0) = \varphi_0^{(s)}(t)\chi(0) + \varphi_1^{(s)}(t)\dot\chi(0) + \cdots + \varphi_{l-1}^{(s)}(t)\chi^{(l-1)}(0).$$

Therefore,

$$\begin{aligned}&\varphi_0(T)\chi(t,0) + \varphi_1(T)\dot\chi(t,0) + \cdots + \varphi_{l-1}(T)\chi^{(l-1)}(t,0)\\ &\quad = \varphi_0(T)(\varphi_0(t)\nu(0) + \cdots + \varphi_{l-1}(t)\nu^{(l-1)}(0))\\ &\qquad + \varphi_1(T)(\dot\varphi_0(t)\nu(0) + \cdots + \dot\varphi_{l-1}(t)\nu^{(l-1)}(0))\\ &\qquad + \cdots + \varphi_{l-1}(T)(\varphi_0^{(l-1)}(t)\nu(0) + \cdots + \varphi_{l-1}^{(l-1)}(t)\nu^{(l-1)}(0)).\end{aligned}$$

Thus

$$\begin{aligned}\chi(T,t) &= \varphi_0(T+t)\nu(0) + \cdots + \varphi_{l-1}(T+t)\nu^{(l-1)}(0) + \int\limits_0^t \varphi_{l-1}(T+t-\tau)f(\tau)d\tau\\ &= \chi(T+t,0) + \int\limits_0^t \varphi_{l-1}(T+t-\tau)f(\tau)d\tau.\end{aligned}$$

$\square$

Let $V = \{v : \|v\| \leqslant 1\}$, $\lambda(a, v) = \sup\{\lambda \geqslant 0 : -\lambda a \in V - v\}$.

Lemma 2.4. *Let Assumptions 2.1, 2.2 hold, with $b_i(t) : [0, \infty) \to R^k$, $i = 1, \ldots, p$,*

$$\inf_{v \in V} \max_i \lambda(b_i(T^0 + t), v) > \delta > 0$$

for all $t > 0$. Then there exists a time $T > 0$, such that for any admissible function $v(\cdot)$ there exists a number q, such that

$$1 - e^{-\lambda_s(T^0+T)} \int_0^T \varphi_{l-1}(T^0 + T - \tau)\lambda(b_q(T^0 + T), v(\tau))d\tau \leqslant 0.$$

Proof. Let T be given. We define the functions h_i as

$$h_i(t) = 1 - e^{-\lambda_s(T^0+t)} \int_0^t \varphi_{l-1}(T^0 + T - \tau)\lambda(b_i(T^0 + T), v(\tau))d\tau.$$

$h_i(0) = 1$, h_i is continuous and

$$\sum_i h_i(T) \leqslant p - e^{-\lambda_s(T^0+T)} \int_0^T \varphi_{l-1}(T^0 + T - \tau) \max_i \lambda(b_i(T^0 + T), v(\tau))d\tau.$$

Since $\inf\limits_{v \in V} \max\limits_i \lambda(b_i(T^0 + t), v) > \delta$, we have

$$\sum_i h_i(T) \leqslant p - \delta e^{-\lambda_s(T^0+T)} \int_0^T \varphi_{l-1}(T^0 + T - \tau)d\tau = g(T).$$

By Assumptions 2.1, 2.2 and by Lemma 2.2 of [3, p. 154],

$$\lim_{T \to \infty} g(T) = -\infty.$$

Therefore, there exists a time T that satisfies the condition of the lemma. $\square$

Definition 2.2. We say that the vectors $a_1, \ldots, a_s$ form a positive basis for R^k, if for any $x \in R^k$ there exist positive real numbers $\alpha_1, \ldots, \alpha_s$, such that

$$x = \alpha_1 a_1 + \cdots + \alpha_s a_s.$$

Lemma 2.5. *Let $a_1, \ldots, a_s$ be a positive basis. Then for any b_l, $l = 1, \ldots, s$, there exists $\mu_0 > 0$, such that for all $\mu > \mu_0$*

$$a_1, \ldots, a_{l-1}, b_l + \mu a_l, \ldots, b_s + \mu a_s$$

is a positive basis.

Proof. Assume this is not true. Then there exist $b_l, \dots, b_s$ and a sequence $\mu_n \to \infty$, such that

$$a_1, \dots, a_{l-1}, b_l + \mu_n a_l, \dots, b_s + \mu_n a_s$$

is not a positive basis. Therefore, there exist $\{v_n\}$, $\|v_n\| = 1$, such that for all n

$$(a_r, v_n) \leqslant 0, \quad r = 1, \dots, l-1, \tag{9}$$

$$(a_r + \frac{1}{\mu_n} b_r, v_n) \leqslant 0, \quad r = l, \dots, s. \tag{10}$$

By the compactness of the unit sphere, we can assume without loss of generality that the $v_n \to v_0$, $\|v_0\| = 1$. Going to the limit in (9), (10) we obtain $(a_r, v_0) \leqslant 0$, which contradicts the positivity of the basis. This establishes the lemma. □

Corollary 2.1. *Let $a_1, \dots, a_s$ be a positive basis. Then for any $b_l, \dots, b_s$ $(1 \leqslant l \leqslant s)$ there exists $\mu_0 > 0$, such that for all $\mu > \mu_0$ the systems of vectors*

$$\{a_1, \dots, a_{l-1}, b_l + \mu a_l, \dots, b_s + \mu a_s\},$$

$$\{a_1, \dots, a_{l-1}, a_{\alpha_0}, b_r + \mu a_r, r \neq \alpha_0, r = l, \dots, s\}$$

form positive bases.

Lemma 2.6. *Let $x_i, y_j \in R^k$ and assume*
(a) $n + m \geqslant k + 2$,
(b) *in the set $\{x_i - y_j, y_r - y_q, r \neq q, x_s - x_l, s \neq l\}$ there exist k linearly independent vectors.*
Then

$$\operatorname{rico}\{x_i\} \cap \operatorname{rico}\{y_j\} \neq \emptyset, \tag{11}$$

if and only if $\{x_i - y_j\}$ form a positive basis.

Proof. Assume that (11) holds while the vectors $\{x_i - y_j\}$ do not form a positive basis. Then (see [16]) there exists $p \in R^k$, $\|p\| = 1$, such that $(x_i - y_j, p) \leqslant 0$ for all i, j. Hence the sets $\operatorname{co}\{x_i\}$, $\operatorname{co}\{y_j\}$ can be separated by a hyperplane.

If this were not the case, there would exist a hyperplane H, such that

$$\operatorname{co}\{x_i\} \subset H, \quad \operatorname{co}\{y_j\} \subset H.$$

Hence $x_i - y_j \in L$, $y_r - y_q \in L$, $x_s - x_l \in L$. (Here L is the linear subspace, corresponding to H.) This contradicts the conditions of the lemma.

On the other side, let

$$\operatorname{rico}\{x_i\} \cap \operatorname{rico}\{y_j\} = \emptyset.$$

Then the sets $\operatorname{co}\{x_i\}$, $\operatorname{co}\{y_j\}$ are separable. Therefore, there exists a unit vector p, such that $(x_i - y_j, p) \leq 0$. Hence $\{x_i - y_j\}$ ([16]) do not form a positive basis. □

Lemma 2.7. *Let $x_i, y_j \in R^k$, $n+m \geqslant k+2$, be any $k+1$ affinely independent points such that*

$$\mathrm{co}\{x_1, \dots, x_n\} \cap \mathrm{co}\{y_1, \dots, y_m\} \neq \emptyset.$$

Then there exist sets $I \subset \{1, \dots, n\}$, $J \subset \{1, \dots, m\}$, such that $|I|+|J| = k+2$, and the intersection

$$\mathrm{rico}\{x_\alpha, \alpha \in I\} \cap \mathrm{rico}\{y_\beta, \beta \in J\} \neq \emptyset$$

consists of one point only.

Proof. Let $I = \{1, \dots, n\}$, $J = \{1, \dots, m\}$, $M_1 = \mathrm{co}\{x_i, i \in I\}$, $M_2 = \mathrm{co}\{y_j, j \in J\}$, $x \in M_1 \cap M_2$. Then

$$x = \sum_{i \in I} \beta_i x_i = \sum_{j \in J} \lambda_j y_j, \ \sum_{i \in I} \beta_i = \sum_{j \in J} \lambda_j = 1, \ \beta_i \geq 0, \ \lambda_j \geq 0.$$

Define the sets $I_1 = \{i \in I, \beta_i \neq 0\}$, $J_1 = \{j \in J, \lambda_j \neq 0\}$. Then $\mathrm{co}\{x_i, i \in I_1\} \cap \mathrm{co}\{y_j, j \in J_1\} \neq \emptyset$. Let x belong to this set. Now take as sets I, J the sets I_1, J_1 and repeat this procedure until one obtains the sets I, J, such that for any $x \in M_1 \cap M_2$ one will have $\beta_i > 0$, $\lambda_j > 0$ for all $i \in I$, $j \in J$. Then $M_1 \cap M_2 = \mathrm{ri} M_1 \cap \mathrm{ri} M_2$. It has been shown in [17] that this implies that $M_1 \cap M_2 = \mathrm{aff} M_1 \cap \mathrm{aff} M_2$ and the set $M_1 \cap M_2$ consists of one point only. It follows that $\dim M_1 + \dim M_2 = k$. As $\dim M_1 = |I| - 1$, $\dim M_2 = |J| - 1$, we have $|I| + |J| = k + 2$ and the lemma is proved. □

Lemma 2.8. *Let $\{x_1, \dots, x_n, y_1, \dots, y_m\}$ be a set of points from R^k, $n+m = k+2$, any k vectors from the set $\{x_i - y_j,\ y_r - y_s,\ r \neq s\}$ that are linearly independent and satisfy*

$$\mathrm{rico}\{x_i\} \cap \mathrm{rico}\{y_j\} \neq \emptyset, \quad x_{n+1} - y_{\beta_0} = \mu(y_{\beta_0} - y_1) = -\mu(y_1 - y_{\beta_0}).$$

Here $\mu > 0$, $\beta_0 \in \{2, \dots, m\}$.

Then

$$\mathrm{rico}\{x_i, i = 1, \dots, n+1\} \cap \mathrm{rico}\{y_j, j \in 2, \dots, m\} \neq \emptyset.$$

Proof. By Lemma 2.6 the vectors $\{x_i - y_j\}$ compose a positive basis. Therefore, the vectors

$$\{x_i - y_{\beta_0} + y_{\beta_0} - y_1, x_i - y_j, i = 1, \dots, n, j = 2, \dots m\}$$

compose a positive basis. From this it follows that a positive basis is composed of the vectors

$$\{y_{\beta_0} - y_1, x_i - y_j, i = 1, \dots, n, j = 2, \dots m\}.$$

Replacing $y_{\beta_0} - y_1$ by the vectors $x_{n+1} - y_{\beta_0}$, we get that a positive basis is also composed of the vectors $\{x_i - y_j,\ i = 1, \dots, n+1, j = 2, \dots, m\}$. By Lemma 2.5

$$\mathrm{rico}\{x_i, i = 1, \dots, n+1\} \cap \mathrm{rico}\{y_j, j \in 2, \dots, m\} \neq \emptyset.$$

This establishes the result. □

Let's suppose that the vectors X_i^0, Y_j^0 are such that

(a) if $n > k$, then for any set of indices $I \subset \{1, \dots, n\}$, $\|I\| \geq k+1$

$$\mathrm{Intco}\{X_i^0, i \in I\} \neq \emptyset;$$

(b) any k vectors from the sets $\{X_i^0 - Y_j^0,\ Y_s^0 - Y_t^0,\ s \neq r\}$ are linearly independent.

Theorem 2.1. *Let $n \geqslant k+1$ and*

$$0 \in \mathrm{Intco}\{Z_{i,j}^0\}.$$

Then there is a capture in the game Γ.

Proof. Since $n \geqslant k$, $n+m \geqslant k+2$. By Lemma 2.7 there exist $I \subset \{1, \dots, n\}$ and $J \subset \{1, \dots, m\}$, such that $\{Z_{i,j}^0,\ i \in I,\ j \in J\}$ form a positive basis and $|I|+|J| = k+2$. We will suppose, without loss of generality, that $I = \{1, \dots, q\}$, $J = \{1, \dots, l\}$, $q+l = k+2$.

If $|J| = 1$, then capture follows from [3, §2]. We will assume that $|J| \geqslant 2$.

By Lemma 2.4 of [3, p. 155] it follows that there exists a time T^0, such that

$$\{\xi_{i,j}(T^0+t, 0), i \in I, j \in J\} \tag{12}$$

form a positive basis for all $t \geqslant 0$. For all $i \in I$ and for all $\alpha \neq \alpha_0, \alpha \in J$,

$$\xi_{i,\alpha}(T^0, t) = \xi_{i,\alpha_0}(T^0, t) + C_{\alpha_0,\alpha}(T^0, t).$$

Then

$$\{\xi_{i,\alpha_0}(T^0+t, 0), i \in I, C_{\alpha_0,\alpha}(T^0+t, 0), \alpha \neq \alpha_0, \alpha \in J\}$$

is a positive basis.

Suppose that $\alpha_0 = 1$. Let $J_1 = J \setminus \{1\}$. Then

$$\{\xi_{i,1}(T^0+t, 0), i \in I, C_{1,\alpha}(T^0+t, 0), \alpha \in J_1\}$$

is a positive basis, and the number of vectors in a given set is $|I|+|J|-1 = k+1$.

According to our assumptions $n \geqslant k+1$, therefore, $q+\alpha-1 \in \{q+1, \dots, k+1\}$, at $\alpha \in J_1$.

By Corollary 2.1 of Lemma 2.5 it follows that

$$\{\xi_{i,1}(T^0+t, 0), i \in I, \xi_{q+\alpha-1,1}(T^0+t, 0) + \mu C_{1,\alpha}(T^0+t, 0), \alpha \in J_1\}$$

is a positive basis.

Designate by β_r

$$b_i(t) = \xi_{i,1}(T^0+t, 0), \quad i \in I,$$
$$b_{q+\alpha-1}(t) = \xi_{q+\alpha-1}(T^0+t, 0) + \mu C_{1,\alpha}(T^0+t, 0), \quad \alpha \in J_1,$$

and let

$$T_0 = \min\left\{ t \,\middle|\, t \geqslant 0, \inf_{v(\cdot)} \max_r \; e^{-\lambda_s(T^0+t)} \int_0^t \varphi_{l-1}(T^0+t-\tau)\lambda(\beta_r(T^0+t), v(\tau))d\tau \geqslant 1 \right\}.$$

Then by Lemma 2.4 $T_0 < \infty$.

Let $v : [0, T_0] \to V$ be an admissible control of evaders and let t_1 be the smallest positive root of function h, such that

$$h(t) = 1 - \max_r \int_0^t e^{-\lambda_s(T^0+T_0)} \lambda_r^0(T_0, v(\tau))d\tau.$$

Here

$$\lambda_r^0(T_0, v) = \lambda(e^{-\lambda_s(T^0+T_0)} b_r(T_0), v), \quad r \in I,$$
$$\lambda_{q+\alpha-1}^0(T_0, v) = \lambda(e^{-\lambda_s(T^0+T_0)} b_{q+\alpha-1}(T_0), v), \quad \alpha \in J_1.$$

Suppose that $\lambda(e^{-\lambda_s(T^0+T_0)} b_r(T_0), v(t)) = 0$ at $t \in [t_1, T_0]$. Define the controls of pursuers as

$$u_i(t) = v(t) - \lambda(e^{-\lambda_s(T^0+T_0)} b_i(T_0), v(t)) e^{-\lambda_s(T^0+T_0)} b_i(T_0), \quad i \in I,$$
$$u_{q+\alpha-1}(t) = v(t) - \lambda(e^{-\lambda_s(T^0+T_0)} b_{q+\alpha-1}(T^0+T_0), v(t)) e^{-\lambda_s(T^0+T_0)} b_{q+\alpha-1}(T_0),$$
$$\alpha \in J_1.$$

Then

$$\xi_{i,1}(T^0, t) = \xi_{i,1}(T^0+t, 0) + \xi_{i,1}(T^0+T_0, 0)(h_i(t)-1),$$
$$i \in I$$
$$\xi_{q+\alpha-1,1}(T^0, t) + \mu C_{1,\alpha}(T^0, t) = b_{q+\alpha-1,1}(t) + b_{q+\alpha-1}(T_0)(h_{q+\alpha-1}(t)-1),$$
$$\alpha \in J_1.$$

Here

$$h_i(t) = 1 - e^{-\lambda_s(T^0+T_0)} \int_0^t \varphi_{l-1}(T^0+T_0-\tau)\lambda(e^{-\lambda_s(T^0+T_0)} b_i(T_0), v(\tau))d\tau,$$
$$i \in I,$$
$$h_{q+\alpha-1}(t) = 1 - e^{-\lambda_s(T^0+T_0)}$$

$$\cdot \int_0^t \varphi_{l-1}(T^0+T_0-\tau)\lambda(e^{-\lambda_s(T^0+T_0)}b_{q+\alpha-1}(T_0), v(\tau))d\tau, \quad \alpha \in J_1.$$

By Lemma 2.4 it follows that there exists an r, such that $h_r(T_0)=0$. If $r \in I$, we suppose that $u_i(t)=v(t)$, $t \in [T_0, T^0+T_0]$. Then

$$\begin{aligned} z_r(T^0+T_0) &= \xi_r(T^0+T_0,0) - \int_0^{T^0+T_0} \varphi_{l-1}(T^0+T_0-\tau)(u_r(t)-v(t))dt \\ &= \xi_r(T^0+T_0,0)(1-e^{-\lambda_s(T^0+T_0)} \\ &\quad \cdot \int_0^{t_1} \varphi_{l-1}(T^0+T_0-\tau)\lambda(e^{-\lambda_s(T^0+T_0)}b_r(T_0), v(\tau))d\tau) = 0. \end{aligned}$$

Therefore, there is a capture in the game Γ.

If $h_{q+\alpha_0-1}(T_0)=0$ at some $\alpha_0 \in J_1$, then

$$\xi_{q+\alpha_0-1,1}(T^0,T_0) = -\mu C_{1\alpha_0}(T^0,T_0) = -\mu C_{1,\alpha_0}(T^0+T_0,0).$$

Let's show that the vectors $\{\xi_{i,j}(T^0,T_0),\ i \in I,\ j \in J\}$ form a positive basis.

The selected time T^0 is such that $\{\xi_{i,j}(T^0+t,0), i \in I, j \in J\}$ is a positive basis for all $t>0$. One has $\xi_{i,1}(T^0,T_0)=\xi_{i,1}(T^0+T_0,0)h_i(T_0)$. Therefore,

$$\{\xi_{i,1}(T^0,T_0), \xi_i(T^0+T_0,0)+C_{1,j}(T^0+T_0,0), i \in I, j \in J_1\}$$

forms a positive basis. Therefore, the positive basis contains the vectors

$$\left\{\xi_{i1}(T^0,T_0), \frac{\xi_{i1}(T^0,T_0)}{h_i(T_0)} + C_{1,j}(T^0+T_0,0), i \in I, j \in J_1\right\}.$$

Since

$$h_i(T_0)<1, \quad C_{1j}(T^0+T_0,0)=\xi_{ij}(T^0,T_0)-\xi_{i1}(T^0,T_0),$$

the positive basis contains the vectors

$$\left\{\xi_{i1}(T^0,T_0), \xi_{ij}(T^0,T_0) + \frac{1-h_i(T_0)}{h_i(T_0)}\xi_{i1}(T^0,T_0)\right\}.$$

Therefore, $\{\xi_{i,j}(T^0,T_0),\ i \in I,\ j \in J\}$ forms a positive basis. Then

$$\begin{aligned} &\{\zeta_i(T^0,T_0)-\eta_{\alpha_0}(T^0,T_0)+\eta_{\alpha_0}(T^0,T_0)-\eta_1(T^0,T_0), \\ &\ \zeta_i(T^0,T_0)-\eta_j(T^0,T_0), i \in I, j \in J_1\} \end{aligned}$$

forms a positive basis.

Hence

$$\{\xi_{i,j}(T^0, T_0), i \in I, j \in J_1, -C_{1,\alpha_0}(T^0, T_0)\}$$

forms a positive basis. If we put $\xi_{q+\alpha_0-1,1}(T^0, T_0)$ instead of $-C_{1,\alpha_0}$ (T^0, T_0), then

$$\{\xi_{i,j}(T^0, T_0), i \in I \cup \{q + \alpha_0 - 1\}, j \in J_1\}$$

forms a positive basis.

Let $\alpha_0 = 2$, set $q = q + 1$, and renumber $\{2, \dots, l\}$ as $J = \{1, \dots, l\}$, at $l = l - 1$.

So, we now have an expression similar to (12), but the number of evaders $|J|$ was decreased by one. Let's then assume the time T_0 as the initial time and repeat our reasoning (starting from (12)) until the number of evaders $|J|$ in expression (12) will be equal to one. So now have an an expression similar to (12) with $|J|$ equal to 1. Now we can apply the theorem of Pontrjagin (Theorem 2.1 [3, p. 113]) and the theorem is proved. $\square$

3 "Soft" Capture in Pontrjagin's Problem with Many Participants

In the space $R^k (k \geq 2)$ we consider an $n+1$-person differential game Γ: n pursuers $P_1,\ P_2, \dots, P_n$ and the evader E. The law of motion of each pursuer P_i is defined by (1).

The law of motion of the evader E is defined by (2).

Here $x_i, y_j, u_i, v \in R^k, a_1, \dots, a_l \in R^1, V$ is compact. The initial conditions at $t = 0$ are

$$x_i^{(\alpha)}(0) = x_{i\alpha}^0, \quad y^{(\alpha)}(0) = y_\alpha^0, \quad \alpha = 0, \dots, l-1.$$

Here $x_{i0}^0 \neq y_0^0,\ x_{i1}^0 \neq y_1^0$. Everywhere we will assume that index i takes the values $i = 1, 2, \dots, n$.

Definition 3.1. In the game Γ there is a "soft" capture if there exist $T > 0$ and measurable functions $u_i(t) = u_i(t, x_{i\alpha}^0, y_\alpha^0, v_t(\cdot)) \in V$, such that for any measurable function $v(\cdot), v(t) \in V, t \in [0, T]$ there exist $\tau \in [0, T]$ and a number $q \in \{1, 2, \dots, n\}$, such that

$$x_q(\tau) = y(\tau), \quad \dot{x}_q(\tau) = \dot{y}(\tau).$$

Instead of systems (1)–(3) we consider a system

$$z_i^{(l)} + a_1 z_i^{(l-1)} + \dots + a_l z_i = u_i - v, \quad u_i, v \in V \tag{13}$$

$$z_i(0) = z_{i0}^0 = x_{i0}^0 - y_0^0, \dots, z_i^{(l-1)}(0) = z_{il-1}^0 = x_{il-1}^0 - y_{l-1}^0. \tag{14}$$

Assumption 3.1. All roots of the characteristic equation (7) are real and non-positive.

Let's designate the roots of Equation (7) as $\lambda_1 < \cdots < \lambda_s$, and their orders as $k_1, \ldots, k_s$.

Lemma 3.1. *Let Assumption 3.1 hold with* $\lambda_s = 0$*. Then* $\varphi_{l-1}(t) \geq 0$, $\dot{\varphi}_{l-1}(t) \geq 0$ *for all* $t \geq 0$.

Lemma 3.2. *Let Assumption 3.1 hold and let* $\lambda_s < 0$*. Then*
(1) $\varphi_{l-1}(t) \geq 0$ *for all* $t > 0$;
(2) *there exist* $\tau_0 > 0$*, such that* $\dot{\varphi}_{l-1}(t) > 0$, $t \in (0, \tau_0)$, $\dot{\varphi}_{l-1}(t) < 0$, $t \in (\tau_0, \infty)$.

The assertions of Lemmas 3.1, 3.2 follow from the well-known results in ([18], p. 136).

Let

$$\xi_i(t) = \sum_{k=0}^{l-1} \varphi_k(t) z_{ik}^0.$$

Then ξ_i can be written as

$$\xi_i(t) = \sum_{j=1}^{l-1} e^{\lambda_j t} P_{ji}(t).$$

Here P_{ji} are polynomials. Let's suppose that the $\deg P_{si} = k_s - 1 = \gamma$ for all i, for otherwise pursuers obtain the given conditions, first by selecting their controls as $u_i(t)$ on a sufficiently small time interval such that the coefficients for t^γ in polynomials P_{si} are not equal to zero.

Let's consider further

$$\lambda_s = 0, \quad k_s \geq 2 \tag{15}$$

and designate

$$M(t, \tau) = \min\left\{ \frac{\varphi_{l-1}(t-\tau)}{t^\gamma}, \ \frac{\dot{\varphi}_{l-1}(t-\tau)}{\gamma t^{\gamma-1}} \right\},$$

$$R(f, t, \tau) = \sum_{j=1}^{s-1} \frac{e^{\lambda_j (t-\tau)} f_j(t-\tau)}{\gamma (t-\tau)^{\gamma-1}}.$$

Lemma 3.3. *Suppose that Assumption 3.1 and condition (15) are satisfied. Then for any* $T > 0$

$$\lim_{t \to \infty} \int_T^t M(t, \tau) d\tau = \infty.$$

Proof. The functions $\varphi_{l-1},\ \dot\varphi_{l-1}$ can be represented as

$$\varphi_{l-1}(t-\tau)=a_\gamma(t-\tau)^\gamma[1+g_1(t-\tau)],$$
$$\dot\varphi_{l-1}(t-\tau)=a_\gamma\gamma(t-\tau)^{\gamma-1}[1+g_2(t-\tau)].$$

Here

$$g_1(t-\tau)=\sum_{l=0}^{\gamma-1}\frac{a_l}{(t-\tau)^{\gamma-l}}+R(P,t,\tau), g_2(t-\tau)=\sum_{l=1}^{\gamma-1}\frac{b_l}{(t-\tau)^{\gamma-l}}+R(Q,t,\tau).$$

Let $\varepsilon\in(0,1),\ \tau\in[0,\varepsilon t]$. Then $t-\tau\geq(1-\varepsilon)t$ and

$$|g_1(t-\tau)|\leq\sum_{r=1}^{\gamma}\frac{|a_{\gamma-r}|}{t^r(1-\varepsilon)^r}+\Sigma^1(t)=\Delta_1(t),$$
$$|g_2(t-\tau)|\leq\sum_{r=1}^{\gamma-1}\frac{|b_{\gamma-r}|}{t^r(1-\varepsilon)^r}+\Sigma^2(t)=\Delta_2(t).$$

Here

$$\Sigma^k(t)=\sum_{j=1}^{s-1}e^{\lambda_j(1-\varepsilon)t}c_j^k(t),$$

where

$$c_j^1(t)=\frac{\max\limits_{\tau\in[0,\varepsilon t]}|P_j(t-\tau)|}{t^\gamma(1-\varepsilon)^\gamma},\quad c_j^2(t)=\frac{\max\limits_{\tau\in[0,\varepsilon t]}|Q_j(t-\tau)|}{\gamma t^{\gamma-1}(1-\varepsilon)^{\gamma-1}}.$$

Since $\Delta_1(t),\ \Delta_2(t)\to 0$ at $t\to\infty$, there exists T_0, such that $|\Delta_1(t)|\leq 1/2$, $|\Delta_2(t)|\leq 1/2$ for all $t>T_0$. Therefore,

$$\varphi_{l-1}(t-\tau)\geq 1/2a_\gamma(t-\tau)^\gamma,\quad \dot\varphi_{l-1}(t-\tau)\geq 1/2a_\gamma\gamma(t-\tau)^{\gamma-1}$$

for all $t>T_0,\ \tau\in[0,\varepsilon t]$.

Hence for all (t,T), such that $T>T_0,\ \varepsilon t>T$,

$$\int_T^t M(t,\tau)d\tau\geq\int_T^{\varepsilon t}M(t,\tau)d\tau\geq\int_T^{\varepsilon t}\frac{a_\gamma}{2}\frac{(t-\tau)^\gamma}{t^\gamma}d\tau\to\infty\quad\text{at } t\to\infty.$$

□

Let

$$z_i^0=\lim_{t\to\infty}P_{si}(t)/t^\gamma,$$
$$\lambda(A,v)=\sup\{\lambda\mid\lambda\geq 0,\ -\lambda A\cap(V-v)\neq\emptyset\},$$
$$\delta=\inf_{v\in V}\max_i\lambda(z_i^0,v)>0.$$

Assumption 3.2. The functions $\lambda(z_i^0, v)$ are continuous at (z_i^0, v) whenever $\lambda(z_i^0, v) > 0$.

Lemma 3.4. *Let Assumptions 3.1, 3.2 and condition (15) hold, with $\delta > 0$. Then there exists a time T, such that for any admissible function v there exists an index i, such that $h_i(T) \leq 0$. Here*

$$h_i(t) = 1 - \int_0^T \beta_i(T, \tau, v(\tau))\ d\tau \leq 0,\ \ \mathcal{L}(f(r), t) = \left\| \begin{array}{c} f(r)/t^\gamma \\ \dot{f}(r)/(\gamma t^{\gamma-1}) \end{array} \right\|,$$

$$\beta_i(t, \tau, v) = \sup\Big\{\lambda | \lambda \geq 0, -\lambda \mathcal{L}(\xi_i(t), t) \in \mathcal{L}(\varphi_{l-1}(t-\tau), t)(V - v)\Big\}.$$

Proof. Let's note that

$$\beta_i(t, \tau, v) = \min\Big\{\frac{\varphi_{l-1}(t-\tau)}{t^\gamma}\lambda(\frac{\xi_i(t)}{t^\gamma}, v),\ \ \frac{\dot{\varphi}_{l-1}(t-\tau)}{\gamma t^{\gamma-1}}\lambda(\frac{\dot{\xi}_i(t)}{\gamma t^{\gamma-1}}, v)\Big\}.$$

Since $z_i^0 = \lim\limits_{t\to\infty} \frac{\xi_i(t)}{t^\gamma} = \lim\limits_{t\to\infty} \frac{\dot{\xi}_i(t)}{\gamma t^{\gamma-1}}$, there exists a time T_0, such that

$$\max_i \lambda(\frac{\xi_i(t)}{t^\gamma}, v) \geq 1/2\delta, \quad \max_i \lambda(\frac{\dot{\xi}_i(t)}{\gamma t^{\gamma-1}}, v) \geq 1/2\delta$$

for all $t > T_0,\ v \in V$. Let's consider continuous functions h_i, such that

$$h_i(0) = 1,\ \sum_i h_i(T) \leq n - \int_0^T \max_i \beta_i(T, \tau, v(\tau))d\tau.$$

Let $T > T_0$. Then

$$\max_i \beta_i(T, \tau, v(\tau)) \geq 1/2\delta M(T, \tau).$$

Therefore,

$$\int_0^T \max_i \beta_i(T, \tau, v(\tau))d\tau \geq 1/2\delta \int_{T_0}^T M(T, \tau)\ d\tau = g(T).$$

Therefore, $\sum\limits_i h_i(T) \leq n - g(T)$. As $\lim\limits_{T\to\infty} g(T) = +\infty$, there exists a time T_1 and index i, such that $h_i(T_1) \leq 0$. The lemma is established. □

Let

$$\hat{T} = \inf\{T \geq 0 : \inf_{v(\cdot)\in\Omega(T)} \max_i \int_0^T \beta_i(T, \tau, v(\tau))d\tau \geq 1\}.$$

Here $\Omega(T)$ is a set of all admissible functions v, defined on interval $[0, T]$ with values from V. By Lemma 3.5, $\hat{T} < \infty$.

Theorem 3.1. *Let Assumptions 3.1, 3.2 and condition (15) hold, and $\delta > 0$. Then there is a "soft" capture in the game Γ.*

Proof. Let $v : [0, \hat{T}] \to V$ be a random admissible control of evader E and let t_1 be a least positive root of function h. Here

$$h(t) = 1 - \max_i \int_0^t \beta_i(\hat{T}, \tau, v(\tau))\; d\tau.$$

Let $\hat{u}_i(\tau)$ be the lexicographical minimum among the roots of system

$$-\beta_i(\hat{T}, \tau, v(\tau))\mathcal{L}(\xi_i(\hat{T}), \hat{T}) = \mathcal{L}(\varphi_{l-1}(\hat{T} - \tau), \hat{T})(u - v(\tau)).$$

Define the controls of pursuers P_i as $u_i(\tau) = \hat{u}_i(\tau)$ and consider that the $\beta_i(\hat{T}, \tau, v(\tau))0$ at $\tau \in [t_1, \hat{T}]$. Then

$$\mathcal{L}(z_i(\hat{T}), \hat{T}) = \mathcal{L}(\xi_i(\hat{T}), \hat{T}) + \int_0^{\hat{T}} \mathcal{L}(\varphi_{l-1}(\hat{T} - \tau), \hat{T})(u_i(\tau) - v(\tau))d\tau$$

$$= \mathcal{L}(\xi_i(\hat{T}), \hat{T})h_i(\hat{T}) = \mathcal{L}(\xi_i(\hat{T}), \hat{T})(1 - \int_0^{t_1} \beta_i(\hat{T}, \tau, v(\tau))d\tau) = 0.$$

This proves the theorem. $\square$

Let's consider a case

$$\lambda_s = 0, \;\; k_s = 1 \tag{16}$$

and designate

$$M_1(t, \tau) = \min\Big\{\varphi_{l-1}(t - \tau),\;\; \frac{\dot{\varphi}_{l-1}(t - \tau)e^{-\lambda_{s-1}t}}{t^\mu}\Big\},$$

$$\mathcal{L}_1(f(r), t) = \left\| \begin{matrix} f(r) \\ \dot{f}(r)e^{-\lambda_{s-1}t}/t^\mu \end{matrix} \right\|.$$

In that case

$$\dot{\varphi}_{l-1}(t) = \sum_{r=1}^{s-1} e^{\lambda_r t}Q_r(t),\;\; \xi_i(t) = \sum_{j=1}^{s-1} e^{\lambda_j t}P_{ji}(t) + z_i^0, \dot{\xi}_i(t) = \sum_{j=1}^{s-1} e^{\lambda_j t}Q_{ji}(t).$$

Let $\deg Q_{s-1}(t) = \mu$. We will consider that $Q_{s-1i}(t) \neq 0$ and $\deg Q_{s-1i}(t) = \mu$ for all i. Let $z_i^1 = \lim\limits_{t\to\infty} Q_{s-1i}/t^\mu$.

Lemma 3.5. *Let Assumption 3.1 and condition (16) hold. Then for all $T > 0$*

$$\lim_{t\to\infty}\int\limits_T^t M_1(t,\tau)d\tau = \infty.$$

Proof. The functions $\varphi_{l-1}, \dot{\varphi}_{l-1}$ can be represented as

$$\begin{aligned}
&\varphi_{l-1}(t-\tau) = a_\gamma + g_1(t-\tau),\\
&\dot{\varphi}_{l-1}(t-\tau) = e^{\lambda_{s-1}(t-\tau)}(t-\tau)^\mu \lambda_{s-1} b_\mu [1 + g_2(t-\tau)].
\end{aligned}$$

Here

$$\begin{aligned}
g_1(t-\tau) &= \sum_{j=1}^{s-1} e^{\lambda_j(t-\tau)} P_j(t-\tau),\\
g_2(t-\tau) &= \sum_{j=1}^{s-2} \frac{e^{(\lambda_j-\lambda_{s-1})(t-\tau)} Q_j^1(t-\tau)}{(t-\tau)^\mu} + \sum_{r=0}^{\mu-1} \frac{b_r}{(t-\tau)^{\mu-r}}.
\end{aligned}$$

Suppose $\varepsilon \in (0,1)$, $\tau \in [0, \varepsilon t]$. Therefore, the inequality $t-\tau \geq (1-\varepsilon)t$ is true. And therefore, if $g_1(t-\tau), g_2(t-\tau)$, then

$$g_j(t-\tau) \leq \Delta_j(t)$$

$\Delta_j(t) \to 0$ at $t \to \infty$. Therefore, there exists T_0, such that

$$\varphi_{l-1}(t-\tau) \geq 1/2 a_\gamma, \quad \dot{\varphi}_{l-1}(t-\tau) \geq 1/2 e^{\lambda_{s-1}(t-\tau)}(t-\tau)^\mu b_\mu \lambda_{s-1}$$

for all $t > T_0$ and $\tau \in [0, \varepsilon t]$. Hence

$$\frac{\dot{\varphi}_{l-1}(t-\tau)e^{-\lambda_{s-1}t}}{t^\mu} \geq 1/2(1-\varepsilon)^\mu b_\mu \lambda_{s-1}.$$

Then for all $T > T_0$

$$\int\limits_T^t M_1(t,\tau)d\tau \geq \int\limits_T^{\varepsilon t} a d\tau \ \to \infty \ \ at \ \ t \to \infty.$$

This establishes the lemma. □

Let further

$$\delta = \inf_{v\in V} \max_i \min\{\lambda(z_i^0, v), \lambda(z_i^1, v)\}.$$

Assumption 3.3. The functions $\lambda(z_i^0, v)$, $\lambda(z_i^1, v)$ are continuous at all points (z_i^0, v), (z_i^1, v), such that $\lambda(z_i^0, v) > 0$, $\lambda(z_i^1, v) > 0$.

Lemma 3.6. *Let Assumptions 3.1, 3.3 hold with condition (16) and $\delta > 0$. Then for any admissible function v there exist time T and index i, such that $h_i(T) \le 0$. Here*

$$\beta_i(t, \tau, v) = \sup\Big\{\lambda \,\Big|\, \lambda \ge 0, -\lambda\mathcal{L}_1(\xi_i(t), t) \in \mathcal{L}_1(\varphi_{l-1}(t-\tau), t)(V - v)\Big\}.$$

Proof. Since $\delta > 0$ it follows that for any $v \in V$ there exists i such that $\lambda(z_i^0, v) > 0$, $\lambda(z_i^1, v) > 0$. By Assumption 3.3 and condition $z_i^0 = \lim\limits_{t\to\infty} \xi_i(t)$, $z_i^1 = \lim\limits_{t\to\infty} \dot{\xi}_i(t)e^{-\lambda_{s-1}t}/t^\mu$ we obtain that there exists a time T_1, such that for all $t > T_1$ the inequality

$$\inf_v \max_i \min\{\lambda(\xi_i(t), v),\ \lambda(\dot{\xi}_i(t)e^{-\lambda_{s-1}t}/t^\mu, v)\} \ge 1/2\delta$$

holds. Let $T > T_1$. Then

$$\sum_i h_i(T) \le n - 1/2\delta \int\limits_{T_1}^{T} M_1(t, \tau)d\tau = n - g(T).$$

By Lemma 3.5, $g(T) \to \infty$ at $T \to \infty$. Therefore, there exist a time T_0 and index i, such that $h_i(T_0) \le 0$. The lemma is proved. □

Theorem 3.2. *Let Assumptions 3.1, 3.3 hold together with condition (16) and $\delta > 0$. Then there is a "soft" capture in the game Γ.*

Proof. Let $v : [0, \hat{T}] \to V$ be a random admissible control of evader E and t_1 a least positive root of function h. Let $\hat{u}_i(\tau)$ be a lexicographical minimum among the roots of system

$$-\beta_i(\hat{T}, \tau, v(\tau))\mathcal{L}_1(\xi_i(\hat{T}), \hat{T}) = \mathcal{L}_1(\varphi_{l-1}(\hat{T} - \tau), \hat{T})(u - v(\tau)).$$

Let's assign the controls of pursuers P_i, as $u_i(\tau) = \hat{u}(\tau)$. We will consider that the $\beta_i(\hat{T}, \tau, v(\tau)) = 0$ at $\tau \in [t_1, \hat{T}]$. Then

$$\mathcal{L}_1(z_i(\hat{T}), \hat{T}) = \mathcal{L}_1(\xi_i(\hat{T}), \hat{T}) + \int\limits_0^{\hat{T}} \mathcal{L}_1(\varphi_i(\hat{T} - \tau), \hat{T})(u_i(\tau) - v(\tau))d\tau$$

$$= \mathcal{L}_1(\xi_i(\hat{T}), \hat{T})h_i(\hat{T}) = \mathcal{L}_1(\xi_i(\hat{T}), \hat{T})(1 - \int\limits_0^{t_1} \beta_i(\hat{T}, \tau, v(\tau))d\tau) = 0.$$

This establishes the theorem. □

Let

$$\mathcal{L}_2(f(r),t)=\frac{e^{-\lambda_s t}}{t^\gamma}\left\|\begin{matrix} f(r)\\ \dot f(r)/\lambda_s\end{matrix}\right\|,$$

$$M_2(t,\tau)=\min\Big\{\frac{\varphi_{l-1}(t-\tau)e^{-\lambda_s t}}{t^\gamma},\ \frac{-\dot\varphi_{l-1}(t-\tau)e^{-\lambda_s t}}{t^\gamma}\Big\}.$$

Lemma 3.7. *Let Assumption 3.1 hold, with $\lambda_s<0$, $\varepsilon\in(0,1)$. Then there exists T_0, such that for any $T>T_0$*

$$\lim_{t\to\infty}\int\limits_T^{\varepsilon t} M_2(t,\tau)d\tau=\infty.$$

Proof. The functions $\varphi_{l-1}, -\dot\varphi_{l-1}$ can be represented as

$$\varphi_{l-1}(t-\tau)=a_\gamma(t-\tau)^\gamma e^{\lambda_s(t-\tau)}(1+g_1(t,\tau)),$$
$$-\dot\varphi_{l-1}(t-\tau)=a_\gamma(-\lambda_s)(t-\tau)^\gamma e^{\lambda_s(t-\tau)}(1+g_2(t,\tau)).$$

Here

$$g_1(t,\tau)=\sum_{j=1}^{s-1}e^{(\lambda_j-\lambda_s)(t-\tau)}\frac{P_j(t-\tau)}{(t-\tau)^\gamma a_\gamma}+\sum_{l=0}^{\gamma-1}\frac{a_l}{(t-\tau)^{\gamma-l}},$$

$$g_2(t,\tau)=\sum_{j=1}^{s-1}e^{(\lambda_j-\lambda_s)(t-\tau)}\frac{Q_j(t-\tau)}{(t-\tau)^\gamma a_\gamma(-\lambda_s)}+\sum_{l=0}^{\gamma-1}\frac{b_l}{(t-\tau)^{\gamma-l}}.$$

Let $\tau\in(0,\varepsilon t)$. Then $t-\tau\ge(1-\varepsilon)t$; therefore,

$$|g_1(t,\tau)|\le\Delta_1(t),\quad |g_2(t,\tau)|\le\Delta_2(t).$$

$\Delta_1(t)$, $\Delta_2(t)\to 0$ at $t\to\infty$.

Therefore, there exists a time T_0, such that $|g_1(t,\tau)|\le 1/2$, $|g_2(t,\tau)|\le 1/2$ for all $t>T_0$, $\tau\in(0,\varepsilon t)$.

Therefore,

$$\frac{\varphi_{l-1}(t-\tau)e^{-\lambda_s t}}{t^\gamma}\ge\frac{a_\gamma(t-\tau)^\gamma e^{-\lambda_s t}}{t^\gamma},$$
$$-\frac{\dot\varphi_{l-1}(t-\tau)e^{-\lambda_s t}}{t^\gamma}\ge\frac{a_\gamma(t-\tau)^\gamma e^{-\lambda_s t}(-\lambda_s)}{t^\gamma}$$

for all $t>T_0$, $\tau\in(0,\varepsilon t)$.

Let $T>T_0$, $\varepsilon t>T$, $t(1-\varepsilon)\ge\tau_0$, $\tau\in(0,\varepsilon t)$. Then

$$\int\limits_T^{\varepsilon t}M_2(t,\tau)d\tau\ge\int\limits_T^{\varepsilon t}\frac{c(t-\tau)^\gamma e^{-\lambda_s t}}{t^\gamma}d\tau\to\infty\quad\text{at } t\to\infty.$$

The lemma is proved. □

Let $z_i^0 = \lim\limits_{t\to\infty} \xi_i(t)e^{-\lambda_s t}/t^\gamma$, $\delta = \inf\limits_{v\in V} \max\limits_i \lambda(z_i^0, v)$. and note that the

$$z_i^0 = \lim_{t\to\infty} \frac{\dot{\xi}_i(t)e^{-\lambda_s t}}{t^\gamma \lambda_s}.$$

Lemma 3.8. *Let Assumptions 3.1, 3.2 hold, with $\lambda_s < 0$, $\delta > 0$. Then there exists a time T, such that for any admissible function v there exists an index i, such that $h_i(T) \le 0$. Here*

$$\beta_i(T,\tau,v) = \begin{cases} \beta_i^1(T,\tau,v), & \text{if } T-\tau > \tau_0 \\ 0, & \text{if } T-\tau \le \tau_0 \end{cases},$$

$$\beta_i^1(t,\tau,v) = \sup\Big\{\lambda \;\Big|\; \lambda \ge 0, -\lambda\mathcal{L}_2(\xi_i(t),t) \in \mathcal{L}_2(\varphi_{l-1}(t-\tau),t)(V-v)\Big\},$$

and τ_0 is a positive root of function $\dot{\varphi}_{l-1}$.

The proof of this lemma is similar to the proof of Lemma 3.6.

We can now state our last theorem.

Theorem 3.3. *Let Assumptions 3.1, 3.2 hold, with $\lambda_s < 0$, $\delta > 0$. Then there is a "soft" capture in the game Γ.*

The proof of this theorem is similar to the proofs of the corresponding theorems for $\lambda_s = 0$.

REFERENCES

[1] Pshenichnij B.N., The simple pursuit by several objects, *Kibernetika*, 3, 145–146, 1976.

[2] Pontrjagin L.S., *The Selected Scientific Works*, Nauka, Moscow, 1988.

[3] Petrov N.N., *The Theory of Games*, Izhevsk, Izd-vo Udmurt University un-ta, 1997.

[4] Petrosjan L.A., *The Differential Games of Pursuit*, Izd-vo Leningrad State University un-ta, 1977.

[5] Subbotin A.I. and Chentsov A.G., *The Optimization of Guarantee in the Problems of Control*, Nauka, Moscow, 1981.

[6] Chikrij A.A., *Conflict Controllable Processes*, Nauk.dumka, Kiev, 1992.

[7] Chernousko F.L., The one problem of evasion from many pursuers, *PMM*, 40, 1, 14–24, 1976.

[8] Ivanov R.P., To the question about the soft capture in differential games with many pursuers and single evader, *Tr. Matem. in-ta AN SSSR*, 185, 74–83, 1988.

[9] Grigorenko N.L., *The Mathematical Methods of Control by Several Dynamical Processes*, Izd-vo MGU, 1990.

[10] Petrov N.N., Petrov N. Nikandr. About the differential game "kazaki–razboiniki", *Diff. urav.*, 19, 8, 1366–1374, 1983.

[11] Petrov N.N. and Prokopenko V.A., About one problem of pursuit of the group of evaders, *Diff. urav.*, 23, 4, 724–726, 1987.

[12] Chikrij A.A. and Prokopovich P.V., The linear problem of evasion under the interaction of the groups of objects, *PMM*, 58, 4, 12–21, 1994.

[13] Chikrij A.A. and Prokopovich P.V., About the problem of evasion under the interaction of groups of movable objects, *Kibernetika*, 5, 59–63, 1989.

[14] Satimov N.U. and Mamatov M.SH., About the problem of pursuit and evasion from the encounter in the differential games between groups of pursuers and evaders, *DAN. Uzb.SSR.*, 4, 3–6, 1983.

[15] Vagin D.A. and Petrov N.N., The problem of pursuit of the group of hardunited evaders, *Izv. RAN. Teorija i sistemi upravlenija*, 5, 75–79, 2001.

[16] Petrov N.N., About the controllability of the autonomous systems, *Diff. urav.*, 4, 4, 606–617, 1968.

[17] Rockafellar R., *Convex Analysis*, Mir, Moscow, 1973.

[18] Pontrjagin L.S., Boltjanskij V.G., Gamkrelidze R.V. and Mischenko E.F., *The Mathematical Theory of Optimal Processes*, Nauka, Moscow, 1969.

Cooperative Stochastic Games

Leon A. Petrosjan
St. Petersburg State University
St. Petersburg, Russia
spbuoasis7@peterlink.ru

Abstract

A cooperative stochastic n-person game on a finite graph tree is considered. The subtree of cooperative trajectories maximizing the sum of expected players' payoffs is defined, and the solution of the game along the paths of this tree is investigated. The new notion of cooperative payoff distribution procedure (CPDP) is defined, and the time-consistent Shapley value is constructed.

Key words. Stochastic game, stage game, behavior strategy, Shapley value, time-consistency.

1 Cooperative Game

Consider a finite graph tree $G = (Z, L)$ where Z is the set of all vertices and L is a point-to-set mapping ($L_z \subset Z, z \in Z$). In our setting each vertex $z \in Z$ is considered as an n-person simultaneous move (one-stage) game

$$\Gamma(z) = \langle N; X_1^z, \ldots, X_n^z; K_1^z, \ldots, K_n^z \rangle \,,$$

where $N = \{1, \ldots, n\}$ is the set of players which is the same for all $z \in Z$, X_i^z is the set of strategies of Player $i \in N$, and $K_i^z(x_1^z, \ldots, x_n^z)$ (with $K_i^z \geq 0$) is the payoff of Player i ($i \in N,\ x_i^z \in X_i^z$). The n-tuple $x^z = (x_1^z, \ldots, x_n^z)$ is called the situation in the game $\Gamma(z)$. The game $\Gamma(z)$ is called a *stage game*. For each $z \in Z$ the transition probabilities

$$p(z, y; x_1^z, \ldots, x_n^z) = p(z, y; x^z) \geq 0,$$

$$\sum_{y \in L_z} p(z, y; x^z) = 1$$

are given, where $p(z, y; x^z)$ is the probability that the game $\Gamma(y)$, $y \in L_z$ will be played next after the game $\Gamma(z)$, under the condition that in $\Gamma(z)$ the situation $x^z = (x_1^z, \ldots, x_n^z)$ was realized. We set $p(z, y; x^z) \equiv 0$ if $L_z = \emptyset$.

Suppose that the path $z_0, z_1, \dots, z_l$ ($L_{z_l} = \emptyset$) is realized during the game. Then the payoff of Player $i \in N$ is defined as

$$K_i(z_0) = \sum_{j=0}^{l} K_i^{z_j}(x^{z_j}).$$

Due to the stochastic transition from one stage game to the other we consider the mathematical expectation of the player's payoff

$$E_i(z_0) = \exp K_i(z_0).$$

The following relation holds:

$$E_i(z_0) = K_i^{z_0}(x^{z_0}) + \sum_{y \in L_{z_0}} p(z_0, y; x^{z_0}) E_i(y), \tag{1}$$

where $E_i(y)$ is the mathematical expectation of a Player i payoff in the stochastic subgame starting from the stage game $\Gamma(y)$, $y \in L_{z_0}$.

A strategy $\pi_i(\cdot)$ for Player $i \in N$ is a mapping which determines for each stage game $\Gamma(y)$ which local strategy x_i in this stage game is to be selected. Thus $\pi_i(y) = x_i^y$ for $y \in Z$.

We shall denote the stochastic game described above as $\overline{G}(z_0)$. We denote by $\overline{G}(z)$ any subgame of $\overline{G}(z_0)$ starting from the stage game $\Gamma(z)$ (played on a subgraph of the graph G starting from vertex $z \in Z$).

If $\pi_i(\cdot)$ is a strategy of Player $i \in N$ in $\overline{G}(z_0)$, then the trace of this strategy $\pi_i^y(\cdot)$, defined on a subtree $G(y)$ of G, is a strategy in a subgame $\overline{G}(y)$ of the game $\overline{G}(z_0)$.

The following version of (1) holds for a subgame $\overline{G}(z)$ (for the mathematical expectation of Player i's payoff in $\overline{G}(z)$):

$$E_i(z) = K_i^z(x^z) + \sum_{y \in L_z} p(z, y; x^z) E_i(y).$$

As mixed strategies in $\overline{G}(z_0)$ we consider behavior strategies [2]. Denote them $q_i(\cdot)$, $i \in N$, and denote the corresponding situation as

$$q^N(\cdot) = (q_1(\cdot), \dots, q_n(\cdot)).$$

Here $q_i(z)$ for each $z \in Z$ is a mixed strategy of Player i in a stage game $\Gamma(z)$. Denote by $\bar{\pi}^N(\cdot) = (\bar{\pi}_1(\cdot), \dots, \bar{\pi}_n(\cdot))$ the n-tuple of pure strategies in $\overline{G}(z_0)$ which maximizes the sum of expected players' payoffs (*cooperative solution*). Denote this maximal sum by $V(z_0)$

$$V(z_0) = \max E(z_0) = \max \left[\sum_{i \in N} E_i(z_0) \right].$$

It can be easily seen that the maximization of the sum of the expected payoffs of players over the set of behavior strategies does not exceed $V(z_0)$.

In the same way we can define a cooperative n-tuple of strategies for any subgame $\overline{G}(z)$, $z \in Z$, starting from the stage game $\Gamma(z)$. From Bellman's optimality principle it follows that each of these n-tuples is a trace of $\pi^N(\cdot)$ in the subgame $\Gamma(z)$. The following Bellman equation holds [1]:

$$
\begin{aligned}
V(z) &= \max_{\substack{x_i^z \in X_i^z \\ i \in N}} \left\{ \sum_{i \in N} K_i^z(x_i^z) + \sum_{y \in L_z} p(z, y; x^z) V(y) \right\} \\
&= \sum_{i \in N} K_i^z(\bar{x}^z) + \sum_{y \in L_z} p(z, y; \bar{x}^z) V(y)
\end{aligned} \tag{2}
$$

with the initial condition

$$
V(z) = \max_{\substack{x_i^z \in X_i^z \\ i \in N}} \sum_{i \in N} K_i^z(x^z), \; z \in \{z : \; L_z = \emptyset\}. \tag{3}
$$

The maximizing n-tuple $\bar{\pi}^N(\cdot) = (\bar{\pi}_1(\cdot), \dots, \bar{\pi}_n(\cdot))$ defines the probability measure over the game tree $G(z_0)$. Consider a subtree $\hat{G}(z_0)$ of $G(z_0)$ which consists of paths in $G(z_0)$ having a positive probability under the measure generated by $\bar{\pi}^N(\cdot)$. We call $\hat{G}(z_0)$ the cooperative subtree and $CZ \subset Z$ the set of vertices in $\hat{G}(z_0)$.

For each $z \in CZ$ let us define a zero-sum game over the structure of the graph $G(z)$ between the coalition $S \subset N$ considered as the maximizing player and anti-coalition $N \setminus S$ as the minimizing one. Let $V(S, z)$ be the value of this game in behavior strategies (the existence follows from [2]). Thus for each subgame $\bar{G}(z)$, $z \in CZ$, we can define a characteristic function $V(S, z)$, $S \subset N$, with $V(N, z) = V(z)$ defined by Equations (2) and (3).

Consider now the cooperative version $\overline{\overline{G}}(z)$, $z \in Z$, of a subgame $\overline{G}(z)$ with characteristic function $V(S, z)$. Let $I(z)$ be the imputation set in $\overline{\overline{G}}(z)$

$$
I(z) = \left\{ \alpha^z : \; \sum_{i \in N} \alpha_i^z = V(z) = V(N, z), \; \alpha_i^z \geq V(\{i\}, z) \right\}. \tag{4}
$$

A *solution* to $\overline{\overline{G}}(z)$ will be a particular subset $C(z) \subset I(z)$. This can be any classical cooperative solution (nucleous, core, NM-solution, Shapley value). In what follows we suppose that $C(z)$ is the Shapley value

$$
C(z) = Sh(z) = \{Sh_1(z), \dots, Sh_n(z)\} \in I(z),
$$

but all conclusions extend automatically to any other cooperative solution concept.

2 Cooperative Payoff Distribution Procedure (CPDP)

The vector function $\beta(z) = (\beta_1(z), \dots, \beta_n(z))$ is called a CPDP if

$$\sum_{i\in N} \beta_i(z) \le \max_{\substack{x_i^z \in X_i^z \\ i\in N}} \sum_{i\in N} K_i^z(x_1^z, \dots, x_n^z) = w(N, z). \tag{5}$$

Here $w(N, z)$ is the maximal total payoff of the players in a stage game $\Gamma(z)$. In each subgame $\overline{\overline{G}}(z)$, with each path $\bar{z} = z_0, \dots, z_m$ in this subgame we associate the sum of $\beta_i(z)$ along this path $\bar{z}$. Denote $B_i(z)$ the expected value of this sum in $\overline{\overline{G}}(z)$.

It can be easily seen that $B_i(z)$ satisfies the following functional equation:

$$B_i(z) = \beta_i(z) + \sum_{y\in L_z} p(z, y; x^z) B_i(y). \tag{6}$$

Calculate the Shapley value for each subgame $G(z)$ for $z \in CZ$

$$Sh_i(z) = \sum_{\substack{S\subset N \\ i\in S}} \frac{(|S|-1)!(n-|S|)!}{n!} (V(S, z) - V(S\setminus\{i\}, z)), \tag{7}$$

where $|S|$ is the number of elements in S.

Define $\gamma_i(z)$ by the formula

$$Sh_i(z) = \gamma_i(z) + \sum_{y\in Z} p(z, y; x^z) Sh_i(y). \tag{8}$$

Since $Sh(z) \in I(z)$ we get from (8)

$$V(N; z) = \sum_{i\in N} \gamma_i(z) + \sum_{y\in L_z} p(z, y; x^z) V(N; y). \tag{9}$$

Comparing (9) and (2) we get that $\sum_{i\in N} \gamma_i(z) = \sum_{i\in N} K_i^z(\tilde{x}^z)$ for some $\tilde{x}^z = (\tilde{x}_1^z, \dots, \tilde{x}_n^z)$, $\tilde{x}_i^z \in X_i^z$, $i \in N$ and thus

$$\sum_{i\in N} \gamma_i(z) \le w(N; z). \tag{10}$$

We have that $\gamma_i(z)$ satisfies (5) and the following lemma holds.

Lemma 2.1. *$\gamma(z) = (\gamma_1(z), \dots, \gamma_n(z))$ defined by (8) is CPDP.*

Definition 2.1. The Shapley value $\{Sh(z_0)\}$ is called time consistent in $\overline{\overline{G}}(z_0)$ if there exists a nonnegative CPDP ($\beta_i(z) \ge 0$) such that the following condition holds:

$$Sh_i(z) = \beta_i(z) + \sum_{y\in L_z} p(z, y; x^z) Sh_i(y), \ \ i \in N, \ \ z \in Z. \tag{11}$$

From (11) we get

$$\beta_i(z) = Sh_i(z) - \sum_{y \in L_z} p(z, y; x^z) Sh_i(y)$$

and the nonnegativity of CPDP $\beta_i(z)$ follows from the monotonicity of the Shapley value along the paths of a cooperative subgame $\hat{\hat{G}}(z_0)$.

As before, we call $B_i(z)$ the expected value of the sums of $\beta_i(y)$ from (11), $y \in Z$ along the paths in the cooperative subgame $\hat{\hat{G}}(z)$ of the game $\hat{\hat{G}}(z_0)$.

Lemma 2.2.

$$B_i(z) = Sh_i(z), \ i \in N. \tag{12}$$

Proof. We have for $B_i(z)$ Equation (6)

$$B_i(z) = \beta_i(z) + \sum_{y \in L_z} p(z, y; x^z) B_i(y) \tag{13}$$

with initial condition

$$B_i(z) = Sh_i(z) \text{ for } z \in \{z : \ L_z = \emptyset\}, \tag{14}$$

and for the Shapley value we have

$$Sh_i(z) = \beta_i(z) + \sum_{y \in L_z} p(z, y; x^z) Sh_i(y). \tag{15}$$

From (13), (14), (15) it follows that $B_i(z)$ and $Sh_i(z)$ satisfy the same functional equations with the same initial condition (14), and the proof follows from backward induction. □

Lemma 2.2 provides a natural interpretation for CPDP $\beta_i(z)$ which can be seen as the instantaneous payoff which Player i has to get in a stage game $\Gamma(z)$ when this game actually occurs along the paths of the cooperative subtree $\hat{\hat{G}}(z_0)$, if his payoff in the whole game has to be equal to the ith component of the Shapley value. So the CPDP shows the distribution in time of the Shapley value in such a way that the players in each subgame are oriented to get the current Shapley value of this subgame.

3 Regularization

In this section we propose a procedure, similar to the one used in differential cooperative games [3], which will guarantee the existence of a time-consistent Shapley value in the cooperative stochastic game (monotonicity of the Shapley value).

Introduce

$$\bar{\beta}_i(z) = \frac{\sum_{i\in N} K_i(\bar{x}_1^z, \ldots, \bar{x}_n^z)}{V(N,z)} Sh_i(z), \tag{16}$$

where $\bar{x}^z = (\bar{x}_1^z, \ldots, \bar{x}_n^z)$, $z \in Z$ is the realization of the n-tuple of strategies $\bar{\pi}(\cdot) = (\bar{\pi}_1(\cdot), \ldots, \bar{\pi}_n(\cdot))$ maximizing the sum of players' payoffs in the game $\overline{G}(z_0)$ (cooperative solution) and $V(N,z)$ is the value of the characteristic function for the grand coalition N in a subgame $\overline{\overline{G}}(z)$. Since

$$\sum_{i\in N} Sh_i(z) = V(N,z),$$

it follows from (16) that $\bar{\beta}_i(z)$, $i \in N$, $z \in Z$, is a CPDP. From (16) it also follows that the instantaneous payoff of the player in a stage game $\Gamma(z)$ must be proportional to the Shapley value in a subgame $\overline{\overline{G}}(z)$ of the game $\overline{\overline{G}}(z_0)$.

We now define by induction the *regularized Shapley value* (RSV) in $\overline{\overline{G}}(z)$ as follows:

$$\hat{S}h_i(z) = \frac{\sum_{i\in N} K_i(\bar{x}^z)}{V(N,z)} Sh_i(z) + \sum_{y\in L_z} p(z,y;\bar{x}^z)\hat{S}h_i(y) \tag{17}$$

with the initial condition

$$\hat{S}h_i(z) = \frac{\sum_{i\in N} K_i(\bar{x}^z)}{V(N,z)} Sh_i(z) = Sh_i(z) \text{ for } z \in \{z :\ L_z = \emptyset\}. \tag{18}$$

Since $K_i(x) \geq 0$ it follows from (16) that $\bar{\beta}_i(z) \geq 0$, and thus the RSV $\hat{S}h_i(z)$ is time consistent.

Introduce the new characteristic function $\hat{V}(S,z)$ in $\overline{\overline{G}}(z)$ by induction using the formula ($S \subset N$)

$$\hat{V}(S,z) = \frac{\sum_{i\in N} K_i(\bar{x}^z)}{V(N,z)} V(S,z) + \sum_{y\in L_z} p(z,y;\bar{x}^z)\hat{V}(S,y) \tag{19}$$

with the initial condition

$$\hat{V}(S,z) = V(S,z) \text{ for } z \in \{z :\ L_z = \emptyset\}.$$

Here $V(S,z)$ is superadditive, so is $\hat{V}(S,z)$, and $\hat{V}(N,z) = V(N,z)$ since both functions $\hat{V}(N,z)$ and $V(N,z)$ satisfy the same functional equation (2) with the initial condition (3). Rewriting (19) for $\{S \setminus i\}$ we get

$$\hat{V}(S\setminus i,z) = \frac{\sum_{i\in N} K_i(\bar{x}^z)}{V(N,z)} V(S\setminus i,z) + \sum_{y\in L_z} p(z,y;\bar{x}^z)\hat{V}(S\setminus i,y). \tag{20}$$

Subtracting (20) from (19), multiplying by $\frac{(|S|-1)!(n-|S|)!}{n!}$ and summing upon $S \subset N$, $i \in S$ we get

$$\sum_{\substack{S \subset N \\ i \in S}} \frac{(|S|-1)!(n-|S|)!}{n!} \left[\hat{V}(S,z) - \hat{V}(S \setminus i, z) \right]$$
$$= \left\{ \sum_{\substack{S \subset N \\ i \in S}} \frac{(|S|-1)!(n-|S|)!}{n!} [V(S,z) - V(S \setminus i, z)] \right\} \frac{\sum_{i \in N} K_i(\bar{x}^z)}{V(N,z)}$$
$$+ \sum_{y \in L_z} p(z,y;\bar{x}^z) \left\{ \sum_{\substack{S \subset N \\ i \in S}} \frac{(|S|-1)!(n-|S|)!}{n!} \left[\hat{V}(S,z) - \hat{V}(S \setminus i, z) \right] \right\}. \tag{21}$$

From (17), (18) and (21) it follows that the RSV $\hat{Sh}(z)$ is a Shapley value for the characteristic function $\hat{V}(S,z)$.

This can be summarized in the following theorem.

Theorem 3.1. *The RSV is time consistent and is a Shapley value for the regularized characteristic function $\hat{V}(S,z)$ defined for any subgame $\overline{\overline{G}}(z)$ of the stochastic game $\overline{\overline{G}}(z_0)$.*

REFERENCES

[1] Bellman R., *Dynamic Programming*, Princeton University Press, Princeton, NJ, 1957.

[2] Kuhn H.W., Extensive games and the problem of information, *Annals of Mathematics Studies*, 28, 193–216, 1953.

[3] Petrosjan L.A., The Shapley value for differential games, *Annals of the International Society of Dynamic Games*, 3, 409–419, 1995.

[4] Shapley L.S., A value for n-person games, *Annals of Mathematics Studies*, 28, 307–317, 1953.

[5] Shapley L.S., Stochastic games, *Proceedings of the National Academy of U.S.A.*, 39, 1095–1100, 1953.

The Uniqueness of a Reduced Game in a Characterization of the Core in Terms of Consistency*

Yukihiko Funaki
School of Political Science and Economics
Waseda University
Tokyo 169-8050, Japan
funaki@waseda.jp

Takehiko Yamato
Department of Value and Decision Science and
Department of Social Engineering
Graduate School of Decision Science and Technology
Tokyo Institute of Technology
Tokyo 152-8552, Japan

Abstract

In this paper, we examine the uniqueness of a reduced game in an axiomatic characterization of the core of transferable utility (TU) games in terms of consistency. Tadenuma [10] establishes that the core is the only solution satisfying non-emptiness, individual rationality, and consistency with respect to a natural reduced game due to Moulin [6]. However, the core satisfies consistency with respect to many other reduced games, including unnatural ones. Then we ask whether there are other reduced games that can be used to characterize the core based on the same three axioms. The answer is no: the Moulin reduced game is the only reduced game such that the core is characterized by the three axioms, since for any other reduced game, there is a solution that satisfies the three axioms, but it differs from the core. Many other unnatural reduced games cannot be used to characterize the core based on the three axioms. Funaki [4] provides another axiomatization of the core: the core is the only solution satisfying non-emptiness, Pareto optimality, sub-grand rationality, and consistency with respect to a simple reduced game similar to a so-called subgame. We show that the simple reduced game is the only reduced game that can be used to characterize the core by the four axioms.

*We thank an anonymous referee for useful comments. This research was partially supported by the Grants in Aid for Research 15310023 and 16530126 of the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

1 Introduction

One of the most fundamental conditions on solutions for cooperative games is *consistency*, described as follows. Pick a payoff vector selected by a solution for some game. Suppose that the members of some group S want to renegotiate the payoff distribution among them, while all the members agree on the payoff distribution to the members outside S. A coalition in S may cooperate with some members outside S and pay the agreed-upon payoffs for them. Such a situation is represented by a *reduced game*. The solution is *consistent* if it recommends the same payoff distribution for the reduced game as initially. If a solution was not consistent, then a redistribution of payoffs would be necessary.[1]

This paper considers axiomatic characterizations of the core of transferable utility (TU) games in terms of consistency. In a cooperative game situation, there are several possibilities to make a reduced game from a given game. Different reduced games have been used to characterize the core of TU games based on consistency (e.g., see Peleg [7], Tadenuma [10], and Funaki [4]). Among them, Tadenuma [10] employs a natural reduced game due to Moulin [6], described as follows. Take a TU game, a payoff vector x in a solution, and a player j. The player set of a reduced game is obtained by removing player j from the original player set N.[2] The worth of each subcoalition in $N \setminus \{j\}$ is equal to the worth of the subcoalition with player j minus the payoff x_j, that is, each subcoalition is required to involve player j and to pay him according to the original payoff x_j. Moulin [6] proposes this reduced game in the context of cost allocation problems with quasi-linear preferences. Tadenuma [10] establishes that on the class of TU games for which the core is non-empty, the core is the only solution satisfying non-emptiness, individual rationality, and consistency with respect to the Moulin reduced game.

Besides the Moulin reduced game, however, the core satisfies consistency with respect to many other reduced games, including unnatural ones. An example of such a reduced game is as follows. The worth of the coalition $N \setminus \{j\}$ is equal to the worth of N minus the payoff x_j. For a proper subcoalition S in $N \setminus \{j\}$, if the cardinality of S is odd, then the worth of S is the same as the original game; and if the cardinality of S is even, then the worth of S is equal to the worth of S with player j minus the payoff x_j. The core satisfies consistency with respect to this rather unnatural reduced game.

In this paper, we consider a large class of reduced games for which the core satisfies consistency, containing the Moulin reduced game as well as unnatural reduced games such as the one of the preceding example. Then we ask whether

[1] See Thomson [11] and Driessen [3] for surveys of consistency in game theory and economics.

[2] A reduced game is usually defined by removing a subset of players from the original player set. Our definition of a reduced game is identical to the usual one under a suitable condition. See Remark 2.1.

there are reduced games in the class, other than the Moulin reduced game, that can be used to characterize the core based on the three axioms that Tadenuma [10] employs. The answer is no: for any reduced game except the Moulin reduced game, there is a solution that satisfies non-emptiness, individual rationality, and consistency with respect to that reduced game, but it differs from the core, which recommends payoff vectors that do not belong to the core for some games. In this sense, the Moulin reduced game is the only reduced game such that the core is characterized by the three axioms. Many other unnatural reduced games cannot be used to characterize the core based on the three axioms.

Funaki [4] provides another axiomatization of the core: the core is the only solution satisfying non-emptiness, Pareto optimality, sub-grand rationality, and consistency with respect to a simple reduced game similar to a subgame. This simple reduced game belongs to the class of reduced games we consider. We ask whether there are other reduced games in the class that can be used to characterize the core based on the four axioms. The answer is no once again: for any reduced game other than the simple reduced game, there is a solution that satisfies non-emptiness, Pareto optimality, sub-grand rationality, and consistency with respect to that reduced game, but it differs from the core. The simple reduced game is the unique reduced game such that the core is characterized by the four axioms.

In most work on an axiomatic characterization of a solution in terms of consistency, given a reduced game, it has been examined which axioms characterize the solution. The question that we ask is quite different. Given a set of axioms including consistency, we investigate which reduced games can be used to characterize the core. We find that only the Moulin or simple reduced game can be employed.

The paper is organized as follows. In Section 2, we introduce notation and definitions for TU games, reduced games, and corresponding consistency properties. In Section 3, we prove the uniqueness of a reduced game for the characterization of the core of TU games due to Tadenuma [10] and for that due to Funaki [4]. In Section 4, we make some concluding remarks.

2 TU Games, Reduced Games, and Consistency Properties

There is an infinite set of *potential players* indexed by the natural numbers, $\mathcal{N}$. A *transferable utility (TU) game* is a pair (N, v), where N is a finite set of players drawn from $\mathcal{N}$, and v is a function that associates with each subset of N a real number. We assume that $v(\emptyset) = 0$.

Let a TU game (N, v) be given. A *coalition* in N is a non-empty subset of N. For each coalition S in N, the *worth* of S is $v(S)$. A *payoff vector* for N is a vector $x \in \mathcal{R}^N$, and for each coalition S in N, the restriction of x to $\mathcal{R}^S$ is denoted by $x_S \in \mathcal{R}^S$. For $x \in \mathcal{R}^N$ and $S \subseteq N$, let $x(S) = \sum_{i \in S} x_i$ and

$x(\emptyset) = 0$. The payoff vector x for N is *feasible* in (N, v) if $v(N) \geq x(N)$. The set of all feasible payoff vectors for N in (N, v) is denoted by $X(N, v)$.

A feasible payoff vector $x \in X(N, v)$ is *Pareto optimal* for (N, v) if $x(N) = v(N)$. It is *individually rational* for (N, v) if for all $i \in N$, $x_i \geq v(\{i\})$. Let $PO(N, v)$ and $IR(N, v)$ be the set of Pareto optimal payoff vectors for (N, v) and the set of individually rational payoff vectors for (N, v) respectively.

Let Γ be a class of TU games. A *solution* on Γ is a correspondence Φ which associates with each game $(N, v) \in \Gamma$ a subset $\Phi(N, v)$ of $X(N, v)$.

We now consider one of the most important solutions in game theory and economics.

Definition 2.1. For each $(N, v) \in \Gamma$, the *core* $C(N, v)$ on Γ is defined by

$$C(N, v) \equiv \{x \in X(N, v) \mid x(S) \geq v(S) \text{ for all } S \subseteq N, S \neq \emptyset\}.$$

It is easily checked that $C(N, v) \subseteq PO(N, v) \cap IR(N, v)$ for each $(N, v) \in \Gamma$.

Next we give a general definition of a reduced game including the reduced game due to Moulin [6] and a simple reduced game as special cases. Take a TU game, a payoff vector x in a solution, and a player j. The player set of a reduced game is obtained by removing player j from the original player set N. The worth of the coalition $N \setminus \{j\}$ is equal to the worth of N minus the payoff x_j. It is easy to show that any characteristic function in any reduced game satisfies this property if the solution satisfies Pareto optimality and consistency. On the other hand, there are two possibilities for the worth of a proper subcoalition in $N \setminus \{j\}$: it is equal to either (i) the same worth of the original game, or (ii) the worth of the subcoalition with player j minus the payoff x_j. We assume that whether (i) or (ii) holds depends on the cardinality of the subcoalition.

Definition 2.2. Let $\mathcal{S}$ be a mapping from $\{n \in \mathcal{N} \mid n \geq 3\}$ into $2^{\mathcal{N}}$ such that for all $n \in \mathcal{N}$ with $n \geq 3$, $\mathcal{S}(n) \subseteq \{1, 2, \ldots, n-2\}$. Given a TU game $(N, v) \in \Gamma$, a player $j \in N$, and a payoff vector $x \in \mathcal{R}^N$, the *$\mathcal{S}$-reduced game with respect to j and x* is the game $(N \setminus \{j\}, v_x^{\mathcal{S}})$ where

$$v_x^{\mathcal{S}}(S) = \begin{cases} 0 & \text{if } S = \emptyset, \\ v(N) - x_j & \text{if } S = N \setminus \{j\}, \\ v(S) & \text{if } |S| \in \mathcal{S}(|N|), \\ v(S \cup \{j\}) - x_j & \text{otherwise.} \end{cases}$$

If the cardinality of a proper subcoalition S in the reduced game $(N \setminus \{j\}, v_x^{\mathcal{S}})$ is in $\mathcal{S}(|N|)$, the worth of the subcoalition is equal to $v(S)$. Otherwise it is equal to $v(S \cup \{j\}) - x_j$. Thus the mapping $\mathcal{S}$ determines the structure of a reduced game definitively. Then we call $\mathcal{S}$ a *reduced game structure.*

In the following, we require a reduced game structure to satisfy the following condition:

$$\text{for all } m, n \in \mathcal{N} \text{ with } m \leq n, \mathcal{S}(m) = \mathcal{S}(n) \cap \{k \in \mathcal{N} | k \leq m\}. \qquad (1)$$

That is, the set $\mathcal{S}(m)$ for the smaller number of players is a restriction of the set $\mathcal{S}(n)$ for the larger number of players. We cite some examples of reduced game structures satisfying this condition.

Example 2.1. Let $\mathcal{S}(n) = \emptyset$ for all $n \in \mathcal{N}$ with $n \geq 3$. Then we get the reduced game proposed by Moulin [6]:

$$v_x^{\mathcal{S}}(S) = \begin{cases} 0 & \text{if } S = \emptyset, \\ v(S \cup \{j\}) - x_j & \text{otherwise.} \end{cases}$$

We call this game structure the *M-reduced game structure.*

Example 2.2. Let $\mathcal{S}(n) = \{1, 2, \ldots, n-2\}$ for all $n \in \mathcal{N}$ with $n \geq 3$. Then we get a simple reduced game:

$$v_x^{\mathcal{S}}(S) = \begin{cases} v(S) & \text{if } S \subset N \setminus \{j\}, \\ v(N) - x_j & \text{if } S = N \setminus \{j\}. \end{cases}$$

We call this reduced game structure the *SIM-reduced game structure.*

Example 2.3. Let $\mathcal{S}(n) = \{1, 3, 5, \ldots, n-2\}$ if n is odd with $n \geq 3$, and $\mathcal{S}(n) = \{1, 3, 5, \ldots, n-3\}$ if n is even with $n \geq 3$. Then we get the following unnatural reduced game:

$$v_x^{\mathcal{S}}(S) = \begin{cases} v(S) & \text{if } |S| \text{ is odd,} \\ v(S \cup \{j\}) - x_j & \text{if } |S| \text{ is even,} \\ v(N) - x_j & \text{if } S = N \setminus \{j\}. \end{cases}$$

Remark 2.1. A reduced game (T, v_x^T) is usually defined by reducing N to a subset T. In this case, all the members in $N \setminus T$ are supposed to go out at once. If each person in $N \setminus T$ goes out in an order, then we can make this type of a reduced game from a reduced game structure $\mathcal{S}$. In general, if the orders of going out differ, then the corresponding reduced games derived from a reduced game structure could differ. In this sense, our successive elimination approach is not equivalent to the simultaneous elimination approach. However, for the reduced game due to Moulin [6] and the simple reduced game, the eliminating order does not matter, that is, for any order of going out, the same reduced game is derived from a reduced game structure.

We provide a general definition of the consistency property of a solution with respect to a reduced game structure. Let Φ be a solution on Γ and $\mathcal{S}$ be a reduced game structure.

$\mathcal{S}$-Consistency, $\mathcal{S}$-*CONS*: For all $(N, v) \in \Gamma$ with $|N| \geq 2$, if $j \in N$ and $x \in \Phi(N, v)$, then $(N \setminus \{j\}, v_x^{\mathcal{S}}) \in \Gamma$ and $x_{N \setminus \{j\}} \in \Phi(N \setminus \{j\}, v_x^{\mathcal{S}})$.

$\mathcal{S}$-$CONS$ implies that given a TU game (N, v), if x is a solution payoff vector for (N, v), then for every player $j \in N$, the payoff vector for $N \setminus \{j\}, x_{N\setminus\{j\}}$, must be a solution payoff vector for the reduced game $\Phi(N \setminus \{j\}, v_x^{\mathcal{S}})$. It is a kind of internal consistency requirement to guarantee that players respect recommendations by the solution.

For any possible reduced game structure $\mathcal{S}$ including the Moulin reduced game structure and the simple reduced game structure, the core satisfies $\mathcal{S}$-$CONS$ as the following proposition shows. Let Σ be the class of all possible reduced game structures.

Proposition 2.1. *For all reduced game structures $\mathcal{S} \in \Sigma$, the core C satisfies $\mathcal{S}$-$CONS$.*

Proof. Let a reduced game structure $\mathcal{S} \in \Sigma$ and a TU game $(N, v) \in \Gamma_c$ be given. Take any $j \in N$ and any $x \in C(N, v)$. Then $x(N) = v(N)$ and $x(N \setminus \{j\}) = v(N) - x_j = v_x^{\mathcal{S}}(N \setminus \{j\})$. Consider any $S \subset N \setminus \{j\}$. Then $x(S) \geq v(S)$. Also, $x(S \cup \{j\}) \geq v(S \cup \{j\})$, i.e., $x(S) \geq v(S \cup \{j\}) - x_j$. These imply that $x(S) \geq v_x^{\mathcal{S}}(S)$. Therefore, $x_{N\setminus\{j\}} \in C(N \setminus \{j\}, v_x^{\mathcal{S}})$, and $C(N \setminus \{j\}, v_x^{\mathcal{S}}) \neq \emptyset$. Thus $(N \setminus \{j\}, v_x^{\mathcal{S}}) \in \Gamma_c$. □

3 Characterizations of the Core

Let Γ be a class of TU games and Φ be a solution on Γ.

Non-emptiness, **NE**: For all $(N, v) \in \Gamma, \Phi(N, v) \neq \emptyset$.

Pareto optimality, **PO**: For all $(N, v) \in \Gamma, \Phi(N, v) \subseteq PO(N, v)$.

Individual rationality, **IR**: For all $(N, v) \in \Gamma, \Phi(N, v) \subseteq IR(N, v)$.

Individual rationality for games of at most *k* persons, $\mathbf{IR^k}$: For all $(N, v) \in \Gamma$ with $|N| \in \{1, 2, \ldots, k\}, \Phi(N, v) \subseteq IR(N, v)$.

NE, PO, and IR are standard properties for solutions. IR^k says that for any TU game of at most n players, if x is a solution payoff vector, then it must be individually rational. In this paper, we consider the individual rationality axiom for $k = 1$, IR^1, and the axiom for $k = 2$, IR^2. IR^1 requires $\Phi(\{i\}, v) = v(\{i\})$ for any i in $\mathcal{N}$. Under this weak condition, consistency implies Pareto optimality for any $\mathcal{S}$.

Lemma 3.1. *For all reduced game structures $\mathcal{S} \in \Sigma$, if a solution Φ satisfies IR^1 and $\mathcal{S}$-$CONS$, then it satisfies PO.*

Proof. By way of contradiction, suppose that Φ does not satisfy PO. Then for some $(N, v) \in \Gamma$ and some $x \in \Phi(N, v)$, $x(N) < v(N)$. Since Φ satisfies IR^1, $\Phi(\{i\}, v) = v(\{i\})$ for any i in $\mathcal{N}$. Therefore, $|N| \geq 2$. We denote the set of players by $N = \{1, 2, \ldots, n\}$ where $n = |N|$. For the game $(N \setminus \{n\}, v_x^{\mathcal{S}})$, $v_x^{\mathcal{S}}(N \setminus \{n\}) = v(N) - x_n > x(N) - x_n = x(N \setminus \{n\})$. Next for the game $(N \setminus \{n, n-1\}, (v_x^{\mathcal{S}})^2)$ where $(v_x^{\mathcal{S}})^2 \equiv (v_x^{\mathcal{S}})_x^{\mathcal{S}}, (v_x^{\mathcal{S}})^2(N \setminus \{n, n-1\}) = v_x^{\mathcal{S}}(N \setminus \{n\}) - x_{n-1} > x(N \setminus \{n\}) - x_{n-1} = x(N \setminus \{n, n-1\})$. By repeating the argument $n - 1$ times, we have

$$(v_x^{\mathcal{S}})^{n-1}(\{1\}) \equiv ((\ldots (v_x^{\mathcal{S}})_x^{\mathcal{S}})_x^{\mathcal{S}} \ldots)_x^{\mathcal{S}}(\{1\}) > x_1.$$

By $\mathcal{S}$-$CONS$, $x \in \Phi(N, v)$ implies $x_{N \setminus \{n\}} \in \Phi(N \setminus \{n\}, v_x^{\mathcal{S}})$, $x_{N \setminus \{n\}} \in \Phi(N \setminus \{n\}, v_x^{\mathcal{S}})$ implies $x_{N \setminus \{n, n-1\}} \in \Phi(N \setminus \{n, n-1\}, (v_x^{\mathcal{S}})^2)$, and so on. By repeating the argument $n - 1$ times, we get $x_{\{1\}} \in \Phi(\{1\}, (v_x^{\mathcal{S}})^{n-1})$. By $IR^1, x_1 \geq (v_x^{\mathcal{S}})^{n-1}(\{1\})$. Therefore, we have a contradiction. □

Let $\Gamma_c \equiv \{(N, v) \mid C(N, v) \neq \emptyset\ \}$ be the class of all balanced TU games. Tadenuma [10] establishes that the core is the only solution on Γ_c satisfying non-emptiness, individual rationality, and consistency with respect to the M-reduced game structure specified in Example 2.1. We ask whether there are other reduced game structures that can be used to characterize the core based on the same three axioms. The answer is no, as the following theorem shows.

Theorem 3.1. *There is a unique reduced game structure $\mathcal{S} \in \Sigma$ such that a solution Φ on Γ_c satisfies NE, IR, and $\mathcal{S}$-CONS if and only if $\Phi = C$, and it is the M-reduced game structure.*

Remark 3.1. Theorem 3.1 states that there is no reduced game structure other than the M-reduced game structure that characterizes the core by the three axioms NE, IR, and $\mathcal{S}$-$CONS$. However a more general result is shown in the proof of the theorem:

For any reduced game structure $\mathcal{S} \in \Sigma$ other than the M -reduced game structure, there is a solution Φ which satisfies NE, IR, and $\mathcal{S}$-$CONS$ such that $\Phi(N, v) \supseteq C(N, v)$ for all games $(N, v) \in \Gamma_c$ and $\Phi(N, v) \supset C(N, v)$ for some game $(N, v) \in \Gamma_c$.

In other words, there is a super-solution of the core that satisfies the three axioms except for the M-reduced game structure. Moreover this means that those three axioms cannot be used to characterize any subsolution of the core for any possible game structure.

We have the characterization due to Tadenuma [10] as an immediate corollary of Theorem 3.1. Let us call the consistency property with respect to the M-reduced game structure *M-Consistency* (M-$CONS$).

Corollary 3.1 (Tadenuma [10]). *A solution Φ on Γ_c satisfies NE, IR, and M-CONS if and only if $\Phi = C$.*

Proof of Theorem 3.1. We will prove Theorem 3.1 by using the following auxiliary lemma. We first define some conditions on the solutions to describe the lemma.

Condition $\mathcal{S}$: for all $(N,v) \in \Gamma$ with $|N| = n \geq 3$, for all $x \in \Phi(N,v)$, $x(S) \geq v(S)$ for all $S \subseteq N$ such that $S \neq N$ and

$$|S| \in \begin{cases} \{1\} \cup K(n) & \text{if } 1 \notin \mathcal{S}(n), \\ K(n) & \text{if } 1 \in \mathcal{S}(n), \end{cases}$$

where $K(n) \equiv \{s \mid 2 \leq s \leq n-1, s-1 \in \mathcal{S}(n), s \notin \mathcal{S}(n)\}$.

Lemma 3.2. *Let a reduced game structure $\mathcal{S} \in \Sigma$ be given. A solution Φ on Γ_c satisfies $NE, IR^2, \mathcal{S}$-$CONS$, and Condition $\mathcal{S}$ if and only if $\Phi = C$.*

The proof of Lemma 3.2 is provided in the Appendix.

It was Tadenuma [10] who first proved that the core is the unique solution Φ on Γ_c that satisfies NE, IR, and M-$CONS$. Here we give a simple alternative proof of this result by applying Lemma 3.2 to the M-reduced game structure. Let $\mathcal{S}(n) = \emptyset$ for all $n \in \mathcal{N}$ with $n \geq 3$. Clearly, IR implies IR^2. We prove that IR implies Condition $\mathcal{S}$. Let $(N,v) \in \Gamma_c$, $j \in N$, and $x \in \Phi(N,v)$ be given. Note that since $\mathcal{S}(|N|) = \emptyset$, $1 \notin \mathcal{S}(|N|)$ and $K(n) = \{s \mid 2 \leq s \leq n-1, s-1 \in \mathcal{S}(n), s \notin \mathcal{S}(n)\} = \emptyset$. If Φ satisfies IR, then $x(T) \geq v(T)$ for all $T \in \{S \subseteq N \mid |S| = 1\} = \{S \subseteq N \mid S \neq N \text{ and } |S| \in \{1\} \cup K(n)\}$, so that Φ satisfies Condition $\mathcal{S}$. By Lemma 3.2, we have the desired result.

Take any reduced game structure $\mathcal{S} \in \Sigma$ other than the M-reduced game structure. We show that there is a solution Φ on Γ_c that satisfies NE, IR, and $\mathcal{S}$-$CONS$, but $\Phi \neq C$. Consider the solution Φ defined as follows: for each $(N,v) \in \Gamma_c$, $\Phi(N,v) \equiv \{x \in PO(N,v) \cap IR(N,v) \mid x(T) \geq v(T) \text{ for all } T \subseteq N$ such that $|T| \notin K(n)\}$, where $n = |N|$. First of all, we prove that $\Phi \neq C$. Since $\mathcal{S}$ is different from the M-reduced game structure, $\mathcal{S}(n^*) \neq \emptyset$ for some $n^* \in \mathcal{N}$ with $n^* \geq 3$. We show that $K(n^*) = \{s|2 \leq s \leq n^*-1, s-1 \in \mathcal{S}(n^*), s \notin \mathcal{S}(n^*)\} \neq \emptyset$. Suppose that $K(n^*) = \emptyset$. Then $s \in \mathcal{S}(n^*)$ for all $s-1 \in \mathcal{S}(n^*)$. This and $\mathcal{S}(n^*) \neq \emptyset$ together imply $n^*-2 \in \mathcal{S}(n^*)$. Since $n^*-1 \notin \mathcal{S}(n^*) \subseteq \{1,2,\ldots,n^*-2\}$, it follows that $n^*-1 \in K(n^*)$, which is a contradiction. Therefore, $\Phi \neq C$.

Clearly, Φ satisfies NE and IR. We prove that it also satisfies $\mathcal{S}$-$CONS$. Let $(N,v) \in \Gamma_c$ with $n = |N|$, $j \in N$, and $x \in \Phi(N,v)$ be given. First we remark that

$$\begin{aligned} \{1,2,\ldots,n-1\} \setminus K(n) = \{&s|1 \leq s \leq n-1, s \in \mathcal{S}(n)\} \\ &\cup \{s \in \mathcal{N}|1 \leq s \leq n-1, s-1 \notin \mathcal{S}(n), s \notin \mathcal{S}(n)\}. \end{aligned}$$

Since $x \in \Phi(N,v)$, $x(T) \geq v(T)$ for all $T \subset N$ such that $|T| \in \{1,2,\ldots,n-1\} \setminus K(n)$. Consider the $\mathcal{S}$-reduced game $(N \setminus \{j\}, v_x^{\mathcal{S}})$. We show that $x(T) \geq v_x^{\mathcal{S}}(T)$ for any subset $T \subset N \setminus \{j\}$ such that

$$\begin{aligned} |T| &\in \{1,2,\ldots,n-2\} \setminus K(n-1) \\ &= \{s \mid 1 \leq s \leq n-2, s \in \mathcal{S}(n-1)\} \end{aligned}$$

$$\cup \{s \in \mathcal{N} | 1 \leq s \leq n-2, s-1 \notin \mathcal{S}(n-1), s \notin \mathcal{S}(n-1)\}.$$

Consider any T such that $|T| \in \{s \mid 1 \leq s \leq n-2, s \in \mathcal{S}(n-1)\}$. Because $\{s \mid 1 \leq s \leq n-2, s \in \mathcal{S}(n-1)\}$ is included in $\{s \mid 1 \leq s \leq n-1, s \in \mathcal{S}(n)\}$ by the condition (1) and $x \in \Phi(N, v)$, $v_x^{\mathcal{S}}(T) = v(T) \geq x(T)$.

Next consider any T such that $|T| \in \{s \in \mathcal{N} \mid 1 \leq s \leq n-2, s-1 \notin \mathcal{S}(n-1), s \notin \mathcal{S}(n-1)\}$. There are three cases to examine.

Case 1: $|T| = n-2 \in \mathcal{S}(n)$. Since $|T| \in \{1, 2, \ldots, n-1\} \setminus K(n)$, $x(T) \geq v(T) = v_x^{\mathcal{S}}(T)$.

Case 2: $|T| = n-2 \notin \mathcal{S}(n)$. Since $n-1 \notin \mathcal{S}(n)$, $|T \cup \{j\}| = n-1 \in \{1, 2, \ldots, n-1\} \setminus K(n)$, so that $x(T \cup \{j\}) \geq v(T \cup \{j\})$. Moreover $v_x^{\mathcal{S}}(T) = v(T \cup \{j\}) - x_j$ by $|T| \notin \mathcal{S}(n)$. Thus $x(T) \geq v_x^{\mathcal{S}}(T)$.

Case 3: $|T| \leq n-3$. In this case, since $|T|-1 \notin \mathcal{S}(n-1)$ and $|T| \notin \mathcal{S}(n-1)$, it follows from the condition (1) that $|T|-1 \notin \mathcal{S}(n)$ and $|T| \notin \mathcal{S}(n)$. Therefore, $|T \cup \{j\}| \in \{1, 2, \ldots, n-1\} \setminus K(n)$, and so $x(T \cup \{j\}) \geq v(T \cup \{j\})$. Moreover, $v_x^{\mathcal{S}}(T) = v(T \cup \{j\}) - x_j$ by $|T| \notin \mathcal{S}(n)$. Thus $x(T) \geq v_x^{\mathcal{S}}(T)$.

Finally consider $T = N \setminus \{j\}$. In this case, $v_x^{\mathcal{S}}(N \setminus \{j\}) = v(N) - x_j$. Moreover, $x(N) \geq v(N)$ by Pareto optimality of the solution Φ. Therefore, $x(N \setminus \{j\}) \geq v_x^{\mathcal{S}}(N \setminus \{j\})$. $\square$

For the SIM-reduced game structure in Example 2.2, Funaki [4] shows that the core is the only solution on Γ_c satisfying non-emptiness, Pareto optimality, consistency with respect to that reduced game structure, and the following axiom.

Sub-grand rationality (*SGR*): For all $(N, v) \in \Gamma$, if $x \in \Phi(N, v)$, then $x(N \setminus \{i\}) \geq v(N \setminus \{i\})$ for all $i \in N$.

Sub-grand rationality requires group rationality for every coalition consisting of all players except one. We ask whether there are other reduced game structures that can be used to characterize the core based on the same four axioms. Again, the answer is no, as the following theorem indicates.

Theorem 3.2. *There is a unique reduced game structure $\mathcal{S} \in \Sigma$ such that a solution Φ on Γ_c satisfies NE, PO, $\mathcal{S}$-CONS, and SGR if and only if $\Phi = C$, and it is the SIM-reduced game structure.*

Remark 3.2. Theorem 3.2 states that there is no reduced game structure other than the SIM-reduced game structure that characterizes the core by the four axioms NE, PO, SGR, and $\mathcal{S}$-$CONS$. However a more general result is shown in the proof of the theorem.

For any reduced game structure $\mathcal{S} \in \Sigma$ other than the SIM-reduced game structure, there is a solution Φ which satisfies NE, PO, SGR, and $\mathcal{S}$-$CONS$

such that $\Phi(N,v) \supseteq C(N,v)$ for all games $(N,v) \in \Gamma_c$ and $\Phi(N,v) \supset C(N,v)$ for some game $(N,v) \in \Gamma_c$.

In other words, there is a super-solution of the core that satisfies the four axioms except for the SIM-reduced game structure. Moreover, this means that those four axioms cannot be used to characterize any subsolution of the core for any possible game structure.

We obtain the characterization due to Funaki [4] as an immediate corollary of Theorem 3.2. Let us call the consistency property with respect to the SIM-reduced game structure *SIM-Consistency* (SIM-$CONS$).

Corollary 3.2 (Funaki [4]). *A solution Φ on Γ_c satisfies NE, PO, SIM-$CONS$, and SGR if and only if $\Phi = C$.*

Proof of Theorem 3.2. Funaki [4] first proved that the core is the unique solution Φ on Γ_c that satisfies NE, PO, SIM-$CONS$, and SGR. Here we give an alternative simple proof of this result by applying Lemma 3.2 to the SIM-reduced game structure. Let $\mathcal{S}(n) = \{1, 2, \dots, n-2\}$ for all $n \in \mathcal{N}$ with $n \geq 3$. Clearly, PO and SGR together imply IR^2. We prove that SGR implies Condition $\mathcal{S}$. Let $(N,v) \in \Gamma_c$ with $|N| \geq 3$, $j \in N$, and $x \in \Phi(N,v)$ be given. Since $\mathcal{S}(|N|) = \{1, 2, \dots, |N|-2\}$, $1 \in \mathcal{S}(|N|)$ and $K(|N|) = \{s \mid 2 \leq s \leq n-1,$ $s-1 \in \mathcal{S}(|N|),\ s \notin \mathcal{S}(|N|)\} = \{|N|-1\}$. If Φ satisfies SGR, then $x(T) \geq v(T)$ for all $T \in \{S \subseteq N \mid |S| = |N|-1\} = \{S \subseteq N \mid S \neq N$ and $|S| \in K(|N|)\}$, so that Φ satisfies Condition $\mathcal{S}$. By Lemma 3.2, we have the desired result.

Take any reduced game structure $\mathcal{S} \in \Sigma$ other than the SIM-reduced game structure. We show that there is a solution Φ on Γ_c that satisfies NE, PO, $\mathcal{S}$-$CONS$, and SGR, but $\Phi \neq C$. First of all, notice that by the condition (1), for every $\mathcal{S} \in \Sigma$, either (i) $1 \in \mathcal{S}(n)$ for all $n \in \mathcal{N}$ with $n \geq 3$, or (ii) $1 \notin \mathcal{S}(n)$ for all $n \in \mathcal{N}$ with $n \geq 3$. Let us consider the first case (i). Define the solution Φ as follows: for each $(N,v) \in \Gamma_c$, $\Phi(N,v) \equiv \{x \in PO(N,v) \cap IR(N,v) \mid x(T) \geq v(T)$ for all $T \subseteq N$ such that $|T| = n-1$ and $|T| \notin K(n) \setminus \{n-1\}\}$, where $n = |N|$. First, we prove that $\Phi \neq C$. Since $\mathcal{S}$ is different from the SIM-reduced game structure, $\mathcal{S}(n^*) \neq \{1, 2, \dots, n^*-2\}$ for some $n^* \in \mathcal{N}$ with $n^* \geq 4$. (If $n = 3$ and $1 \in \mathcal{S}(n)$, then there is only one possible $\mathcal{S}(n)$, that is, $\mathcal{S}(3) = \{1\}$.) We show that $K(n^*) \setminus \{n^*-1\} = \{s \mid 2 \leq s \leq n^*-2, s-1 \in \mathcal{S}(n^*), s \notin \mathcal{S}(n^*)\} \neq \emptyset$. Suppose that $K(n^*) \setminus \{n^*-1\} = \emptyset$. Then $1 \in \mathcal{S}(n^*)$ implies $2 \in \mathcal{S}(n^*)$. Since $2 \in \mathcal{S}(n^*)$ and $K(n^*) \setminus \{n^*-1\} = \emptyset$, it follows that $3 \in \mathcal{S}(n^*)$, and so on. Finally, since $n^*-3 \in \mathcal{S}(n^*)$ and $K(n^*) \setminus \{n^*-1\} = \emptyset$, it follows that $n^*-2 \in \mathcal{S}(n^*)$. Therefore, $\mathcal{S}(n^*) = \{1, 2, \dots, n^*-2\}$, which is a contradiction. Accordingly, $\Phi \neq C$.

Clearly Φ satisfies NE, PO, and SGR. We show that it also satisfies $\mathcal{S}$-$CONS$. Let $(N,v) \in \Gamma_c$ with $n = |N|$, $j \in N$, and $x \in \Phi(N,v)$ be given. Note that

$$\{1, 2, \dots, n-1\} \setminus K(n) = \{s \mid 1 \leq s \leq n-1, s \in \mathcal{S}(n)\}$$

$$\cup \{s \in \mathcal{N} \mid 1 \leq s \leq n-1, s-1 \notin \mathcal{S}(n), s \notin \mathcal{S}(n)\}.$$

Since $x \in \Phi(N,v)$, $x(T) \geq v(T)$ for all $T \subset N$ such that $|T| \in \{n-1\} \cup \{1,\ldots,n-1\} \setminus K(n)$. Consider the $\mathcal{S}$ -reduced game $(N \setminus \{j\}, v_x^{\mathcal{S}})$. By using an argument similar to that in the proof of Theorem 3.1, it is easy to prove that $x(T) \geq v_x^{\mathcal{S}}(T)$ for $T = N \setminus \{j\}$ and any subset $T \subset N \setminus \{j\}$ such that

$$\begin{aligned}|T| &\in \{1,2,\ldots,n-2\} \setminus K(n-1)\\ &= \{s \mid 1 \leq s \leq n-2, s \in \mathcal{S}(n-1)\}\\ &\quad \cup \{s \in \mathcal{N} | 1 \leq s \leq n-2, s-1 \notin \mathcal{S}(n-1), s \notin \mathcal{S}(n-1)\}.\end{aligned}$$

Next consider the second case (ii): $1 \notin \mathcal{S}(n)$ for all $n \in \mathcal{N}$ with $n \geq 3$. Define the solution Φ as follows: for each $(N,v) \in \Gamma_c$,

$\Phi(N,v) \equiv \{x \in PO(N,v) \mid x(T) \geq v(T)$ for all $T \subseteq N$ such that $|T| = n-1$ and $|T| \notin \{1\} \cup K(n) \setminus \{n-1\}\}$, where $n = |N|$. Clearly, $\Phi \neq C$ and Φ satisfies NE, PO, and SGR. We prove that it also satisfies $\mathcal{S}$-$CONS$. Let $(N,v) \in \Gamma_c$ with $n = |N|$, $j \in N$, and $x \in \Phi(N,v)$ be given. Note that

$$\begin{aligned}\{2,\ldots,n-1\} \setminus K(n) &= \{s \mid 2 \leq s \leq n-1, s \in \mathcal{S}(n)\}\\ &\quad \cup \{s \in \mathcal{N} \mid 2 \leq s \leq n-1, s-1 \notin \mathcal{S}(n), s \notin \mathcal{S}(n)\}.\end{aligned}$$

Since $x \in \Phi(N,v)$, $x(T) \geq v(T)$ for all $T \subset N$ such that $|T| \in \{n-1\} \cup \{2,\ldots,n-1\} \setminus K(n)$. Consider the $\mathcal{S}$-reduced game $(N \setminus \{j\}, v_x^{\mathcal{S}})$. By using an argument similar to that in the proof of Theorem 3.1, it is easy to prove that $x(T) \geq v_x^{\mathcal{S}}(T)$ for $T = N \setminus \{j\}$ and any subset $T \subset N \setminus \{j\}$ such that

$$\begin{aligned}|T| &\in \{2,\ldots,n-2\} \setminus K(n-1)\\ &= \{s | 2 \leq s \leq n-2, s \in \mathcal{S}(n-1)\}\\ &\quad \cup \{s \in \mathcal{N} | 2 \leq s \leq n-2, s-1 \notin \mathcal{S}(n-1), s \notin \mathcal{S}(n-1)\}. \qquad \square\end{aligned}$$

4 Concluding Remarks

We have considered a class of reduced game structures that are dependent only on the cardinality of a set of players. However, this class of reduced game structures Σ does not contain the reduced game structure due to Davis and Maschler [2]. We give a more general definition of reduced game structures including the reduced game due to Davis and Maschler as a special case. In the following definition, a reduced game structure depends on a set of players N, a characteristic function v, a player $j \in N$, and a payoff vector x. For a finite set of players $N \subset \mathcal{N}$, let $\mathcal{P}^N \equiv \{S \subseteq N \mid S \neq N, \emptyset\}$, i.e., $\mathcal{P}^N$ is the set of all proper subcoalitions in N.

Definition 4.1. A *reduced game structure* is a mapping $\mathcal{S}$ from $\{(N,v,j,x) \mid (N,\ v) \in \Gamma$ with $|N| \geq 3, j \in N, x \in \mathcal{R}^N\}$ into $\cup_{|N|\geq 3}\mathcal{P}^N$ such that for all

$(N, v) \in \Gamma$ with $|N| \geq 3$ and all $(j, x) \in N \times \mathcal{R}^N$, $\mathcal{S}(N, v, j, x) \subseteq \mathcal{P}^{N\setminus\{j\}}$. Given a reduced game structure $\mathcal{S}$, a TU game $(N, v) \in \Gamma$, a player $j \in N$, and a payoff vector $x \in \mathcal{R}^N$, the $\mathcal{S}$-*reduced game with respect to* j *and* x is the game $(N \setminus \{j\}, v_x^{\mathcal{S}})$ where

$$v_x^{\mathcal{S}}(S) = \begin{cases} 0 & \text{if } S = \emptyset, \\ v(N) - x_j & \text{if } S = N \setminus \{j\}, \\ v(S) & \text{if } S \in \mathcal{S}(N, v, j, x), \\ v(S \cup \{j\}) - x_j & \text{otherwise.} \end{cases}$$

Let Σ^* be the class of all possible such reduced game structures.

Example 4.1. For all $(N, v) \in \Gamma$ and all $(j, x) \in N \times \mathcal{R}^N$, let $\mathcal{S}(N, v, j, x) = \{S \in \mathcal{P}^{N\setminus\{j\}} \mid v(S) \geq v(S \cup \{j\}) - x_j\}$. Then we have the reduced game introduced by Davis and Maschler [2]:

$$v_x^{\mathcal{S}}(S) = \begin{cases} 0 & \text{if } S = \emptyset, \\ v(N) - x_j & \text{if } S = N \setminus \{j\}, \\ \max\{v(S \cup \{j\}) - x_j, v(S)\} & \text{otherwise.} \end{cases}$$

We call this reduced game structure $\mathcal{S} \in \Sigma^*$ the *DM-reduced game structure.*

Our uniqueness results no longer hold for the class of reduced game structures Σ^* that is larger than Σ. Let us consider the following reduced game structure that belongs to Σ^*, but not to Σ.

Example 4.2. Given $N \subset \mathcal{N}$, denote the set of players by $N = \{1, 2, \ldots, n\}$. Consider a reduced game structure $\mathcal{S}^1 \in \Sigma^*$ defined as follows: for all $(N, v) \in \Gamma$ with $|N| \geq 3$ and all $(j, x) \in N \times \mathcal{R}^N$, $\mathcal{S}^1(N, v, j, x) = \{\{j + 1\}\}$, where we regard player $n + 1$ as player 1 if $j = n$. Then the reduced game $(N \setminus \{j\}, v_x^{\mathcal{S}^1})$ takes the following form:

$$v_x^{\mathcal{S}^1}(S) = \begin{cases} 0 & \text{if } S = \emptyset, \\ v(\{j + 1\}) & \text{if } S = \{j + 1\}, \\ v(S \cup \{j\}) - x_j & \text{otherwise.} \end{cases}$$

Proposition 4.1. *A solution* Φ *on* Γ_c *satisfies* NE, IR^2*, and* $\mathcal{S}^1$*-CONS if and only if* $\Phi = C$*.*

The proof of Proposition 4.1 is in the Appendix.

Since IR implies IR^2, it follows from Proposition 4.1 that the core is the only solution satisfying NE, IR, and $\mathcal{S}^1$-$CONS$, that is, there exists a reduced game structure, $\mathcal{S}^1 \in \Sigma^* \setminus \Sigma$, other than the M-reduced game structure such that the core is characterized by non-emptiness, individual rationality, and consistency with respect to it.

Moreover, since PO and SGR together imply IR^2, it follows from Proposition 4.1 that the core is the only solution satisfying NE, PO, SGR, and $\mathcal{S}^1$-$CONS$, that is, there exists a reduced game structure, $\mathcal{S}^1 \in \Sigma^* \backslash \Sigma$, other than the SIM-reduced game structure such that the core is characterized by non-emptiness, Pareto optimality, sub-grand rationality, and consistency with respect to it.

For a set of axioms that is different from those we examine, however, a uniqueness result on a reduced game in Σ^* might hold. For example, Peleg [7] employs the DM-reduced game structure, and he proves that the core is the only solution satisfying consistency with respect to it, non-emptiness, individual rationality, and super-additivity. It is an open question to investigate whether there are other reduced game structures in Σ^* that can be used to characterize the core by the four axioms.[3]

5 Appendix

Proof of Lemma 3.2. By Proposition 2.1, the core satisfies $\mathcal{S}$-$CONS$. Clearly, the core satisfies NE, IR^2, and Condition $\mathcal{S}$ on Γ_c.

Next we prove that if Φ satisfies NE, IR^2, $\mathcal{S}$-$CONS$, and Condition $\mathcal{S}$, then $\Phi = C$. The proof consists of three steps:

Claim 5.1. If Φ satisfies $IR^2, \mathcal{S}$-$CONS$, and Condition $\mathcal{S}$, then $\Phi(N, v) \subseteq C(N, v)$ for every $(N, v) \in \Gamma_c$.

Proof. Since IR^2 implies IR^1, it follows from Lemma 3.1, that Φ satisfies PO. If $|N| = 1$, then $\Phi(N, v) \subseteq C(N, v)$ by PO. If $|N| = 2$, then $\Phi(N, v) \subseteq C(N, v)$ by IR^2 and PO. Assume that $\Phi(N, v) \subseteq C(N, v)$ for any $(N, v) \in \Gamma_c$ with $|N| = k$ where $k \geq 2$. Consider any $(M, u) \in \Gamma_c$ with $|M| = k + 1$. Pick any $x \in \Phi(M, u)$ and any $j \in M$. By $\mathcal{S}$-$CONS$, $x_{M \setminus \{j\}} \in \Phi(M \setminus \{j\}, u_x^{\mathcal{S}})$. By the assumption, $\Phi(M \setminus \{j\}, u_x^{\mathcal{S}}) \subseteq C(M \setminus \{j\}, u_x^{\mathcal{S}})$. Therefore, for any $S \subseteq M \setminus \{j\}$ with $|S| \in \mathcal{S}(|M|)$, $x(S) \geq u_x^{\mathcal{S}}(S) = u(S)$. Moreover, for any $S \subseteq M \setminus \{j\}$ with $|S| \notin \mathcal{S}(|M|)$, $x(S) \geq u_x^{\mathcal{S}}(S) = u(S \cup \{j\}) - x_j$, and hence $x(S \cup \{j\}) \geq u(S \cup \{j\})$. That is, $x(S) \geq u(S)$ for any S with $|S| \in \{t \in \mathcal{N} \mid 2 \leq t \leq |M| - 1, t - 1 \notin \mathcal{S}(|M|)\ \}$. Thus, $x(S) \geq u(S)$ for any $S \in \{S \subseteq M \setminus \{j\} \mid |S| \in \mathcal{S}(|M|)\} \cup \{S \mid |S| \in \{t \in \mathcal{N} \mid 2 \leq t \leq |M| - 1, t - 1 \notin \mathcal{S}(|M|)\ \}$. Since Φ satisfies Condition $\mathcal{S}$, it follows that $x(S) \geq u(S)$ for any $S \in \{S \subseteq M \mid S \neq M, \emptyset\}$. This together with PO implies that $x \in C(M, u)$. □

[3]Funaki and Yamato [5] provide an axiomatization of the core for any given reduced game structure in Σ^*. The axiomatization of the core due to Peleg [7] can be obtained as a corollary of their axiomatization.

Claim 5.2. Let $(N,v) \in \Gamma_c$ and $x \in C(N,v)$ be given. Then there exists $(M,u) \in \Gamma_c$ such that (i) $M = N \cup \{i\}$ where $i \notin N$; (ii) $C(M,u) = \{(x,0)\}$; and (iii) $(N, u_y^{\mathcal{S}}) = (N,v)$ where $y \equiv (x,0) \in \mathcal{R}^{N\cup\{i\}}$.

Proof. Define $(M,u) \in \Gamma_c$ as follows: $u(\{i\}) = 0, u(S) = v(S)$ and $u(S\cup\{i\}) = x(S)$ for $S \in \mathcal{S}(|M|)$, and $u(S) = x(S)$ and $u(S\cup\{i\}) = v(S)$ for $S \in \{S \subseteq M \setminus \{i\} \mid S \neq M\setminus\{i\}, \emptyset$ and $|S| \notin \mathcal{S}(|M|)\}$. First we show that $y \in C(M,u)$. Clearly, $u(M) = v(N) = x(N) = y(M)$ and $u(\{i\}) = y_i$. Moreover, if $|S| \in \mathcal{S}(|M|)$, then $y(S) = x(S) \geq v(S) = u(S)$ and $y(S \cup \{i\}) = x(S) = u(S \cup \{i\})$. Further, if $S \in \{S \subseteq M \setminus \{i\} \mid S \neq M \setminus \{i\}, \emptyset$ and $|S| \notin \mathcal{S}(|M|)\}$, then $y(S) = x(S) = u(S)$ and $y(S \cup \{i\}) = x(S) \geq v(S) = u(S \cup \{i\})$. Thus $y \in C(M,u)$. Next we prove that $C(M,u) = \{y\}$. Let $z \in C(M,u)$ be given. First, we claim that $z_i = 0$. If $|N| \in \mathcal{S}(|M|)$, then $u(N) = v(N)$. If $|N| \notin \mathcal{S}(|M|)$, then $u(N) = x(N) = v(N)$. Therefore, $u(N) = v(N)$. Since $z \in C(M,u), z(N) \geq u(N) = v(N) = u(M) = z(N) + z_i$ and $z_i \geq u(\{i\}) = 0$. Hence $z_i = y_i = 0$.

Take any $j \in N$. If $1 \in \mathcal{S}(|M|)$, then $z_j = z_i + z_j \geq u(\{i,j\}) = y_j$. If $1 \notin \mathcal{S}(|M|)$, then $z_j \geq u(\{j\}) = y_j$. Therefore, $z_j \geq y_j$ for all $j \in N$. Further, $z(N) = z(M) = u(M) = v(N) = y(N)$. These imply that $z_j = y_j$ for all $j \in N$. Therefore, $z = y$.

Finally, we show that $(N, u_y^{\mathcal{S}}) = (N,v)$. Let $S \subset N, S \neq \emptyset$ be given. If $|S| \in \mathcal{S}(|M|)$, then $u_y^{\mathcal{S}}(S) = u(S) = v(S)$. If $|S| \notin \mathcal{S}(|M|)$, then $u_y^{\mathcal{S}}(S) = u(S \cup \{i\}) - y_i = u(S \cup \{i\}) = v(S)$. Moreover, $u_y^{\mathcal{S}}(N) = u(M) - y_i = u(M) = v(N)$. □

Let Φ be a solution on Γ_c satisfying $NE, IR^2, \mathcal{S}$-$CONS$, and Condition $\mathcal{S}$. Let $(N,v) \in \Gamma_c$ and $x \in C(N,v)$ be given. By Claim 5.2, there exists a TU game (M,u) satisfying (i), (ii), and (iii). Claim 5.1 and NE imply that $\Phi(M,u) = C(M,u) = \{(x,0)\}$. By $\mathcal{S}$-$CONS$, $x \in \Phi(N, u_y^{\mathcal{S}}) = \Phi(N,v)$. Thus, $C(N,v) \subseteq \Phi(N,v)$. This fact and Claim 5.1 together imply that $\Phi(N,v) = C(N,v)$. □

Proof of Proposition 4.1. It is easy to check that the core satisfies NE, IR^2, and $\mathcal{S}^1$-$CONS$ on Γ_c.

Next we prove that if Φ satisfies NE, IR^2, and $\mathcal{S}^1$-$CONS$, then $\Phi = C$. In the following, we denote the reduced game structure $\mathcal{S}^1(N,v,j,x)$ simply by $\mathcal{S}^1(N,j)$ for all $(N,v) \in \Gamma$ with $|N| \geq 3$ and all $(j,x) \in N \times \mathcal{R}^N$, since $\mathcal{S}^1$ is dependent on a set of players N and a player $j \in N$, but independent of a characteristic function v and a payoff vector x. The proof consists of three steps:

Claim 5.3. If Φ satisfies $IR^2$and $\mathcal{S}^1$-$CONS$, then $\Phi(N,v) \subseteq C(N,v)$ for every $(N,v) \in \Gamma_c$.

Proof. Since IR^2 implies IR^1 , it follows from Lemma 3.1 that Φ satisfies PO. If $|N| = 1$, then $\Phi(N,v) \subseteq C(N,v)$ by PO. If $|N| = 2$, then $\Phi(N,v) \subseteq C(N,v)$

by IR^2 and PO. Assume that $\Phi(N,v) \subseteq C(N,v)$ for any $(N,v) \in \Gamma_c$ with $|N| = k$ where $k \geq 2$. Consider any $(M,u) \in \Gamma_c$ with $|M| = k+1$. Pick any $x \in \Phi(M,u)$ and any $j \in M$. By $\mathcal{S}^1$-$CONS$, $x_{M\setminus\{j\}} \in \Phi(M \setminus \{j\}, u_x^{\mathcal{S}^1})$. By the assumption, $\Phi(M \setminus \{j\}, u_x^{\mathcal{S}^1}) \subseteq C(M \setminus \{j\}, u_x^{\mathcal{S}^1})$. Therefore, for any $S \subseteq M \setminus \{j\}$ with $S \in \mathcal{S}^1(M,j)$, $x(S) \geq u_x^{\mathcal{S}}(S) = u(S)$. Moreover, for any $S \subseteq M \setminus \{j\}$ with $S \notin \mathcal{S}^1(M,j)$, $x(S) \geq u_x^{\mathcal{S}}(S) = u(S \cup \{j\}) - x_j$, and hence $x(S \cup \{j\}) \geq u(S \cup \{j\})$. That is, $x(S) \geq u(S)$ for any $S \in \mathcal{T}(M,j;\mathcal{S}^1) \equiv \{S \in \mathcal{P}^M \mid S \ni j, S \neq \{j\}, S \setminus \{j\} \notin \mathcal{S}^1(M,j)\}$. Thus, $x(S) \geq u(S)$ for any $S \in \cup_{j\in M}(\mathcal{S}^1(M,j) \cup \mathcal{T}(M,j;\mathcal{S}^1))$. Note that $\cup_{j\in M}\mathcal{S}^1(M,j) = \{S \in \mathcal{P}^M \mid |S| = 1\}$. Also, $\cup_{j\in M}\mathcal{T}(M,j;\mathcal{S}^1) = \{S \in \mathcal{P}^M \mid |S| \geq 2\}$. These imply $\cup_{j\in M}(\mathcal{S}^1(M,j) \cup \mathcal{T}(M,j;\mathcal{S}^1)) = \mathcal{P}^M$. Therefore, $x(S) \geq u(S)$ for any $S \in \mathcal{P}^M$. This together with PO implies that $x \in C(M,u)$. □

Claim 5.4. Let $(N,v) \in \Gamma_c$ and $x \in C(N,v)$ be given. Then there exists $(M,u) \in \Gamma_c$ such that (i) $M = N \cup \{i\}$ where $i \notin N$; (ii) $C(M,u) = \{(x,0)\}$; and (iii) $(N, u_y^{\mathcal{S}^1}) = (N,v)$ where $y \equiv (x,0) \in \mathcal{R}^{N\cup\{i\}}$.

Proof. Define $(M,u) \in \Gamma_c$ as follows: $u(\{i\}) = 0$, $u(S) = v(S)$, and $u(S \cup \{i\}) = x(S)$ for $S \in \mathcal{S}^1(M,i)$, and $u(S) = x(S)$ and $u(S \cup \{i\}) = v(S)$ for $S \in \mathcal{P}^{M\setminus\{i\}} \setminus \mathcal{S}^1(M,i)$. First we show that $y \in C(M,u)$. Clearly, $u(M) = v(N) = x(N) = y(M)$ and $u(\{i\}) = y_i$. Moreover, if $S \in \mathcal{S}^1(M,i)$, then $y(S) = x(S) \geq v(S) = u(S)$ and $y(S \cup \{i\}) = x(S) = u(S \cup \{i\})$. Further, if $S \in \mathcal{P}^{M\setminus\{i\}} \setminus \mathcal{S}^1(M,i)$, then $y(S) = x(S) = u(S)$ and $y(S \cup \{i\}) = x(S) \geq v(S) = u(S \cup \{i\})$. Thus $y \in C(M,u)$. Next we prove that $C(M,u) = \{y\}$. Let $z \in C(M,u)$ be given. First, we claim that $z_i = 0$. If $N \in \mathcal{S}^1(M,i)$, then $u(N) = v(N)$. If $N \notin \mathcal{S}^1(M,i)$, then $u(N) = x(N) = v(N)$. Therefore, $u(N) = v(N)$. Since $z \in C(M,u), z(N) \geq u(N) = v(N) = u(M) = z(N) + z_i$, and $z_i \geq u(\{i\}) = 0$. Hence $z_i = y_i = 0$.

Take any $j \in N$. If $\{j\} \in \mathcal{S}^1(M,i)$, then $z_j = z_i + z_j \geq u(\{i,j\}) = y_j$. If $\{j\} \notin \mathcal{S}^1(M,i)$, then $z_j \geq u(\{j\}) = y_j$. Therefore, $z_j \geq y_j$ for all $j \in N$. Further, $z(N) = z(M) = u(M) = v(N) = y(N)$. These imply that $z_j = y_j$ for all $j \in N$. Therefore, $z = y$.

Finally, we show that $(N, u_y^{\mathcal{S}^1}) = (N,v)$. Let $S \subset N, S \neq \emptyset$ be given. If $S \in \mathcal{S}^1(M,i)$, then $u_y^{\mathcal{S}^1}(S) = u(S) = v(S)$. If $S \notin \mathcal{S}^1(M,i)$, then $u_y^{\mathcal{S}^1}(S) = u(S \cup \{i\}) - y_i = u(S \cup \{i\}) = v(S)$. Moreover, $u_y^{\mathcal{S}^1}(N) = u(M) - y_i = u(M) = v(N)$. □

Let Φ be a solution on Γ_c satisfying NE, IR^2, and $\mathcal{S}^1$-$CONS$. Let $(N,v) \in \Gamma_c$ and $x \in C(N,v)$ be given. By Claim 5.4, there exists a TU game (M,u) satisfying (i), (ii), and (iii). Claim 5.3 and NE imply that $\Phi(M,u) = C(M,u) = \{(x,0)\}$. By $\mathcal{S}^1$-$CONS$, $x \in \Phi(N, u_y^{\mathcal{S}^1}) = \Phi(N,v)$. Thus, $C(N,v) \subseteq \Phi(N,v)$. This together with Claim 5.3 implies that $\Phi(N,v) = C(N,v)$. □

REFERENCES

[1] Aumann R.J. and Drèze J.H., Cooperative games with coalitional structures, *International Journal of Game Theory*, 3, 2127–2237, 1974.

[2] Davis M. and Maschler M., The kernel of a cooperative game, *Naval Research Logistics Quarterly*, 12, 223–259, 1965.

[3] Driessen T.S.H., A survey of consistency properties in cooperative game theory, *SIAM Review*, 33, 43–59, 1991.

[4] Funaki Y., Dual axiomatizations of solutions of cooperative games, mimeo, 1995.

[5] Funaki Y. and Yamato T., The core and consistency properties: A general characterization, *International Game Theory Review*, 3, 175–187, 2001.

[6] Moulin H., The separability axiom and equal-sharing methods, *Journal of Economic Theory*, 36, 120–148, 1985.

[7] Peleg B., On the reduced game property and its converse, *International Journal of Game Theory*, 15, 187–200, 1986; Correction, *International Journal of Game Theory*, 16, 290, 1987.

[8] Schmeidler D., The nucleolus of a characteristic function game, *SIAM Journal of Applied Mathematics*, 17, 1163–1170, 1969.

[9] Sobolev A.I., The characterization of optimality principles in cooperative games by functional equations, in N.N. Vorobev (ed.), *Mathematical Methods in the Social Sciences*, 6, 94–151, Vilnius University, Lithuania, 1975 (in Russian).

[10] Tadenuma K., Reduced games, consistency, and the core, *International Journal of Game Theory*, 20, 325–334, 1992.

[11] Thomson W., The consistency principle in economics and game theory, in T. Ichiishi, A. Neymand and Tauman (eds.), *Game Theory and Applications*, Academic Press, 187–215, 1990.

The Formation of Adaptive Coalitions

Jürgen Scheffran
University of Illinois at Urbana-Champaign
Champaign, Illinois 61820, USA
scheffra@uiuc.edu

Abstract

This chapter develops a framework for analyzing the interaction between individual players (actors) and collective players (coalitions) who mutually adapt the allocation of investment to their values and each other's decisions. The dynamic process of coalition formation can be described by a coupled evolutionary game of allocation controls. Potential fields of applications are outlined, and an example analyzing the management of energy and carbon emissions is discussed in more detail.

1 Introduction

The theory of coalitions is well established in cooperative game theory and appropriate for explaining a variety of phenomena dealing with the distribution of coalition values and the stability of coalitions. In the following, we go beyond the established theory to analyze dynamically changing situations and the formation of coalitions with mixed individual contributions changing over time. Related fields are cooperative dynamic games (e.g., [8]), evolutionary games [10] and the concept of fuzzy coalitions, extending the assumption of strict membership by the level of participation of individual players in a coalition [1,2,16]. This takes note of the experience that there can be different levels of support for different coalitions and a gradual shift between coalitions, leading to the increase or decline of coalitions in the course of time.

In this regard, coalition formation can be defined as a process in which two or more actors adjust their coalition support to the expected benefit from this support. Each actor can join several coalitions by allocating a certain amount of resources to each of them (mixed membership). Coalitions acquire resources from the individual actors and invest them into actions to achieve coalition benefits that are distributed to the individuals. Issue-based coalitions can be formed to set targets and seek agreement on issues of common interest to undertake joint actions. Assuming that coalitions can decide and act themselves and thus adapt to changing circumstances, a dynamic framework for adaptive coalition games is presented, building on previous efforts to analyze the interaction and cooperation in dynamic games [17,14,13]. This chapter introduces a new concept of adaptive coalitions that links multiplayer interaction, cooperative dynamic games, fuzzy coalition theory and evolutionary games. It provides a framework to analyze phenomena of dynamic coalition formation in different fields of application, using deeper mathematical analysis and application of solution concepts.

2 Interaction between Actors and Coalitions

The framework of dynamic games and adaptive coalition formation includes two types of players (see Figure 1):

- *Individual actors* (agents, players) A_i with $i = 1, \ldots, n$ at a given time have an amount of resources $C_i^+(t)$ under their control which they can invest within the resource limits directly to actions and/or to coalitions to take joint action with others.
- *Coalitions* (collective players) $\mathcal{A}^I$ with $I = 1, \ldots, N$ use the resources acquired from the individual actors to allocate them to joint actions which are evaluated by the individuals to reallocate their resources to the coalitions.

We assume that the actors allocate a fraction p_i^I of their invested resources (costs) C_i to each of the coalitions $\mathcal{A}^I$, such that $\sum_I^N p_i^I = 1$, which may include

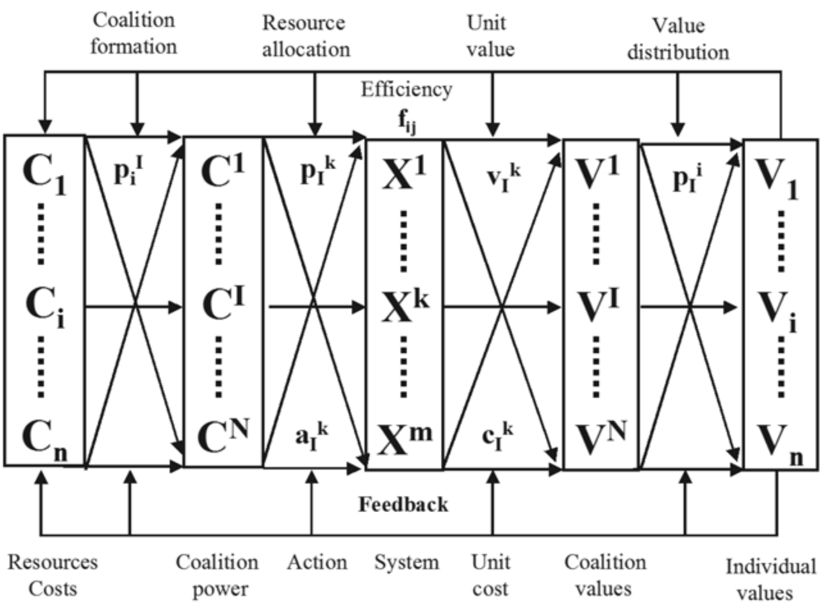


Figure 1: The feedback cycle of multiplayer interaction and coalition formation.

the coalition of each actor acting alone. Each coalition can use its accumulated resources

$$C^I = \sum_{j=1}^{n} p_j^I \cdot C_j$$

to act on various issues $k_I = 1, \ldots, m_I$.[1] The impact of coalition actions (measured by the change of system variables)

$$a_I^k = \frac{p_I^k C^I}{c_I^k}$$

depends on the unit costs c_I^k of each action, and the resource allocation p_I^k of each coalition $\mathcal{A}^I$. Then the value of actor A_i acquired from coalition actions a_I^k is:

$$V_i = \sum_{I=1}^{N} \sum_{k=1}^{m_I} v_i^k a_I^k = \sum_{I=1}^{N} \sum_{k=1}^{m_I} \sum_{j=1}^{n} \frac{v_i^k}{c_I^k} p_j^I p_I^k C_j, \tag{1}$$

where v_i^k is the unit value of actor A_i for action a_I^k of coalition $\mathcal{A}^I$.

Using $f_i^I = \sum_k^{m_I} (v_i^k / c_I^k) p_I^k$ as an indicator for the benefit-cost efficiency of coalition $\mathcal{A}^I$ for actor A_i with regard to the set of coalition actions $k = 1, \ldots, m_I$, the problem can be represented as an interaction between the individual actors, where the value of each actor $(i = 1, \ldots, n)$

$$V_i = \sum_{j=1}^{n} \sum_{I=1}^{N} p_j^I f_i^I C_j = \sum_{j=1}^{n} f_{ij} C_j$$

is influenced by the invested resources C_j of all other actors. The mutual impacts $f_{ij} = \sum_{I=1}^{N} p_j^I f_i^I$ depend on the allocation and efficiency associated with the coalitions.

Coalitions can allocate their resources to various action paths and may have their own value functions to evaluate them. An important factor for a coalition are the total resources C^I under its control which determine its power compared to other coalitions and also with regard to the actors. In addition, a coalition may be interested in the value generated which may be a function of its own resources and the resources of other coalitions. A value generated by a coalition can be either accumulated or distributed to the actors which may be required to ensure continued support of the actors for the coalitions and to avoid their withdrawal. It seems reasonable that an actor tends to support a coalition as long as he/she is better off with it rather than with another coalition (in particular, compared to acting alone).

[1] Where possible, lower indices are used. Upper indices are required for two or more sets of indices. To simplify notation we use k rather than k_I if index I is already given.

3 The Adaptation Process

3.1 Levels of Control

Both actors and coalitions can take decisions on resource allocation, and the interaction between them can be treated as an interconnected decision-making problem, considering three levels of control:

(1) **Actors allocate to coalitions:** Individual actors A_i $(i = 1, \dots , n)$ control their investment flows $0 \leq C_i \leq C_i^+$ (with upper limit C_i^+) as well as their allocation $0 \leq p_i^I \leq 1$ to coalitions $\mathcal{A}^I$ $(I = 1, \dots , N)$, which receive the investment $p_i^I C_i$.

(2) **Coalitions allocate to issues:** Coalition $\mathcal{A}^I$ $(I = 1, \dots , N)$ controls its resources C^I and the allocation $0 \leq p_I^k \leq 1$ to the coalition actions $k = 1, \dots , m_I$ with $\sum_k^{m_I} p_I^k = 1$.

(3) **Coalition values are distributed to actors:** Coalition actions a_I^k generate unit values v_i^k for the individual players, either directly as a result of the action taken by a coalition and/or indirectly via a coalition value v_I^k which is then distributed to the actors A_i. Their unit value $v_i^k = p_I^i v_I^k$ depends on the fraction p_I^i assigned to actor A_i. For $\sum_I^N p_I^i = 1$ the coalition value is completely distributed to the actors and nothing is spent on other costs, an assumption that can be modified. For a given p_I^i actor A_i receives a total share $V_i^I = p_I^i V^I$ of coalition value which can be distributed in different ways to the individual actors.

With these three levels of control the interaction can be quite complicated for a larger number of actors and coalitions. The dynamics will be essentially determined by the rules for selecting or changing the allocation of resources p_i^I to a particular coalition, and from the coalitions to the actions p_I^k which are evaluated by the individual actors. The distribution of coalition values to the actors is subject to negotiations between both, a field where established cooperative game theory comes into play. Whether a distribution is seen as fair or effective is widely debated. A special rule, which satisfies the condition $\sum_i p_I^i = 1$, is to have the fraction proportionate to the resource input of actor A_i:

$$p_I^i = \frac{p_i^I C_i}{C^I}.$$

If actor A_i incrementally changes the allocation by $dp_i^I = -dp_i^J$ from coalition $\mathcal{A}^I$ to coalition $\mathcal{A}^J$, the value functions incrementally change by

$$dV_i = \frac{\partial V_i}{\partial p_i^I} dp_i^I + \frac{\partial V_i}{\partial p_i^J} dp_i^J = (f_i^J - f_i^I) C_i \, dp_i^J .$$

For a complete transition from coalition $\mathcal{A}^I$ to $\mathcal{A}^J$ for actor A_i we have $dp_i^I = -1 = -dp_i^J$ and write $\Delta V_i = (f_i^J - f_i^I)\, C_i$. To maximize value, actor A_i would adapt to the coalition $\mathcal{A}^J$ with the highest efficiency f_i^J, either in a single optimizing time step or gradually through an adaptation process.

3.2 An Evolutionary Game

Such an adaptation process can be well represented by a dynamic system, analogous to the replicator equation in an evolutionary dynamic game for p_i^I:

$$\Delta p_i^I(t) = \alpha_i\, p_i^I(t)\, (V_i^I(t) - \sum_{J=1}^{N} p_i^J(t)\, V_i^J(t)). \tag{2}$$

Here $V_i^I = \partial V_i/\partial p_i^I = \sum_k^{m_I} p_i^k v_i^k C_i/c_I^k$ is the value of actor A_i obtained from coalition $\mathcal{A}^I$ for $p_i^I = 1$ and the second term is the mixed value for the current allocation $p_i^J(t)$ over all coalitions $\mathcal{A}^J$. For continuous time, we have $\Delta p_i^I(t) = \dot{p}_i^I(t)$, for discrete time $\Delta p_i^I(t) = p_i^I(t+1) - p_i^I(t)$. In both cases, the dynamic system satisfies the condition $\sum_I \Delta p_i^I(t) = 0$ and thus $\sum_I p_i^I(t) = 1$ at all times. As in standard evolutionary games $\Delta p_i^I(t)$ increases with the "fitness" of a coalition $\mathcal{A}^I$ for actor A_i, compared to an average fitness. Thus those coalitions that better serve the actors' interests will grow more. The main difference is that we have individual players rather than populations and p_i^I represents a control variable (allocation), not a probability. Using the evolutionary game formalism allows us to apply the methodology from this field, including equilibrium and stability concepts.[2]

The adaptation of coalition actions can be modelled in a similar way. First of all, a coalition would be interested in increasing its power, measured by the sum of acquired resources $C^I = \sum_j p_j^I C_j$. Using the dynamical equation for Δp_j^I, the coalition power evolves according to

$$\Delta C^I(t) = \sum_j \frac{\partial C^I}{\partial p_j^I} \Delta p_j^I = \sum_j \alpha_j p_j^I(t) C_j (V_j^I(t) - \sum_{J=1}^{N} p_j^J(t) V_j^J(t)). \tag{3}$$

Thus, the power dynamics of coalitions is closely related to the values that these coalitions provide to the single actors, weighted by the resource input $p_j^I C_j$ that these actors provide to the coalitions.

Actions taken by coalitions depend on the values these actions generate, for the two cases mentioned before:

[2] The concept of adaptive coalitions can be integrated into a probabilistic coalition theory which makes the use of evolutionary games even more reasonable. This approach defines the probability that an actor joins a coalition on a particular position, which contributes to the aggregated probability that several actors join that coalition [5].

(1) Coalitions which generate a coalition value $V^I = \sum_{k=1}^{m_I} p_I^k V_I^k$ tend to take action k with a maximum of

$$V_I^k = \frac{\partial V^I}{\partial p_I^k} = \frac{v_I^k}{c_I^k} C^I,$$

where v_I^k is the unit coalition value. This again corresponds to an evolutionary dynamic game

$$\Delta p_I^k(t) = \alpha_I \, p_I^k(t) \, (V_I^k(t) - \sum_{l=1}^{m_I} p_I^l(t) \, V_I^l(t)). \tag{4}$$

(2) Coalitions which do not produce a coalition value can still affect the value of actor A_i with coalition actions k_I according to $V_i^{k_I} = (v_i^{k_I}/c_I^k)C^I$. Which action a coalition will take depends on an aggregated function $\bar{V}_I^k = \sum_i w_i^I V_i^{k_I}$ of the values of all actors A_i, weighted with a factor w_i^I according to the relevance for the coalition $\mathcal{A}^I$. In accordance with the previous case, the dynamics of coalition actions k_I can be represented by a replicator equation

$$\Delta p_I^k(t) = \alpha_I \, p_I^k(t) \, (\bar{V}_I^k(t) - \sum_{l=1}^{m_I} p_I^l(t) \, \bar{V}_I^l(t)). \tag{5}$$

Which weights $w_i^I(t)$ a coalition assigns to the value of each actor A_i depends on the coalition's preferences. A natural choice would take into account the coalition's self-interest in increasing its power ΔC^I which according to Equation (3) depends on the value and contribution of each actor. Thus with $w_i^I(t) = p_i^I(t) C_i / C^I$ a coalition would prefer those actors which provide the strongest support due to interest in the coalition actions.

With changing allocations according to the replicator equations, the induced value changes of the actors A_i $(i = 1, \dots, n)$ are

$$\Delta V_i(t) = \sum_j \sum_I V_{ij}^I \Delta p_j^I(t) + \sum_I \sum_k V_i^{k_I} \Delta p_I^k(t) \tag{6}$$

with $V_{ij}^I = \partial V_i / \partial p_j^I = f_i^I C_j$ and $V_i^{k_I} = \partial V_i / \partial p_I^k = v_i^k / c_I^k C^I$. This provides a system of differential/difference equations describing the evolution of resource allocation and values.

4 Limits to Growth of Coalitions

A coalition can potentially grow indefinitely, as long as the resource inputs and the coalition values have no limits. Under certain conditions, however, the size of a coalition may be limited if there are no incentives for further support from additional actors to join the coalition. Consider, for instance, a coalition value function which is quadratic in resources

$$V_I = a_I C_I + b_I (C_I)^2$$

with $a_I > 0$, b_I arbitrary contents. Assuming that actor A_i receives a share of the coalition value

$$V_i^I = \frac{p_i^I C_i}{C^I} \cdot V_I = p_i^I C_i (a_i + b_I C_I)$$

proportionate to the resource input to the coalition, then A_i is indifferent to further contributions to coalition $\mathcal{A}^I$ for the optimality condition

$$v_i^I = \frac{\partial V_i^I}{\partial p_i^I} = (b_I p_i^I C_i + a_I + b_I C_I) C_i = 0,$$

which leads to the optimal allocation

$$\bar{p}_i^I = \frac{a_I + b_I C_I}{-b_I C_i}.$$

We distinguish two cases:

(1) For $b_I > 0$ (increasing returns), we have $\bar{p}_i^I < 0$ and thus an increasing value $v_i^I > 0$ of actor A_i for $p_i^I > \bar{p}_i^I$ which is always the case.
(2) For $b_I < 0$ (decreasing returns), we have $\bar{p}_i^I > 0$ and thus an increasing value $v_i^I > 0$ of actor A_i for $p_i^I < \bar{p}_i^I$. The condition $0 \leq \bar{p}_i^I \leq 1$ corresponds to an interval for coalition power within which an optimal allocation for actor A_i exists:
$$\frac{a_I}{-b_I} - C_i \leq C_I \leq \frac{a_I}{-b_I}.$$

For low C_I actor A_i tends to increase support for coalition $\mathcal{A}^I$, for high C_I there is no incentive to do so. In between there is the optimal allocation threshold $\bar{p}_i^I$ which declines with coalition power C_I and individual resources C_i in the interval.

5 The Case of Two Coalitions and Two Issues

We now treat the special case of $i = 1, \ldots, n$ actors A_i and $I = 1, 2$ coalitions $\mathcal{A}^I$ which allocate their acquired resources C^I to $k = 1, 2$ issues. Both actors and coalitions have one control variable:

- Actors A_i control the allocation $p_i = p_i^2 = 1 - p_i^1$ to coalitions $\mathcal{A}^2$ and $\mathcal{A}^1$.
- Coalitions $\mathcal{A}^I$ control the allocation $p^I = p_I^2 = 1 - p_I^1$ to actions a_I^2 and a_I^1.

Then the two coalitions have resources $C^1 = (1 - p_1)C_1 + (1 - p_2)C_2$ and $C^2 = p_1C_1 + p_2C_2$. The value functions of $i = 1, \dots, n$ actors A_i are given as

$$V_i = f_i^1(p^1)\, C^1 + f_i^2(p^2)\, C^2 = f_{ii}C_i + f_{ij}C_j$$

with the efficiencies of the coalitions $f_i^I = \sum_{k=1}^2 p_I^k v_i^k / c_I^k$ and the mutual impacts

$$f_{ij} = p_j[f_i^2(p^2) - f_i^1(p^1)] + f_i^1(p^1) \qquad (i, j = 1, \dots, n).$$

Actor A_i tends to increase the allocation towards coalition $\mathcal{A}^2$ if its action is more efficient than for coalition $\mathcal{A}^1$

$$\frac{\partial V_i}{\partial p_i} = (f_i^2(p^2) - f_i^1(p^1))C_i > 0$$

until the upper bound $p_i = 1$ is reached which is the case for

$$f_i^2(p^2) = f_i^{21}(1 - p^2) + f_i^{22}p^2 > f_i^{11}(1 - p^1) + f_i^{12}p^1 = f_i^1(p^1).$$

Both coalitions can influence this condition by adapting their allocation p^I to actor A_i's needs, however at the cost of potentially losing other actors. Assuming the case that the second issue offers higher efficiency for each of the coalitions, $f_i^{I2} > f_i^{I1}$ (otherwise the options can be renumbered), then coalition $\mathcal{A}^2$ would gain support from those actors A_i for which

$$p^2 > \frac{p^1(f_i^{12} - f_i^{11}) + f_i^{11} - f_i^{21}}{f_i^{22} - f_i^{21}} \equiv \bar{p}_i^2(p^1).$$

Interesting cases are those for which the allocation threshold is within the boundaries $0 \leq \bar{p}_i^2 \leq 1$ (otherwise one of the coalitions succeeds with regard to actor A_i, whatever action is taken). This leads to the combined requirement

$$0 \leq \frac{f_i^{21} - f_i^{11}}{f_i^{12} - f_i^{11}} \leq p^1 \leq \frac{f_i^{22} - f_i^{11}}{f_i^{12} - f_i^{11}} \leq 1,$$

which is satisfied for the order of factors

$$f_i^{11} \leq f_i^{21} \leq f_i^{22} \leq f_i^{12}.$$

Under these conditions for each of the actors there is a dividing line (reaction curve) in the allocation space of both coalitions. By crossing these lines a coalitions gains or loses support from the respective actor because he/she would gain more value from the other coalition. The associated evolutionary game consists of $n \times 2 \times 2$ coupled dynamic equations which represent this competition among coalitions for actors to support them.

6 Examples of Adaptive Coalitions

6.1 Voting on Coalitions

Voting can be interpreted as an adaptation process between actors (voters) and the coalitions (parties) they vote for. Each voter A_i $(i = 1, \ldots, n)$ has a number of votes $0 \leq C_i \leq C_i+$ which can be allocated to the different parties $\mathcal{A}^I$. Their voting power $C^I = \sum_j p_j^I C_j$ is used to influence issues of interest for the voters:

$$a_i^k = \sum_{I=1}^{N} \frac{p_I^k \; C^I}{c_I^k} \qquad (k = 1, \ldots, m),$$

where c_I^k is the cost of translating voting power into real action, such as efforts of communication and negotiation to acquire a sufficient number of votes to exceed a threshold (quota, critical mass) C^{k*} given for this issue. The art of negotiation is to convince enough decision-makers to pass the threshold. Examples are political decisions by governments on allocation of the State budget or taxes. Coalition actions in return generate value for each of the actors $V_i = \sum_k^{m_I} v_i^k a_I^k$. The typical voting situation is characterized by special conditions:

- One man, one vote: $C_i = 1$ for all $i = 1, \ldots, n$.
- A_i can only vote for one party $\mathcal{A}^{I*}$ at a time (no mixed vote): $p_i^I = 1$ for $I = I^*$ and $p_i^I = 0$ for $I \neq I^*$.
- Issue a_I^k can be pursued only if the combined voting power for this issue exceeds the quota, i.e., for $\sum_{I=1}^{N} p_I^k \cdot C^I > C^{k*}$. In some voting situations the fraction of votes supporting an issue depends on the issue $p_I^k(a_I^k)$, e.g., either increases or decreases with a_I^k.

6.2 Production and Consumption in a Market Economy

The framework of adaptive coalitions is appropriate for analyzing the interplay between consumers (individual actors) and producers (coalitions) of economic goods in a market economy. Producers allocate their investment to the production of economic goods, sold on a market to consumers who in return allocate their income to the producers, raising their profit and thus the possibility to invest in the next time step. Both are characterized as follows:

- *Producers* $\mathcal{A}^I$ $(I = 1, \ldots, N)$ use their investment C^I to produce $k = 1, \ldots, m_I$ economic goods $x_I^k = p_I^k \cdot C^I / c_I^k$, where p_I^k is the fraction allocated to each good and c_I^k is the unit cost. For a price q_I^k for each good, the profit

  $$V^I = \sum_{k=1}^{m_I} q_I^k x_I^k = \sum_k \frac{q_I^k p_I^k}{c_I^k} C^I$$

  serves as the coalition value of $\mathcal{A}^I$.

- *Consumers* A_i $(i = 1, \dots, n)$ allocate their income C_i to buying $k_I = 1, \dots, m_I$ goods $x_i^{k_I} = C_i \cdot p_i^{k_I} / q_i^{k_I}$ from the different producers $\mathcal{A}^I$, given by allocation $p_i^{k_I}$ and prices $q_i^{k_I}$. The value from all producers is

$$V_i = \sum_{I=1}^{N} \sum_{k_I=1}^{m_I} v_i^{k_I} x_i^{k_I} = \sum_{I}^{N} \sum_{k_I}^{m_I} \frac{v_i^{k_I} p_i^{k_I}}{q_i^{k_I}} C_i,$$

where $v_i^{k_I}$ is the value of A_i per unit of good $x_i^{k_I}$ produced by $\mathcal{A}^I$.

Adapting allocation allows producers and consumers to maximize their value. This process, which may be built on the outlined evolutionary game, is further shaped by the price dynamics which continues until the demand-supply relationship achieves a balance, given by the market equilibrium for each product:

$$x^k = \sum_i x_i^k = \sum_i \frac{p_i^k C_i}{q^k} = \sum_I \frac{p_I^k C^I}{c_I^k} = \sum_I x_I^k,$$

which corresponds to the equilibrium market price for each good,

$$q^k = \frac{\sum_i p_i^k C_i}{\sum_I p_I^k C^I / c_I^k} \qquad (k = 1, \dots, m).$$

With this price the value functions of producers and consumers become

$$\begin{aligned} V_i &= \sum_I \sum_k \kappa_i^k \frac{v_i^k}{c_I^k} p_I^k C^I \\ V^I &= \sum_i \sum_k \kappa_I^k p_i^k C_i, \end{aligned}$$

where $\kappa_i^k = p_i^k C_i / (\sum_j p_j^k C_j)$ is the market share of consumer A_i for product x^k and $\kappa_I^k = (p_I^k C^I / c_I^k) / [(\sum_J p_J^k C^J) / c_J^k]$ is the market share of producer $\mathcal{A}^I$ for product x^k, both depending on the allocation of all actors to the different goods. Both value functions show the interdependence of consumers (actors) and producers (coalitions): If producers invest more, this generally increases the consumer values and vice versa.

6.3 Energy Management and Carbon Emissions

Dynamic games have found a wide range of applications in environmental policy and economics [3,4,7]. Using dynamic games can provide interesting methodological tools to understand and find cooperative solutions in energy and climate policy, including technology transfer and emissions trading [9,21]. Climate games deal with decision-making and interaction among multiple actors on global, regional and local levels of climate policy [6,12]. Including the concept of adaptive coalitions extends previous efforts to apply dynamic games in this field [11,18,19], as well as in fishery management [15,20].

In the following, we analyze energy production by $I = 1, \ldots, N$ energy providers $\mathcal{A}^I$ (coalitions) which receive investment from $i = 1, \ldots, n$ customers A_i (such as consumers, companies, governments) to produce value (such as consumption, wealth, profit) and we can select between $k_I = 1, \ldots, m_I$ primary energy sources. The possible actions by the providers $\mathcal{A}^I$ are the amounts of energy produced of each type

$$a_I^k = E_I^k = \frac{p_I^k}{c_I^k} C^I$$

for a given budget C^I, cost per energy unit c_I^k and allocation p_I^k. We assume that each customer A_i receives a share $p_I^i = p_i^I C_i / C^I$ of energy, proportionate to the investment given to this provider, and decides how much energy of which type is consumed. Then the energy of type k received by customer A_i from all providers $\mathcal{A}^I$ is $E_i^k = \sum_I p_I^i E_I^k = \sum_I p_i^I p_I^k C_i / c_I^k$. This energy is used for the production or consumption of economic output Q_i and at the same time generates carbon emissions G_i according to

$$\begin{aligned} Q_i &= \sum_k q_i^k E_i^k \\ G_i &= \sum_k g_i^k E_i^k, \end{aligned}$$

where q_i^k is the productivity for a primary energy unit of type k (which can differ because of different energy efficiencies) and g_i^k is the amount of carbon emission per energy unit. Then the value of customer A_i is defined as

$$V_i = u_i Q_i - d_i \sum_j G_j = \sum_k \left[(u_i q_i^k - d_i g_i^k) E_i^k - \sum_{j \neq i} d_i g_j^k E_j^k \right],$$

where u_i translates the economic output Q_i into value (e.g., representing the utility of consumers from a unit of production, the wealth created per unit of GDP or the price of goods sold by firms) and d_i represents the damage/cost per emission unit (e.g., damages related to climate change, energy taxes or costs for emission trading). The customer values can be expressed as $V_i = \sum_j f_{ij} C_j$ where $f_{ij} = \sum_I \sum_k p_j^I p_I^k (v_{ij}^k / c_I^k)$ with $v_{ii}^k = u_i q_i^k - d_i g_i^k$ and $v_{ij}^k = -d_i g_j^k$. Then the dynamic interaction between energy providers and customers can be described by the evolutionary game outlined above.

Example 6.1. We apply the general methodology to a specific example with two energy paths (called the “old” and the “new” path, marked in Figure 2 as A

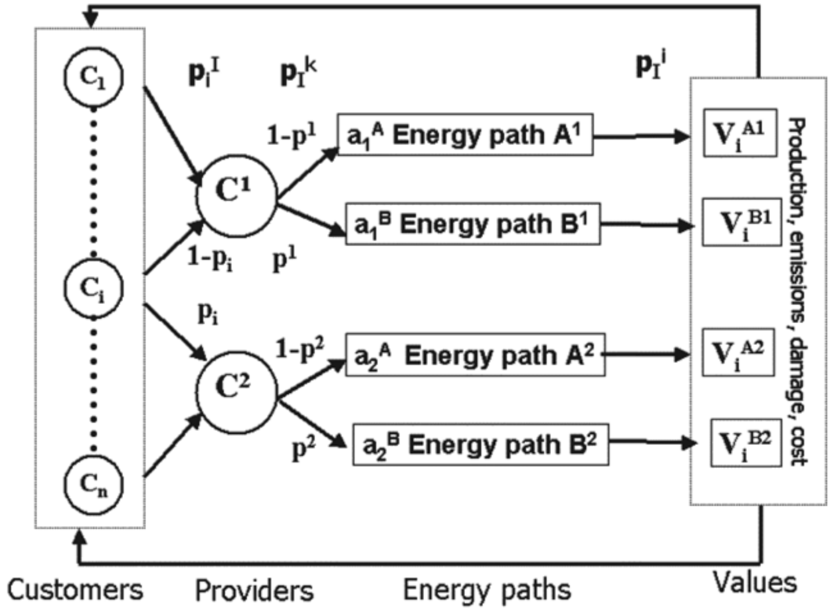


Figure 2: The interplay between energy providers and customers for different energy paths (marked here by A and B).

and B), two energy providers and six customers. The second new energy path is generally more expensive than the old one, where in addition the first provider is cheaper in producing energy 1 and the second provider is cheaper in energy 2. The customers differ with regard to the energy efficiency of production, where the efficiency for energy 1 declines from customer 1 to 6 and the efficiency for energy 2 increases from customer 1 to 6. The range both for unit cost and efficiency is a factor of two. Carbon emissions for the new energy are half of those for the old energy, which is $g_i^1 = 1$ for all $i = 1, \ldots, 6$. Damages per emission unit gradually increase from $d_1 = 0.2$ for the first customer to $d_6 = 0.4$ for the sixth customer. The benefit per unit of production is $u_i = 1$ for all actors. Finally, the customers are assumed to build capital; that is, they use a fraction of their produced net value $V_i(t) - C_i(t)$ in period t to increase the stock of maximum investment $C_i^+(t+1)$ in the following period. The parameters $q_i^k, g_i^k, c_i^k, u_i, d_i$ can be functions of time, energy and investment, but will be treated here as constants in a given moment of decision.

Applying the replication equations of the evolutionary game in p_i^I and p_I^k, we study an example for the evolution of actors and coalitions over time (see Figure 3). The model runs, starting with the same initial condition for actors and coalitions, show a growth in the resources of actors and coalitions, where

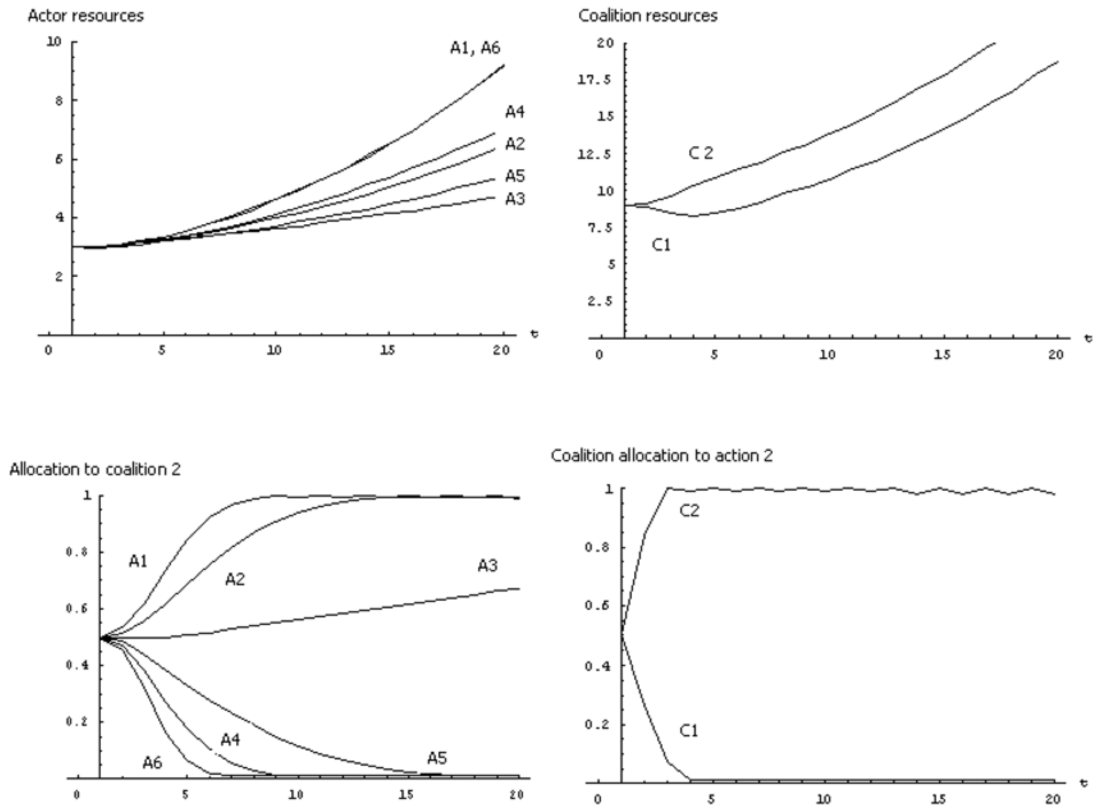


Figure 3: Simulation of the evolutionary game of adaptive coalition formation with two energy paths for two energy providers (coalitions), represented by C1 and C2, and six customers (actors) A1 to A6 using energy to produce value.

actors A_1 and A_6 and coalition $\mathcal{A}^2$ have the strongest growth rates. Coalition $\mathcal{A}^1$ specializes in the old energy path and coalition $\mathcal{A}^2$ in the new energy path, where both have cost advantages. While A_1 and A_6 show the strongest specialization with regard to the coalitions by completely investing in either $\mathcal{A}^2$ or $\mathcal{A}^1$ after an initial adjustment, customer A_3 is largely indifferent between the coalitions, with a slight tendency towards coalition $\mathcal{A}^2$. However, more customers (three together) still prefer coalition $\mathcal{A}^1$ representing the established cheaper energy path with higher emissions.

7 Summary and Outlook

The concept of adaptive coalitions developed in this paper provides a framework for studying the interaction between individual and collective players and the evolution of coalitions in various fields. First applications have been outlined, in particular in energy management and emission reduction. While the concept offers further development potential and requires more mathematical elaboration in the context of the established theory of dynamic and cooperative games, its usefulness is to be demonstrated in future applications in social, economic and environmental sciences.

REFERENCES

[1] Aubin J.P., Cooperative fuzzy games, *Mathematics of Operations Research*, 6, 1, 1–13, 1981.

[2] Aubin J.P., *Optima and Equilibria*, Springer, Berlin/Heidelberg/New York, 1998.

[3] Carraro C. and Filar J.A. (eds.), *Control and Game-Theoretic Models of the Environment*, Birkhäuser, Boston/Basel/Berlin, 1995.

[4] Finus M., *Game Theory and International Environmental Cooperation*, Edward Elgar, Cheltenham/Northhampton, MA, 2001.

[5] Göbeler F. and Scheffran J., Extended power values and dynamic coalition formation, in L. Petrosjan, N.A. Zenkevich (eds.), *The Tenth International Symposium on Dynamic Games and Applications*, July 8–11, 2002, Proceedings Vol. I, International Society of Dynamic Games, St. Petersburg, 338–340, 2002.

[6] Grundig F., Ward H. and Zorick E.R., Modeling global climate negotiations, in U. Luterbacher, D.F. Sprinz (eds.), *International Relations and Global Climate Change*, MIT Press, Cambridge, MA, 2001.

[7] Hanley N. and Folmer H. (eds.), *Game Theory and the Environment*, Edward Elgar, Cheltenham/Northhampton, MA, 1998.

[8] Haurie A., A historical perspective on cooperative differential games, in E. Altmann, O. Pourtaillier (eds.), *Advances in Dynamic Games and Applications*, Annals of the International Society of Dynamic Games, Birkhäuser, Boston/Basel/Berlin, 19–29, 2001.

[9] Haurie A. and Viguier L., A stochastic dynamic game of carbon emissions trading, *Environmental Modeling and Assessment*, 8, 239–248, 2003.

[10] Hofbauer J. and Sigmund K., *Evolutionary Games and Population Dynamics*, Cambridge University Press, London/New York, 1998.

[11] Ipsen D., Rösch R. and Scheffran J., Cooperation in global climate policy: Potentialities and limitations, *Energy Policy*, 29, 4, 315–326, 2001.

[12] Kemfert C., Climate coalitions and international trade: assessment of cooperation incentives by issue linkage, *Energy Policy*, 32, 4, 455–465, 2004.

[13] Krabs W. and Pickl S.W., *Analysis, Controllability and Optimization of Time-Discrete Systems and Dynamical Games*, Springer, Berlin/Heidelberg/ New York, 2003.

[14] Krabs W., Pickl S.W. and Scheffran J., An n-person game under linear side conditions, in E.J. Dockner, R.F. Harte, M. Luptucik, and G. Sorger (eds.), *Optimization, Dynamics and Economic Analysis*, Essays in Honor of Gustav Feichtinger, Springer/Physica, 76–85, 2000.

[15] Kropp J., Eisenack K. and Scheffran J., Sustainable marine resource management, in C. Pahl-Wostl, S. Schmidt, A.E. Rizzoli, and A.J. Jakeman (eds.), *Complexity and Integrated Resources Management*, Transactions 2nd Biennial Meeting, Internat. Environmental Modelling and Software Society, Vol. I, 104–109, Manno, 2004.

[16] Mares M., Fuzzy coalition structures, fuzzy sets and systems, *Elsevier Sciences*, 114, 23–33, 2000.

[17] Scheffran J., The dynamic interaction between economy and ecology—cooperation, stability and sustainability for a dynamic-game model of resource conflicts, *Mathematics and Computers in Simulation*, 53, 371–380, 2000.

[18] Scheffran J., Stability and control of value-cost dynamic games, *Central European Journal of Operations Research*, 9, 7, 197–225, 2001.

[19] Scheffran J., Interaction in climate games: The case of emissions trading, in J. Geldermann, M. Treitz (eds.), *Entscheidungstheorie und -praxis in industrieller Produktion und Umweltforschung*, Shaker, Aachen, 1–18, 2004.

[20] Scheffran J. and Pickl S., Control and game-theoretic assessment of climate change—options for joint implementation, *Annals of Operations Research*, 97, 203–212, 2000.

[21] Svirezhev Y.M., von Bloh W. and Schellnhuber H.J., "Emission game": some applications of the theory of games to the problem of CO_2 emission, *Environmental Modelling and Assessment*, 4, 235–242, 1999.

On Assignment Games

T. E. S. Raghavan*
Department of Mathematics, Statistics and Computer Science
University of Illinois at Chicago
Chicago, IL, U.S.A.
ter@uic.edu

Peter Sudhölter
Department of Economics
University of Southern Denmark at Odense
Odense, Denmark
psu@sam.sdu.dk

Abstract

This chapter surveys recent developments on some basic solution concepts, like stable sets, the core, the nucleolus and the modiclus for a very special class of cooperative games, namely assignment games with transferable utility. The existence of a stable set for assignment games is still an open problem.

1 Introduction

In this survey we concentrate on a subclass of transferable utility (TU) games called *assignment games* and on some properties of their solutions. Assignment games with side payments were introduced by Shapley and Shubik [23]. These games are models of two-sided markets. Players on one side, called sellers, supply exactly one unit of some indivisible good, say, a house in exchange for money, with players from the other side, called buyers. Each buyer has a demand for exactly one house. When a transaction between seller i and buyer j takes place, a certain profit $a_{ij} \geq 0$ accrues. The worth of a coalition is given by an assignment of the players within the coalition which maximizes the total profit of the assigned pairs. Therefore the characteristic function is fully determined by the profits of the mixed pairs.

Assignment games have a nonempty core and are simply dual optimal solutions to the associated optimal assignment problem. It is known that prices which competitively balance supply and demand correspond to elements in the core. The nucleolus, lying in the lexicographic center of the nonempty core, has

*Partially funded by N.S.F. Grant DMS 0072678.

the additional property that it satisfies each coalition as much as possible. The corresponding prices favor neither the sellers nor the buyers, and hence provide some stability for the market.

To find the nucleolus for any general cooperative game, Kohlberg [11] proposed a weighted sum minimization approach leading to a single, but extremely large linear program (LP). In order to ensure that the highest excess gets the largest weight, the second highest excess gets the second largest weight, and so on, coefficients from a very wide range must appear in the constraints (causing serious numerical accuracy problems even for 3-person games). Since all possible permutations of the coalitions must be present among the constraints, the approach enlarges the size of the LP enormously. In Owen's [16] improved version, although one has to solve a somewhat simpler minimization problem, even then the constraints grow exponential in terms of the number of players. Indeed for an n-person game he reduces the problem to a linear programming problem in $2^{n+1}+n$ variables with 4^n+1 constraints, where n is the number of players.

When solving an optimal assignment problem, the ordinary primal simplex method encounters high levels of degeneracy. It is clearly outperformed by specifically designed algorithms, such as Kuhn's [12] well-known Hungarian algorithm. Also for assignment games a method based on general linear programming is not well suited, since the combinatorial structure of the characteristic function cannot be effectively translated into a continuous problem formulation. In the spirit of combinatorial techniques for assignment problems we apply graph-theoretic techniques to replace linear programming for locating the nucleolus for assignment games.

We will survey properties that are unique for the core of assignment games. Besides the nucleolus, as a point solution concept there is yet another point solution concept for all TU cooperative games, called the modiclus. However, it should be remarked that for assignment games another classical solution concept, namely the bargaining set $\mathcal{M}_1^{(i)}$, introduced by Davis and Maschler [5] (see also [1]) simply coincides with the core for assignment games [25] (also see [7]). While the nucleolus is defined by ranking imputations lexicographically via excesses, the modiclus is defined by lexicographical ranking bi-excesses (for definitions see the next section).

2 Preliminaries

A (cooperative TU) *game* is a pair (N,v) such that $\emptyset \neq N$ is finite and $v : 2^N \to \mathbb{R}$, $v(\emptyset) = 0$. A *coalition* is a nonempty subset of N and v is the *coalition function* of (N,v). A *Pareto optimal payoff vector* (*pre-imputation*) is a vector $x \in \mathbb{R}^N$ such that $x(N) = v(N)$, where $x(S) = \sum_{i\in S} x_i$ ($x(\emptyset) = 0$) for every $S \subseteq N$ and every $x \in \mathbb{R}^N$ with $\sum_{i\in\emptyset} x_i = 0$ by convention. A pre-imputation x is an *imputation* if it is *individually rational*, that is, if $x_i \geq v(\{i\})$ for all $i \in N$. A game v is normalized if for any $S \subseteq T \subseteq N, v(S) \leq v(T)$. Let $X(N,v)$ and

$\mathcal{I}(N,v)$ denote the set of pre-imputations and imputations, respectively. Thus, $X(N,v)$ is a nonempty polyhedral set and $\mathcal{I}(N,v)$ is a polytope. Moreover, $\mathcal{I}(N,v) \neq \emptyset$, if and only if $\sum_{i \in N} v(\{i\}) \leq v(N)$. The *core* of (N,v), $\mathcal{C}(N,v)$, is the set of all imputations x such that

$$x(S) \geq v(S) \text{ for all } S \subseteq N. \tag{1}$$

Note that the core is always a polytope, but it may be empty even for games that have imputations. We will assume that $\mathcal{I}(N,v)$ is nonempty.

2.1 The Nucleolus and the Modiclus

Let (N,v) be a game. If $H = (h^k)_{k \in D}$ is a finite family of real-valued functions on $X = X(N,v)$ (the family of *dissatisfaction* functions) and $x \in X$, then let $\theta^H(x) \in \mathbb{R}^d$ (where $d = |D|$ denotes the cardinality of D) be the vector whose components are the numbers $h^k(x)$, $k \in D$, arranged in a nonincreasing order, that is,

$$\theta_t^H(x) = \max_{T \subseteq D, |T| = t} \min_{k \in T} h^k(x) \text{ for all } t = 1, \ldots, d.$$

Let $\geq_{lex}$ denote the lexicographical ordering on $\mathbb{R}^d$; that is, $x \geq_{lex} y$, where $x, y \in \mathbb{R}^d$, if $x = y$ or if there exists $1 \leq t \leq d$ such that $x_k = y_k$ for all $1 \leq k < t$ and $x_t > y_t$. The *nucleolus* of H, $\mathcal{N}(H)$ is defined (see [9]) by

$$\mathcal{N}(H) = \{x \in X \mid \theta^H(x) \geq_{lex} \theta^H(y) \text{ for all } y \in X\}. \tag{2}$$

Let the class H be taken to be the dissatisfactions of coalitions at any $x \in \mathbb{R}^N$ measured by $e(S,x,v) = v(S) - x(S)$ called the *excess* of S at x. Now the *pre-nucleolus* of (N,v) is defined to be the set $\mathcal{N}\left((e(S,\cdot,v))_{S \subseteq N}\right)$. Indeed the pre-nucleolus [19], $\mathcal{N}\left((e(S,\cdot,v))_{S \subseteq N}\right)$, is a singleton, abbreviated by $\nu(N,v)$. It is also called the *nucleolus* if the domain of excesses is restricted to the imputation set.

In order to define the modiclus of (N,v) we proceed similarly. Instead of the ordered vector of excesses we take the nonincreasingly ordered vector of bi-excesses. (Here the *bi-excess* of a pair (S,T), $S,T \subseteq N$, at x is the number $e^b(S,T,x,v) = e(S,x,v) - e(T,x,v)$.) The *bi-excess* can be seen as the level of *envy* of S against T at x. The *modiclus* of (N,v) is the set $\mathcal{N}\left((e^b(S,T,\cdot,v))_{S,T \subseteq N}\right)$. The modiclus denoted by $\psi(N,v)$ is a singleton [28].

Recall that the *dual* game of (N,v), is defined by $v^*(S) = v(N) - v(N \setminus S)$ for all $S \subseteq N$. Also, recall that (N,v) is

- *constant-sum* if $v(S) + v(N \setminus S) v(N)$ for all $S \subseteq N$;
- *convex* if $v(S) + v(T) \leq v(S \cap T) + v(S \cup T)$ for all $S, T \subseteq N$;
- *zero-monotonic* if $v(S \cup \{i\}) \geq v(S) + v(\{i\})$ for all $i \in N$ and all $S \subseteq N \setminus \{i\}$.

The following relationships between the modiclus and the pre-nucleolus are of interest. (See [14] and [28,29].)

Proposition 2.1. *Let (N,v) be a game.*

(1) *Let $^*: N \to N^*$ be a bijection such that $N \cap N^* = \emptyset$. If $(N \cup N^*, \widetilde{v})$ is defined by*

$$\widetilde{v}(S \cup T^*) = \max\{v(S) + v^*(T), v^*(S) + v(T)\} \text{ for all } S, T \subseteq N,$$

then $\psi_i(N,v) = \nu_i(N \cup N^*, \widetilde{v}) = \nu_{i^*}(N \cup N^*, \widetilde{v})$ for all $i \in N$.

(2) *If (N,v) is a constant-sum game, then $\psi(N,v) = \nu(N,v)$.*

(3) *If (N,v) is a convex game, then $\psi(N,v) \in \mathcal{C}(N,v)$.*

(4) *If (N,v) is zero-monotonic, then $\psi(N,v)$ and $\nu(N,v)$ are individually rational.*

If $\nu(N,v) \in \mathcal{I}(N,v)$, then $\nu(N,v)$ is the *nucleolus* of (N,v).

2.2 Assignment Games

Let $N = P \cup Q$ where P, Q is a partition of the player set N into two types of players called sellers and buyers respectively. The players could also be colleges and students or even men and women dating. From now on we will stick to calling them sellers and buyers. Neither sellers nor buyers have any interest in mutual cooperation among themselves. Suppose each seller owns an indivisible object, say, a house which he values as worth at least c_i to him. Each buyer j has a ceiling price b_{ij} for the house of seller i. For any $i \in P, j \in Q$ the coalitional worth of the seller-buyer pair $\{i,j\}$ is taken to be $v(\{i,j\}) = a_{ij} = \max(b_{ij} - c_i, 0)$. Any arbitrary coalition $S \subseteq N$ decomposes into sellers S_1 and buyers S_2. Here if $|S_1| \neq |S_2|$ then by introducing either dummy sellers or dummy buyers if necessary we can assume $|S_1| = |S_2|$. We will take $a_{ij} = 0$ if i or j is a dummy player, namely a dummy seller or a dummy buyer respectively. Thus assuming $|S_1| = |S_2|$, let σ_S denote any arbitrary bijection $\sigma_S : S_1 \to S_2$. Given the coalition S and matrix A for player set $N = P \cup Q$ we define the assignment game with characteristic function given by

$$v_A(S) = \max_{\sigma_S} \sum_{i \in S_1} a_{i\sigma_S(i)}.$$

If $S \subseteq T, v_A(S) \leq v_A(T)$ and $v_A(\{i\}) = 0$ for all $i \in N$. Thus the pre-nucleolus is the same as its nucleolus. The following theorem is due to Shapley and Shubik [23].

Theorem 2.1. *The game (N, v_A) has a nonempty core. The worth of the grand coalition N of v_A is given by the following linear program:*

$$\begin{aligned}
&\max \textstyle\sum_{k \in P} \sum_{\ell \in Q} a_{k\ell} x_{k\ell} \\
&\textit{subject to} \\
&\textstyle\sum_{\ell \in Q} x_{\widetilde{k}\ell} \leq 1, \\
&\textstyle\sum_{k \in P} x_{k\widetilde{\ell}} \leq 1, \\
&x_{\widetilde{k}\widetilde{\ell}} \geq 0, \quad \forall\, \widetilde{k} \in P, \widetilde{\ell} \in Q.
\end{aligned} \tag{3}$$

The core of the game consists of dual optimal solutions to this linear programming problem. The core of the subgame (S, v) *(that is, defined by* $v(T) = v_A(T)$ *for all* $T \subseteq S$*) is the set of optimal solutions of the dual program. Hence,* (N, v_A) *is* totally balanced*, that is,* (N, v_A) *and each of its subgames* (S, v_A)*,* $\emptyset \neq S \subseteq N$*, have nonempty cores and thus* $\nu(N, v_A) \in \mathcal{C}(N, v_A)$*.*

The following observation was made by Sudhölter [30].

Proposition 2.2. *Given an assignment game* (N, v_A) *with sellers* P *and buyers* Q *the modiclus of* (N, v_A) *treats* P *and* Q *equally, that is,* $\psi(P) = \psi(Q)$ *where* $\psi = \psi(P \cup Q, v_A)$*.*

Example 2.1. [Glove Game] Let $P = \{1, \dots, p\}$, $Q = \{1, \dots, q\}$, $p \leq q$, let $A = (a_{k\ell})_{k \in P, \ell \in Q}$ be given by $a_{k\ell} = 1$, and let $v = v_A$. Then $v(S) = \min\{|S \cap P|, |S \cap Q|\}$ for all $S \subseteq N$. Moreover, let $\nu = \nu(N, v)$ and $\psi = \psi(N, v)$. If $p = q$, then $\nu_i = 1/2 = \psi_i$ for all $i \in N$. If $p < q$, then $\nu_k = 1$ for all $k \in P$ and $\nu_\ell = 0$ for all $\ell \in Q$. Proposition 2.2 and the well-known equal treatment property yield $\psi_k = \frac{1}{2}$ for $k \in P$ and $\psi_\ell = \frac{p}{2q}$ for $\ell \in Q$. Hence, $\psi \in \mathcal{C}(N, v)$ if and only if $p = q$.

Let $(P \cup Q, v_A)$ be an assignment game. As all our solution concepts satisfy the strong null-player property, we shall always assume that $|P| = |Q|$. Moreover, our solution concepts are anonymous. Hence, we shall always assume that

$$P = \{1, \dots, p\}, Q = \{1', \dots, p'\}, \text{ and } v_A(N) = \sum_{i=1}^{p} a_{ii'}, \tag{4}$$

that is, an optimal assignment for N is attained at the *diagonal* $\{\{i, i'\} \mid i = 1, \dots, p\}$.

3 Core Stability and Related Concepts

It was von Neumann and Morgenstern [32] who first introduced the notion of a *stable set*. Stable sets are characterized by the notions of internal stability and external stability. The two definitions hinge on comparing pairs of imputations for a game (N, v). We say imputation x dominates imputation y via coalition $S (x \succ_S y)$ if $x_i > y_i, i \in S$, and $\sum_{i \in S} x_i \leq v(S)$. Intuitively players in coalition S object to their share according to y when they have a better share of the grand coalitional worth according to x which is not a dream, but is within their reach. A set $V \subseteq \mathcal{I}(N, v)$ is called *internally stable* if no imputations in V can dominate another imputation in V. Further the set V is *externally stable* if any imputation not in V is dominated by some imputation in V via a coalition. A set V is called *stable* for a game (N, v) if V is both internally and externally stable.

Since the core when it exists is a polyhedral set, the problem of existence is simply reduced to the existence of solution for a system of linear inequalities. Using the duality theorem, the so-called Bondareva [4] and Shapley [20] theorem sharpens the problem to the existence of balanced collections. Thus the existence is decidable via a simplex algorithm in a constructive fashion. Unfortunately, there is no such constructive approach to the existence of a stable set for an arbitrary game (N, v). In general games that are physically motivated have been found to have a plethora of stable sets. It was Lucas [13] who surprised game theorists by constructing a ten-person game with no stable set. In this connection we have the following.

Open Problem: Do all assignment games admit nonempty stable sets?

There are special classes of games for which the stable set exists and is unique. Perhaps the best-known such class is the class of convex games. In fact Shapley [21] proved that for convex games the core is the unique stable set. Since assignment games have a nonempty core, a natural question is to identify those assignment games whose core is also stable, and hence is the unique stable set. For assignment games we have two special imputations called the seller's corner and buyer's corner. In the seller's corner the seller takes away the full coalitional worth and the optimally matched mate receives nothing. In the buyer's corner, it is the buyer who takes away the coalitional worth, with the optimally matched mate receiving nothing. Since domination of an imputation by another imputation is possible only with buyer-seller coalitional pairs, the above two extreme imputations cannot be dominated by any imputation. Thus for the core to be a stable set, necessarily these two imputations must lie in the core. Interestingly, that condition is also sufficient for core stability [27].

Several other sufficient conditions for the stability of the core have been discussed in the literature.

Given an n-person game (N, v) with a nonempty core, the game admits a *Large core* if and only if for any n-vector x with $x(S) \geq v(S)$, $\forall\ S \subseteq N$, there exists a core element y such that $y \leq x$ coordinatewise. In an unpublished paper Kikuta and Shapley [10] investigated another condition, baptized to extendability of the game in the work of van Gellekom et al. [31]. For a totally balanced game (N, v) the *core is extendable* if and only if any core element x of any subgame (S, v), $S \subseteq N$ is simply the restriction of some core element $y \in \mathcal{C}(N, v)$ to the coordinates in S. The core of a game (N, v) is *exact* if and only if for any coalition S, there is some core element x such that $x(S) = v(S)$. Sharkey [23] and Biswas et al. [3] proved the following.

Theorem 3.1. *For any totally balanced game (N, v) we have the following: Core is Large $\Rightarrow$ Core is extendable $\Rightarrow$ Core is exact.*

A game (N, v) is called symmetric if for any two coalitions S, T with $|S| = |T|$, $v(S) = v(T)$. In fact Biswas et al. [3] proved the following.

Theorem 3.2. *For any totally balanced symmetric game or for games with $|N| < 5$ Core is exact $\Rightarrow$ Core is extendable $\Rightarrow$ core is Large.*

Unfortunately, given the data of the game, (N, v) we have no easy way to verify any of these conditions.

It turns out that for the class of assignment games, Largeness of the core, extendability and exactness of the game are all equivalent conditions, but are strictly stronger than the stability of the core. However for assignment games (v_A, N) many of these implications are equivalent and are easily verifiable via the matrix A defining the assignment game.

Let A be a nonnegative real matrix such that (4) is satisfied. We say that A has a *dominant diagonal* if $a_{ii'} \geq a_{ij'}$ and $a_{ii'} \geq a_{ji'}$ for all $i, j \in P$. Also, we say that A has a *doubly dominant diagonal* if $a_{ii'} + a_{jk'} \geq a_{ik'} + a_{ji'}$ for all $i, j, k \in P$. Now we are able to state the following characterization [27].

Theorem 3.3. *Let $P = \{1, \ldots, p\}$, $Q = \{1', \ldots, p'\}$, let A be a nonnegative real matrix on $P \times Q$ satisfying (4), let $N = P \cup Q$, and let v_A be the coalition function of the corresponding assignment game.*

(N, v_A) has a stable core $\Leftrightarrow$ A has a dominant diagonal.
(N, v_A) has a Large core $\Leftrightarrow$ (N, v_A) has an extendable core $\Leftrightarrow$ (N, v_A) is exact $\Leftrightarrow$ A has a dominant and doubly dominant diagonal.
(N, v_A) is convex $\Leftrightarrow$ A is a diagonal matrix (that is, $a_{ij'} \neq 0$ implies $j = i$).

Despite Example 2.1, from the above theorem we may deduce the following result for the modiclus [18].

Theorem 3.4. *The modiclus of an assignment game is in the core, provided the core is stable.*

The authors present a 15-person game which is exact and has a Large core and hence has a stable core and yet its modiclus is not a member of the core.

4 The Geometric Shape of the Core for Assignment Games

While Shapley and Shubik characterized the core of assignment games as dual optimal solutions of (3), they made another key observation that given any two core elements $(u^1, v^1), (u^2, v^2)$, the elements $(u^1 \vee u^2, v^1 \wedge v^2)$, and $(u^1 \wedge u^2, v^1 \vee v^2)$ are also core elements where $\vee, \wedge$ are the usual lattice operations, namely for vectors $u^1, u^2, (u^1 \vee u^2) = \max(u^1, u^2)$ where max is taken coordinatewise.

Interestingly, the dual inequalities that are used for determining the core as the optimal dual allocations have a special geometric structure. The core is obtained by starting with a cube $b_i \leq u_i \leq e_i$, $i = 1, \ldots, p$ for some constants b_i, e_i $i = 1, \ldots, p$ and then chopping off the 45-45-90 degree triangular cylinders determined by inequalities of the type

$$u_i - u_k \geq d_{ik} \quad \forall\ i, k \in 1, \ldots, p;\ i \neq k$$

for some constants $\{d_{ik}\}$. In fact the converse is also true, namely Quint [17] proved the following.

Theorem 4.1. *Let P be a polytope with elements $(u_1, \dots, u_p) \in \mathbf{R}^p$ satisfying*

$$u_i - u_k \geq d_{ik} \quad \forall\ i, k \in 1, \dots, p; i \neq k$$

$$b_i \leq u_i \leq e_i$$

for some constants $\{d_{ik}\}, b_i \geq 0, e_i \geq 0 \quad i = 1, \dots, p$. Then we can always find an assignment game with p sellers whose u space core coincides with P.

The extreme points of the core of assignment games can also be nicely recognized by the following graph-theoretic technique of Balinsky and Gale [2]. Given any core element (u, v) we can associate with the core element a graph Γ_{uv} with vertices as $P \cup Q$ and with edges (p, q) where $u_p + v_q = a_{pq}$.

Theorem 4.2. *A core element (u, v) of the assignment game v_A is an extreme point if and only if the graph Γ_{uv} is connected.*

The extreme points of the cores of subgames of assignment games have the following extension property [2].

Theorem 4.3. *If $(\tilde{u}, \tilde{v})$ is an extreme point of the core of some subgame on $\tilde{P} \cup \tilde{Q}$ of an assignment game with sellers P and buyers Q and defining matrix A, then there is an extreme point (u, v) of the polyhedron*

$$X = \{(u, v) : u_p + v_q \geq a_{pq}, u_{p0} = 0,\ p \in P, q \in Q\},$$

where 0 denotes the dummy buyer such that (u, v) agrees with $(\tilde{u}, \tilde{v})$ on $\tilde{P} \cup \tilde{Q}$.

The cores of assignment games and convex games share the following common properties [8,15].

Property 4.1. In each extreme point of the core allocations of an assignment game (N, v) there is at least one player i who receives his marginal contribution $v(N) - v(N \setminus \{i\})$.

Property 4.2. Every marginal contribution for any player is attained at some core element.

5 An Algorithm to Compute the Nucleolus

Given a game (N, v) and an imputation x let $f(S, x) = -e(S, x, v)$ (see Section 2.1). Hence $f(S, x)$ is the *satisfaction of the coalition S* at imputation x. As we focus on assignment games, we shall henceforth always assume that (N, v)

is zero-monotonic. Hence the nucleolus of the game is just its pre-nucleolus. We now slightly modify our viewpoint. With $H = (-f(S,\cdot))_{S \subseteq N}$ the nucleolus of a zero-monotonic game (N, v) is the unique member of the set given by the right-hand side of (2) in which we may replace X by $\mathcal{I}(N, v)$ by Proposition 2.1. By the *lexicographic center* of a nonempty closed convex subset D of the imputation set, we mean the unique point $x^* \in D$ which lexicographically minimizes the vector $\theta^H(x)$ over D (that is, the set defined by the right-hand side of (2) with $X = D$ is a singleton as shown by Schmeidler [19]). Even though the determination of the nucleolus is quite difficult in general, it is possible to locate it efficiently for special subclasses of games. We will describe an algorithm [26] to locate the nucleolus for an assignment game. We will reinterpret the game slightly differently as follows.

Stable Real Estate Commissions: House owners $P = \{U_1, U_2, \ldots, U_p\}$, each possessing one house, and house buyers $Q = \{V_1, V_2, \ldots, V_q\}$, each wanting to buy one house, approach a common real estate agent. Not revealing the identity of the buyers and sellers, the agent wants an up-front commission $a_{ij} \geq 0$ if he links seller U_i to buyer V_j. The sellers and buyers prefer fixed commissions $u_1, u_2, \ldots, u_p$ and $v_1, \ldots, v_q$. The agent has no objection if they meet his expectation for every possible link. He guarantees their money's worth in his effort and promises to take no commission from a seller (buyer) if he cannot find a suitable buyer (seller).

We define $P_0 = P \cup \{0\}$, $Q_0 = Q \cup \{0\}$, $a_{i0} = 0 \ \forall\ i \in P_0$, $a_{0j} = 0 \ \forall\ j \in Q_0$, $u_0 := 0$, $v_0 = 0$, and all the constraints in (1) reduce to

$$f_{ij}(u, v) = u_i + v_j - a_{ij} \geq 0 \quad \forall\ (i, j) \in P_0 \times Q_0). \tag{5}$$

If $\sigma \subseteq P \cup Q$ is an optimal assignment and D is the core we get

$$\{i, j\} \in \sigma \Rightarrow f_{ij}(u, v) = 0 \quad \forall\ (u, v) \in D. \tag{6}$$

With the convention that $(0, 0) \in \sigma$ we write $(i, 0) \in \sigma$ ($(0, j) \in \sigma$) if in σ row $i \in P$ (column $j \in N$) is not assigned to any column $j \in N$ (row $i \in Q$). Here σ is extended to a subset of $P_0 \times Q_0$ so that (6) also expresses the fact that D lies in the hyperplane $u_i = 0$ (or $v_j = 0$) for any unassigned row i (column j). It is easily seen that

$$D = \{(u, v) : f_{ij}(u, v) = 0 \ \forall\ (i, j) \in \sigma, f_{ij}(u, v) \geq 0 \ \forall\ (i, j) \notin \sigma\}. \tag{7}$$

Here and from now on $(i, j) \notin \sigma$ is written instead of $(i, j) \in (P_0, Q_0) \setminus \sigma$.

Among many vectors of commissions $(u_0, u_1, \ldots, u_m; v_0, v_1, \ldots, v_n)$ in D for the agent, he wants to choose one that is "neutral" and "stable." The lexicographic center is a possible option that is neutral and stable for all pairs.

For every $(u, v) \in D$ the first $\max(p, q) + 1$ components (those coordinates $k = (i, j)$ corresponding to $(i, j) \in \sigma$) of $\theta^H(u, v)$ are equal to 0. Let

$$\alpha^1 = \max_{(u,v) \in D} \min_{(i,j) \notin \sigma} f_{ij}(u, v).$$

Let

$$D^1 = \left\{ (u,v) \in D : \min_{(i,j) \notin \sigma} f_{ij}(u,v) = \alpha^1 \right\}$$

Let

$$\sigma^1 = \{(i,j) : f_{ij}(u,v) = \text{constant on } D^1\}.$$

The set σ^1 can be regarded as an "assignment" between the equivalence classes of the relation $\sim^1$ defined on M_0 and N_0 by

$$\begin{aligned} i_1 \sim^1 i_2 &\text{ if and only if } u_{i_1} - u_{i_2} \text{ is constant on } D^1, \\ j_1 \sim^1 j_2 &\text{ if and only if } v_{j_1} - v_{j_2} \text{ is constant on } D^1, \end{aligned}$$

respectively.

$$\begin{aligned} \alpha^2 &= \max_{(u,v) \in D^1} \min_{(i,j) \notin \sigma^1} f_{ij}(u,v), \\ D^2 &= \left\{ (u,v) \in D^1 : \min_{(i,j) \notin \sigma^1} f_{ij}(u,v) = \alpha^2 \right\}, \\ \sigma^2 &= \{(i,j) \in (M_0, N_0) : f_{ij}(u,v) \text{ is constant on } D^2\}. \end{aligned}$$

Let $i \sim^2 k$ if and only if $u_i - u_k$ is a constant on D^2. Observe that $\sigma^2 \supseteq \sigma^1 \supseteq \sigma$.

Therefore, after some $t \leq \min(m,n)$ rounds the process terminates with

$$\sigma^t = \{(i,j) = (M_0 \times N_0) : f_{ij}(u,v) \text{ is constant on } D^t\}.$$

This means that a subset of D is found that is parallel to all hyperplanes defining D. Since they include $u_i = 0$ for all $i \in M$ and $v_j = 0$ for all $j \in N$, this subset must consist of a single point. It can be proved [26] that this point is precisely the lexicographic center of D.

Next we illustrate by an example how to implement the procedure leading to the lexicographic center.

Example 5.1. We are given

$$A = \begin{bmatrix} 6 & 7 & 7 \\ 0 & 5 & 6 \\ 2 & 5 & 8 \end{bmatrix},$$

where $P = \{1,2,3\} = Q$. The unique optimal assignment for A is $\sigma = \{(1,1), (2,2), (3,3)\}$, i.e., the entries in the main diagonal. Starting with all commissions collected entirely from sellers, one could use the procedure to be described below to locate the u worst point $(u^1, v^1) = (0,6,4,6 : 0,0,1,2)$ in D. Further with rows numbered 0, 1, 2, 3 and columns numbered 0, 1, 2, 3 we can read off (u^1, v^1) from column 0, and row 0 of the matrix

$$[f_{ij}(u^1, v^1)] = \begin{bmatrix} 0 & 0^* & 1 & 2 \\ 6 & 0 & 0^* & 1 \\ 4 & 4 & 0 & 0^* \\ 6 & 4 & 2 & 0 \end{bmatrix}.$$

Even though the coordinates for the starred entries above are the next set with higher f_{ij} values in the lexicographic ranking, they are still 0. However, from now on there will be strict improvement with higher values when we follow the iteration. We want to move in a direction (s, t) inside D with one end at the extreme solution (u^1, v^1). Let the new point be $(u^2, v^2) = (u^1, v^1) + \beta \cdot (s, t)$ for some $\beta \geq 0$. Since the point (u^1, v^1) is the worst for all sellers in terms of commissions in D, they would like their commissions reduced.

Since (u^1, v^1) is the farthest from the hyperplanes indexed by $(0, 1), (1, 2)$ and $(2, 3)$ (indicated by a $*$ in the above matrix) this translates to the requirements

$$s_0 + t_1 \geq 1, \quad s_1 + t_2 \geq 1, \quad s_2 + t_3 \geq 1 \tag{8}$$

with at least one inequality. Since we must remain in D we also have

$$s_i + t_i = 0, \quad i = 1, 2, 3. \tag{9}$$

Combining (8) and (9) gives

$$t_1 - t_0 \geq 1, \quad t_2 - t_1 \geq 1, \quad t_3 - t_2 \geq 1 \tag{10}$$

with at least one equality to hold. Thus the direction for improvement for sellers is $(s, t) = (0, -1, -2, -3 : 0, 1, 2, 3)$. Next we determine how far we can move along this direction inside D starting from the initial u worst corner of D. That is

$$[f_{ij}((u^1, v^1) + \beta \cdot (s, t))] = \begin{bmatrix} 0 & (0+\beta)^* & 1+2\beta & 2+3\beta \\ 6-\beta & 0 & (0+\beta)^* & 1+2\beta \\ 4-2\beta & 4-\beta & 0 & (0+\beta)^* \\ 6-3\beta & 4-2\beta & (2-\beta)^\diamond & 0 \end{bmatrix},$$

where $*$ refers to the worst satisfied coalition at the current imputation, and $\diamond$ refers to the penultimate coalition. Compared to the worst satisfied mixed coalition consisting of the dummy seller 0 with buyer 1 with satisfaction $0 + \beta$, the next worst hit coalition is the one with seller 3 and buyer 2 and with satisfaction $f_{32} = 2 - \beta$ which is the first one to reach the same level as the worst hit one when improved. To reach this common level we equate $2 - \beta = 0 + \beta$ and we get $\beta = 1$ and $(u^2, v^2) = (0, 5, 2, 3; 0, 1, 3, 5)$. It can be shown that (u^2, v^2) is the u-worst corner (v-best corner) in D^1.

$$\begin{array}{c|c|c|c|cl}
0 & 1^* & 3 & 5 & & \\
\hline
5 & \boxed{0} & 1^* & 3 & -1 & \\
\hline
2 & 3 & \boxed{0} & 1^* & -2 & \beta = 1. \\
\hline
3 & 2 & 1^* & \boxed{0} & -2 & \\
\hline
 & +1 & +2 & +2 & &
\end{array}$$

The updated distance matrix is compactly represented by

Here the left-side border frame is the column vector u^2 and the top border frame is the row vector v^2 for all the non-dummy players. We also have the right-side frame with the column vector $s = (-1, -2, -2)^T$ and the bottom frame with the row vector $t(1, 2, 2)$.

They are derived from the following considerations: To improve further from $(u^2, v^2) = (0, 5, 2, 3; 0, 1, 3, 5)$ we need to find a direction to move inside D^1. Observe that the starred entries represent the worst hit coalitions at the current point. If the rows and columns are numbered 0,1,2,3 as before, the satisfactions of sellers with dummy buyers are given by the left-side frame. The satisfaction of buyers with dummy sellers are given by the entries of the top frame. Thus we have a starred value 1 at entries (0,1), (1,2), (2, 3) and (3, 2). This means that $f_{01}(u, v) \geq 1, f_{12}(u, v) \geq 1, f_{23}(u, v) \geq 1, f_{32}(u, v) \geq 1$ for all $(u, v) \in D^1$. Since on D^1 we have $f_{22}(u, v) = f_{33}(u, v) \equiv 0$, we have $f_{23}(u, v) = f_{32}(u, v) \equiv 1$ $\forall\ (u, v) \in D^1$. Thus the new direction (s, t) must satisfy $s_0 + t_1 \geq 1, s_1 + t_2 \geq 1, s_2 + t_3 = 0, s_3 + t_2 = 0$. Thus the direction is $(s, t)(0, -1, -2, -2 : 0, 1, 2, 2)$.

Thus we notice that on the new set $D^2 \subseteq D^1$ not only the coalitions $(1, 1), (2, 2), (3, 3)$ of buyer-seller pairs have constant value for the satisfaction at the imputations but also have constant satisfaction for the coalitions (2,3), (3,2). Thus what were originally boxed coalitions for D^1 are also boxed for D^2 and so are $(2, 3), (3, 2)$ coalitions. Now to determine the new step size β for the new direction we proceed as follows. Consider the matrix

$$\begin{array}{c|c|c|c|cl}
 & (1+\beta^*) & 3+2\beta & 5+2\beta & & \\
\hline
5-\beta & \boxed{0} & (1+\beta)^\diamond & 2+\beta & -1 & \\
\hline
(2-2\beta)^\diamond & 3-\beta & \boxed{0} & \boxed{1} & -2 & \beta = 1 \\
\hline
(3-2\beta) & (2-\beta) & \boxed{1} & \boxed{0} & -2 & \\
\hline
 & +1 & +2 & +2 & &
\end{array}$$

The decreasing distance $f_{20} = 2 - 2\beta$ is the first to reach the increasing second smallest distance $1 + \beta$. It happens when $\beta = 1/3$. So the maximal distance in this direction is $\beta = 1/3$, and the u-worst corner of the set D^2 of points with the second smallest distance 4/3 is $(u^3, v^3) = (0, 14/3, 4/3, 7/3; 0, 4/3, 11/3, 17/3)$. The updated distance matrix is

$$[f_{ij}(u^3, v^3)] = \begin{array}{c|c|c|c|} 0 & 4/3^* & 11/3 & 17/3 \\ \hline 14/3 & \boxed{0} & 4/3^* & 7/3 \\ \hline 4/3^* & 8/3 & \boxed{0} & \boxed{1} \\ \hline 7/3 & 5/3 & \boxed{1} & \boxed{0} \\ \hline \end{array} \,.$$

Again (u^3, v^3) is the u-worst corner (v-best corner) in D^2. To move inside D^2, we look for direction (s, t). Using the starred entries, (s, t) must satisfy $s_0 + t_1 \geq 4/3$, $s_1 + t_2 \geq 4/3$, $s_2 + t_0 \geq 4/3$. Also since $f_{23}(u, v) = f_{32}(u, v) \equiv 1$ on D^2, we easily find the above system of inequalities inconsistent. Thus no more movement inside is possible. We have reached the lexicographic center.

Remark 5.1. Starting with the worst set of commissions for all sellers and using Kuhn's Hungarian method [12], the algorithm locates the unique set of commissions that again favors all the buyers in the restricted new domain D of commissions. The next step is to locate the unique direction (s, t) and the unique step size β in finding the new set of commissions. We have not used in our example any efficient procedure to find the direction (s, t). Solymosi and Raghavan [26] develop an explicit graph-theoretic algorithm to find this direction. The decomposition of the payoff space and the lattice structure of the feasible set at each iteration are utilized in associating a directed graph. If the graph is acyclic, the problem of finding the new direction (s, t) can be transformed to determine the longest path to each vertex of the graph. Cycles are used to collapse vertices so that the graph has fewer vertices. The algorithm stops when the graph is reduced to just one vertex. The assignment game is the simplest of cooperative games which are balanced and hence have a nonempty core. The Real Estate Game was first considered by Shapley and Shubik [23]. The same problem was viewed in the context of competitive pricing of indivisible goods by Gale [6]. Pooling peoples' utility functions amounts to interpersonal comparisons and hence has remained alien to mainstream economists. For a version of the Real Estate Game without side payments see [22].

REFERENCES

[1] Aumann R.J. and Maschler M., The bargaining set for cooperative games, in M. Dresher, L.S. Shapley, and A.W. Tucker (eds.), *Advances in Game Theory*, volume 52 of *Annals of Mathematical Studies*, 443–476, Princeton University Press, Princeton, NJ, 1964.

[2] Balinsky M.L. and Gale D., On the core of the assignment game, in Lev J. Leifman (ed.), *Functional Analysis, Optimization and Mathematical Economics*, 274–289, Oxford University Press, New York, 1990.

[3] Biswas A.K., Parthasarathy T. and Potters J.A.M., Large cores and exactness, *International Journal of Game Theory*, 28, 1–12, 1999.

[4] Bondareva O.N., Some applications of linear programming methods to the theory of cooperative games, *Problemi Kibernitiki*, 10, 119–139, 1963.

[5] Davis M. and Maschler M., Existence of stable payoff configurations for cooperative games, in M. Shubik (ed.), *Essays in Mathematical Economics in Honor of Oskar Morgenstern*, 39–52, Princeton University Press, Princeton, NJ, 1967.

[6] Gale D., *The Theory of Linear Economic Models*, McGraw-Hill, New York, 1960.

[7] Granot D., On a new bargaining set for cooperative games, Working paper, Faculty of Commerce and Business Administration, University of British Columbia, Vancouver 1994.

[8] Hamers H., Klijn F., Solymosi T., Tijs S. and Villarz J.P., On the extreme points of the core of neighbour games and assignment games, Technical report, University of Tilburg, The Netherlands, 1999.

[9] Justman M., Iterative processes with 'nucleolar' restrictions, *International Journal of Game Theory*, 6, 189–212, 1977.

[10] Kikuta K. and Shapley L.S., Core stability in n-person games, mimeo, 1986.

[11] Kohlberg E., The nucleolus as a solution of a minimization problem, *SIAM Journal on Applied Mathematics*, 23, 34–39, 1972.

[12] Kuhn H.W., The Hungarian method for the assignment problem, *Naval Research Logistics Quarterly*, 2, 83–87, 1955.

[13] Lucas W.F., A game with no solution, *Bulletin of the American Mathematical Society*, 74, 237–239, 1968.

[14] Maschler M., Peleg B. and Shapley L.S., The kernel and bargaining set for convex games, *International Journal of Game Theory*, 1, 73–93, 1972.

[15] Nunez M. and Rafels C., The extreme core allocations of the assignment games, Technical report, 2003.

[16] Owen G., A note on the nucleolus, *International Journal of Game Theory*, 3, 101–103, 1974.

[17] Quint T., Characterization of cores of assignment games, *International Journal of Game Theory*, 19, 413–420, 1991.

[18] Raghavan T.E.S. and Sudhölter P., The modiclus and core stability, *International Journal of Game Theory*, 33, 467–478, 2005.

[19] Schmeidler D., The nucleolus of a characteristic function game, *SIAM Journal on Applied Mathematics*, 17, 1163–1170, 1969.

[20] Shapley L.S., On balanced sets and cores, *Naval Research Logistics Quarterly*, 14, 453–460, 1967.

[21] Shapley L.S., Cores of convex games, *International Journal of Game Theory*, 1, 11–26, 1971.

[22] Shapley L.S. and Scarf H.E., On cores and indivisibility, *Journal of Mathematical Economics*, 1, 23–28, 1974.

[23] Shapley L.S. and Shubik M., The assignment game i: The core, *International Journal of Game Theory*, 1, 111–130, 1972.

[24] Sharkey W.W., Cooperative games with large core, *International Journal of Game Theory*, 11, 175–182, 1982.

[25] Solymosi T., On the bargaining set, kernel and core of superadditive games, *International Journal of Game Theory*, 28, 229–240, 1999.

[26] Solymosi T. and Raghavan T.E.S., An algorithm for finding the nucleolus of assignment games, *International Journal of Game Theory*, 23, 119–143, 1994.

[27] Solymosi T. and Raghavan T.E.S., Assignment games with stable core, *International Journal of Game Theory*, 30, 177–185, 2001.

[28] Sudhölter P., The modified nucleolus as canonical representation of weighted majority games, *Mathematics of Operations Research*, 21, 734–756, 1996.

[29] Sudhölter P., The modified nucleolus: Properties and axiomatizations, *International Journal of Game Theory*, 26, 147–182, 1997.

[30] Sudhölter P., Equal treatment for both sides of assignment games in the modified least core, in M. Holler and G. Owen (eds.), *Power Indices and Coalition Formation*, 175–202, Kluwer Academic Publishers, Dordrecht, 2001.

[31] van Gellekom J.R.G., Potters J.A.M. and Reijnierse J.H., Prosperity properties of tu-games, *International Journal of Game Theory*, 28, 211–217, 1999.

[32] von Neumann J. and Morgenstern O., *Game Theory and Economic Behavior*, third edition, Princeton University Press, Princeton, NJ, 1953.

The Folk Theorems in the Framework of Evolution and Cooperation*

Alexander Vasin
Faculty of Computational Mathematics and Cybernetics
Moscow State University
Moscow 119899, Russia
and New Economic School, Moscow
vasin@cs.msu.ru

Abstract

Proceeding from the latest versions of the Folk theorems, this chapter shows that "natural" evolution of behavior in repeated games in human populations is a very unstable process which may be easily manipulated by outside forces. Any feasible and individually rational payoff of the game may be converted into a globally stable outcome by an arbitrary small perturbation of the payoff functions in the repeated game. We show that this result also holds for a trembling-hand perturbation of the game, and prove a new version of the Folk theorem for this case. This conclusion is in contrast to the result of R. Axelrod, K. Sigmund and M. A. Nowak and some other researches on the evolution of behavior in the repeated Prisoner's Dilemma. We discuss the reasons for these different results.

1 Introduction

The question: Does repetition lead to cooperation? has been widely discussed in the game-theoretic literature since the publication of the book by Axelrod [1]. Until now the answer has been ambiguous. Computer simulations started by this work and continued by Novak and Sigmund [9,10] confirm survival and superiority of the behavior strategies which lead to cooperation. This result was also supported by theoretic researches [4,3]. However, the Folk theorems

*This paper was completed during my visit to the University of Alicante. I am grateful to the *Department of Fundamental Economic Researches* for its hospitality. The work was supported by the Ministry of Education and Culture, and the Direccion General de Enseñanza Superior of Spain, and by the Russian Fund for Basic Researches, Grant 02-01-00610. I thank Arkadiy Kryazhimsky, the participants of the IIASA Workshop "Growth and Adaptation in Evolutionary Systems" and the participants of the Tenth International Symposium on Dynamic Games and Applications for their useful discussions.

for repeated games (see Van Damme [13] for a survey of this field) show that every individually rational outcome of a one-stage game can be supported at some Nash equilibrium of the repeated game. With respect to the evolution of cooperation, this means that cooperation is not distinguished among many other equilibrium forms of behavior.

The possibility has still remained that the cooperative outcome is the only stable one in some sense, or at least it has some robust domain of attraction. In this context, the recent version of the Folk theorem for dominance solutions [16,17] is of special interest. Besides other advantages, the concept of iterated strict dominance elimination (see [7]) is very useful for investigation of the game dynamics. For a wide class of game dynamical systems it is known that the frequencies of all eliminated strategies converge to 0 as time tends to infinity. This class includes, in particular, "Cournot tatonnement" [7], replicator dynamics [14] and selection dynamics [8].

Our result [17] establishes that the set of strict dominance solution payoffs of perturbed finitely repeated games converges to the set of individually rational convex combinations of payoffs in the stage game as the number of repetitions tends to infinity and the perturbation value tends to 0. With respect to evolution of cooperation, this contradicts the conjecture on the special status of cooperative behavior in game dynamics, at least for repeated games with complete information and unbounded rationality of players.

In a typical case, the construction of the dominance solution corresponding to a desirable behavior in the repeated game is similar to the recursive constructions of the subgame perfect equilibrium in Benoit and Krishna [2] and Fudenberg and Maskin [5]. At every stage each individual has to either realize the corresponding path or punish "the last disturber," i.e., the last player (with the smallest index) who deviated from the rule, if he has not already been sufficiently punished.

The most important innovation we introduced is a special end-game construction. It may be interpreted as the perturbation of the repeated game payoffs by some operating center (called the Manipulator), who rewards or penalizes players depending on their behavior during the game. The presence of such a center interested in the outcome of the game is typical for social interactions. For instance, recall the prosecutor in the original version of the Prisoner's Dilemma. He is obviously uninterested in "cooperative" behavior of the "players." The whole concept of cooperative behavior is doubtful in such situations involving several persons with different interests and asymmetric positions.

The players' behavior corresponding to subgame perfect equilibria in [2] and [5] (as well as the dominance solution in our paper [17]) is rather sophisticated. This chapter aims to provide a very simple construction of the dominance solution for every outcome where payoffs to all players are not less than their payoffs at some Nash equilibrium of the one-stage game. After any deviation, players just switch to playing the Nash equilibrium until the end of the repeated game. The "last disturber" loses his/her award. Thus, the idea of the solution is close to that of Radner [11], who considered the Nash equilibria of a similar perturbation of the repeated oligopoly game. This case covers all individually rational outcomes in the Prisoner's Dilemma and in the Coordination game with one efficient equilibrium. Then we generalize this result for a "trembling-hand" perturbation of a stage game according to Selten [12], that is, for the case where players mistake in their actions with a small probability. In conclusion we give a brief survey of results which confirm the convergence of evolutionary dynamics to cooperative behavior strategies and discuss why they differ from our results.

2 The Formal Definitions

Let Γ be a normal form game with the set of players $I = \{1, \dots, n\}$, the sets of strategies $X_1, \dots, X_n$ and the payoff functions $f_1(x), \dots, f_n(x)$, $x \in \bigotimes_i X_i$. Let $(x\|y_i)$ denote the result of substitution of y_i for x_i in the strategy combination $x = (x_i, i \in I)$.

Strategy x_i weakly dominates strategy y_i on the set $\overline{X} \subseteq X$ if there exists $\epsilon \geq 0$ such that for any $z \in \overline{X}$

$$f_i(z\|x_i) \geq f_i(z\|y_i) + \epsilon. \tag{1}$$

In contrast to the standard definition, we do not require the strict inequality for at least one $z \in \bar{X}$.

Strategy combination $\bar{x}$ is a weak dominance solution if it may be obtained by means of the weak dominance elimination procedure (see Moulin [7]). That is, if there exists a sequence of sets $X = X^1 \supset X^2 \supset \cdots \supset X^k = \{\overline{x}\}$, where, for every $l = 1, \dots, k-1$, $X^l = \bigotimes_i X_i^l$, for any $x_i \in X_i^l \setminus X_i^{l+1}$ there exists $y_i \in X_i^{l+1}$ that weakly dominates x_i on X^l ; x_i strictly dominates y_i on $\overline{X}$ if (1) holds for $\epsilon > 0$; $\overline{x}$ is a strict dominance solution, if it can be obtained by means of successive elimination of strictly dominated strategies.

Consider a T-fold repetition Γ_T of a normal form game Γ. In order to avoid confusion between strategies in the Γ_T and strategies in the initial game, the latter will be referred to as actions. At any time t every player knows the actions of all participants at previous periods. Let $x^t \in X$ be the action combination at time t, $h^t = (x^0, \dots, x^{t-1})$ a history at this time, X^t the t-fold Cartesian product of X, $H^T = \bigcup_{t=0}^{T-1} X^t$ the set of all histories to time T ($X^0 \overset{\text{def}}{=} \{0\}$). A strategy of player i is a mapping $m_i : H^{T-1} \to X_i$. This mapping determines the choice of the action for every time t and any history h^t. Each strategy combination $m = (m_i, i \in I)$ induces the path of Γ_T $h(m) = \{x^t(m)\}$, where $x^0(m) = 0$, $x^1(m) = m(0)$, $x^t(m) = m(x^0(m), \dots, x^{t-1}(m))$, $t \geq 1$. The payoff function of player i is given by

$$F_{i,T}(m) = F_{i,T}(h(m)) = \sum_{t=0}^{T-1} f_i(x^t(m))/T.$$

For any game Γ with sets of strategies Y_i and payoff functions $u_i(y)$, $i = 1, \dots, n$, and for any $\epsilon > 0$, let $A(\Gamma, \epsilon)$ denote the set of games with the same sets of players and strategies and payoff functions $u'_i(y)$ which differ from $u_i(y)$ by less than ϵ. Since the repeated game Γ_T is a particular case of a normal form game, we may consider a set $A(\Gamma_T, \epsilon)$ of its perturbations. Let us stress that any perturbed repeated game $\hat{\Gamma} \in A(\tilde{\Gamma}, \epsilon)$ is obtained from $\tilde{\Gamma}$ by variation of its payoff functions. The set $A(\Gamma_T, \epsilon)$ is wider than the set of repeated games corresponding to the perturbed stage games $\Gamma' \in A(\Gamma, \epsilon)$.

Our result in Ref. [17] describes the limit set of the dominance solution payoffs of games $\hat{\Gamma} \in A(\Gamma_T, \epsilon)$ as T tends to infinity and ϵ tends to 0. Let v_i be the minmax payoff of player i:

$$v_i = \min_x \max_{x_i} f_i(x \| x_i).$$

Let $W = \{f(x), x \in X\}$ be the set of payoff vectors for pure strategies, let CoW denote the convex hull of W and let $\Phi = \{w \in CoW \mid w \geq v\}$ be the set of individually rational convex combinations of payoff vectors in the stage game Γ; $M = \max_{x,z \in X} |f(x) - f(z)|$.

Theorem 2.1. *The set of strict dominance solution payoff vectors of games $\hat{\Gamma} \in A(\Gamma_T, \epsilon)$ converges (in the sense of Hausdorff distance) to the set Φ as T tends to ∞ and ϵ tends to* 0.

Proof. The proof is given by Vasin in [17]. □

The construction of the dominance solution and the payoff perturbation are rather sophisticated in the general case. The next section studies the problem for a one-stage game Γ with a Nash equilibrium $\bar{x}$.

3 The Dominance Solution Supporting an Outcome Which Dominates Some Nash Equilibrium

For any sequence $z(1), \dots, z(r)$ of action profiles such that $\sum_k f(z(k))/r \geq f(\bar{x})$ we describe a simple payoff perturbation and a strategy profile m^* which supports repetition of this sequence at the dominance solution of the repeated game with the perturbed payoff functions.

In order to explain the construction of the dominance solution corresponding to a sequence of action profiles, let us consider a version of the coordination game with the set of actions $\{1, 2, 3\}$ and the payoff matrix

$$\begin{array}{cccc} & 1 & 2 & 3 \\ 1 & (5,5,0) & (1,0,0) & (0,3,5) \\ 2 & (0,1,0) & (1,1,2) & (0,0,0) \\ 3 & (3,0,5) & (0,0,0) & (0,0,0) \end{array}$$

where the third component shows the gain of the Manipulator who does not act but is interested in his payoff. More precisely, we assume that he would like to maximize his average payoff in the T-fold repetition. Thus, action profiles $(1, 1)$ and $(2, 2)$ are resp. the "cooperative" and the "bad" Nash equilibrium for the active players, and alternation between $(1, 3)$ and $(3, 1)$ is optimal for the Manipulator among sequences of action profiles which are individually rational for players $1, 2$.

If the Manipulator does not interfere then the natural outcome of the game is repetition of the cooperative equilibrium. Our purpose is to describe a perturbation of the payoff function in Γ_T which supports alternation between $(1, 3)$ and $(3, 1)$ at the dominance solution of the perturbed repeated game. The corresponding strategy of each player is to repeat his/her actions in this sequence until the first deviation and to play constantly action 2 after the deviation if it happens. The payoff function perturbation may be interpreted as "awards" which the Manipulator pays to the players for "good behavior" in this game. The value of the "award" is $6/T$. If nobody deviates from alternating between $(1, 3)$ and $(3, 1)$ then both players get their awards. Otherwise, the player who is the last to deviate from the specified strategies is penalized. If the last deviation involved both players simultaneously then only player 1 does not get the award.

The weak dominance elimination may be realized as follows. At first we can exclude any strategy such that the first player deviates at time $T - 1$ for some history. If player 1 is the "last disturber" according to history and player 2 uses action 2 at this time then this deviation is unprofitable. Otherwise, player 1 gains at most $5/T$ by deviation but loses $6/T$ because he/she either becomes the last disturber or misses the chance to pass this label to player 2. At the next stage every strategy which permits deviation by player 2 at time $T - 2$ is eliminated in a similar way. Now we can continue this reasoning for the time $T-2$, $T-3$, etc. After $2T$ stages of eliminations, we obtain the desirable solution. Thus, by spending $12/T$ for awards, the Manipulator increases his average gain from 0 to 5 in this example.

Now, consider a general case where desirable behavior corresponds to repetition of $z(1), \ldots, z(r)$. Then the strategy profile is as follows. The players repeat this sequence until at least one of them deviates from this choice. Then they switch to playing $\bar{x}$ at every repetition until the end of the game. In order to define the payoff perturbation, consider the last disturber, that is, the last player (with a minimal index) not to conform to the specified behavior rule.

This player is penalized with the fine rM. More formally,

$$m_i^*(h^t) = z_i(\{t \ mod \ m\}+1) \text{ if } x^s = z_i(\{s \ mod \ m\}+1) \text{ for any } s < t, \\ \text{otherwise } m_i^*(h^t) = \bar{x}_i. \quad (2)$$

For any path h^T, let

$$s(h^t) = \max\{s < t \text{ s.t. } x^s \neq m^*(h^s)\}$$

denote the last time of deviation before t, and

$$i(h^t) = \min\{i \text{ s.t. } x_i^{s(h^t)} \neq m_i^*(h^{s(h^t)})\}$$

denote the last disturber to time t. The fine for deviation is

$$\varphi_i(h^T) = rM/T \text{ for } i = i(h^T)$$

and is 0 for any other i. Finally, the perturbed payoff function is

$$\hat{F}_i(m) = F_{i,T}(m) - \varphi_i(h(m)), i = 1, \ldots, n. \quad (3)$$

Proposition 3.1. *Strategy profile m^* defined according to (2) is a weak dominance solution of game $\hat{\Gamma}$ with payoff functions (3).*

Proof. The proof will repeat the arguments of the preceding example. Consider the player 1 and any strategy m_1 such that $m_1(h^{T-1}) \neq m_1^*(h^{T-1})$ for some h^{T-1}. Let us show that strategy $\bar{m}^1$, s.t. $\bar{m}_1(h^t) = m_1(h^t)$ for any $t < T-1$ and $\bar{m}_1(h^{T-1}) = m_1^*(h^{T-1})$ for any h^{T-1}, weakly dominates m_1. Consider any strategy profile m including m_1. If $h^{T-1}(m) = h^{T-1}(m^*)$ then deviating from $m_1^*(h^{T-1}(m))$ is unprofitable since it makes player 1 the last disturber, and he has to pay the fine rM/T while by deviating he gains at most M/T. Otherwise, consider the last disturber to the time $T-1$. If $i(m, T-1) \neq 1$ or $x_j(h^{T-1}(m)) \neq \bar{x}_j$ for some $j \neq 1$ then by deviating player 1 is missing the chance to pass the label of the last disturber to another player, and we can argue as in the previous case. Otherwise, deviating brings nothing since $\bar{x}$ is a Nash equilibrium. We can continue this reasoning by induction for $i = 2, 3, \ldots, n$ and then for $t = T-2, \ldots, 0$. The only correction is that by deviating at time $T-k$ a player can gain at most $\min\{k, r\}M/T$. □

4 The Theorem for a Trembling-Hand Perturbation of the Model

Random mistakes introduce some difficulties in the proposed scheme of behavior regulation. Consider a game $G_{T,d}$ which is a trembling-hand perturbation of the repeated game Γ_T according to Selten [12]. In this game, for any history h^t and strategy m_i, the action of player i in period t is $m_i(h^t)$ with probability $1-d$, other action $x_i \neq m_i(h^t)$ with probability $d/(|X_i|-1)$. Thus m determines a probability distribution $P_d(h^T \mid m)$ over the set of paths h^T. Let each player be interested in his/her expected payoff

$$g_i(m) = \sum_{h^T} P_d(h^T \mid m) F_i(h^T).$$

Assume that the desirable behavior of the players corresponds to action profile x^* such that $f(x^*) \geq f(\bar{x})$ for some Nash equilibrium $\bar{x}$. Consider the strategy profile m^* corresponding to x^* according to the definition of Section 3. Let us define the perturbed payoff functions as follows. For every path h^T, we can identify the last disturber up to time T. Let him pay the fine $\bar{\varphi}$ while other players pay nothing. Formally, the value of the fine is given by the following function:

$$\varphi_i(h^T) = \bar{\varphi} \text{ if } i = i(h^T), \text{ otherwise } \varphi_i(h^T) = 0. \tag{4}$$

Consider a game $\hat{G^T}$ with payoff functions

$$\hat{g}_i(m) = \sum_{h^T} p(h^T \mid m)(F_i(h^T) - \varphi_i(h^T)).$$

Proposition 4.1. *Let the value of the fine $\bar{\varphi} > M/(Ta^T)$, where $a \stackrel{\text{def}}{=} (1-d)^n$. Then m^* is a strict dominance solution of the game $\hat{G^T}$ if $1-d > d/|X_i|$, $i = 1, \ldots, n$.*

Proof. Assume that for every $i = 1, \ldots, n$ we have already eliminated all strategies m_i such that $m_i(h^\tau) \neq m_i^*(h^\tau)$ for some $\tau > t$, h^τ. Consider any h^t and m_1 such that $m_1(h^t) = z_1 \neq m_1^*(h^t)$. Let us show that $\bar{m}_1$, such that $\bar{m}_1(h^t) = m_1^*(h^t)$, $\bar{m}_1(\acute{h}^\tau) = m_1(\acute{h}^\tau)$ for any other $\tau, \acute{h}$, strictly dominates m_1. First consider the case where 1 is the last disturber after h^t. Then, according to m^*, the players have to play $\bar{x}$ from t until the end of the game. Let $\hat{g}_i(h^t, m)$ denote the expected gain of player i in $\hat{G}$ under history h^t and profile m. Then, for any $d > 0$,

$$\begin{aligned}&\hat{g}_1(h^t, m \mid\mid \bar{m}_1) - \hat{g}_1(h^t, m)\\ &\qquad > \min_x((1-d)^2 - (d/\mid X_i \mid)^2)(\hat{g}_1(h^t, x \mid\mid \bar{x}_1, m^*) - \hat{g}(h_1^t, x \mid\mid z_1, m^*)).\end{aligned}$$

Consider the last difference. If $x = \bar{x}$ then player 1 gains nothing by deviating since $\bar{x}$ is a Nash equilibrium, and behavior from time $t+1$ on does not depend on x_1^t. Othervise, by deviating from $\bar{x}_1$, player 1 gains at most M but increases the probability to become the last disturber at time T and to pay the fine $\bar{\varphi}$ in $a^{(T-t-1)}$. Thus, $\bar{m}_1 \succ m_1$.

The same reasoning works if some other player is the last disturber at time t. If nobody is, then the only difference is that the expected total gain in repetitions since $t+1$ decreases by deviation since the players will have to switch from x^* to $\bar{x}$. We can continue this argument for players $2, \dots, n$ and thus complete the proof. □

Proposition 4.2. *The mean time of playing x^* in the game $\hat{G}$ under the strategy profile m^* is*

$$\tau(T) = (1-a^T)/(1-a) - Ta^{T-1}(1-a). \tag{5}$$

Proof. The probability of playing x^* t times, $t < T$, is $a^t(1-a)$, and for T this probability is a^T. Thus, the mean time is $(1-a)\sum_1^{T-1} \tau a^\tau + Ta^T$ which coincides with (5). □

Proceeding from Propositions 4.1 and 4.2, the profile m^* does not ensure the desirable behavior for $T > 1/d$. Let $T = k/d$ for some $k > 1$. Then, for d small enough, $a \approx 1 - nd$, $a^T \approx e^{-nk}$, the mean share of the time of playing x^* is $\tau(T)/T \approx (1-e^{-nk})/nk - n(1-nd)e^{-nk}$ and tends to 0 as k tends to infinity, while the necessary amount of the fine is about dMe^{kn}/k and tends to infinity.

One possibile way to improve the strategy is by partitioning the time of the game into s intervals of the same length $\bar{T} = k/d$ and playing the profile m^* independently in each interval. Let m^{**} denote this strategy combination. The corresponding payoff perturbation assumes that, after each interval, the players pay the fine (4) according to their behavior within this interval. Let $\varphi^{**}(h^T)$ denote the corresponding perturbation of the total payoff. Then the maximal value of the perturbation is about dMe^{kn}/k for any small d and $k \sim O(1/d^{1/s})$, $s \in Z$. Under m^{**}, the mean time of playing x^* related to T is the same as in (5). Thus, for any game Γ with a Nash equilibrium $\bar{x}$, we obtain the following result.

Theorem 4.1. *For any $\varepsilon > 0$, there exist $\acute{d} > 0$ and $\acute{T}$ such that for any $d < \acute{d}$ and $T > \acute{T}$, every action profile x^* such that $f_i(x^*) \geq f_i(\bar{x})$, $i = 1, \dots, n$, may be supported at the strict dominance solution m of game $\hat{G} \in A(G_T, \varepsilon)$ such that the mean share of the time of playing x^* under m exceeds $1-\varepsilon$.*

Proof. Consider the game $\hat{G}_T$ with the payoff perturbation ϕ^{**}. For d small enough, the strategy profile m^{**} is a strict dominance solution if the fine exceeds dMe^{kn}/k. The mean share of the time of playing x^* is given by (4). In order to make this share more than $1-\varepsilon$, let us set k such that $1-e^{-nk} > nk(1-\varepsilon)$. Since $e^{-nk} < 1 - nk + nk^2/2$ it suffices to set $k = 2\varepsilon/n$. Now, in order to make the total payoff perturbation less than ε, choose d such that $ndMe^{2\varepsilon}/2\varepsilon < \varepsilon$. It suffices to take $\bar{d} = \varepsilon^2/nM$. Finally, T should be large enough to divide it into equal intervals of length $k/d = 4M/\varepsilon$. □

Corollary 4.1. *For any game Γ with Nash equilibrium $\bar{x}$ such that $v_i = f_i(\bar{x})$, $i = 1, \ldots, n$, consider a trembling-hand perturbation G_T of a T-fold repetition of Γ. The set of the strict dominance solution payoffs in games $\hat{G}_T \in A(G_T, \varepsilon)$ converges to the set Φ of individually rational convex combinations of payoffs in Γ as T tends to infinity and ε tends to 0.*

Proof. The only difference in the proof of the theorem is that, instead of playing x^* in every repetition, the desirable behavior in this case is a repeated sequence of action profiles $z^1, \ldots, z^r$ such that $\sum_k f(z^k)/r$ approximates a given $w \in \Phi$. The minor modifications include r-fold increasing of the fine. Thus, we have proved the Folk theorem for the dominance solutions of games in the specified class. □

5 Discussion

One important direction in studying the evolution of cooperation is computer simulation of behavior dynamics in the iterated Prisoner's Dilemma. Two players engaged in the Prisoner's Dilemma have to choose between cooperation (C) and defection (D). At any given round, each player receives R if both cooperate and payoff $P < R$ if both defect; but a defector exploiting a cooperator gets T points, while the cooperator receives S (with $T > R > P > S$ and $2R > T + S$). Thus in a single round it is always best to defect, but cooperation may be rewarded in an iterated Prisoner's Dilemma. Axelrod's computer tournaments [1] have shown that the known "tit for tat" strategy which supports cooperation wins the competition with other deterministic strategies where the decision to cooperate or defect at each round depends on the outcome of the three previous rounds. Nowak and Sigmund [9] have considered the case where the decision depends only on the previous round, but is stochastic and not deterministic. Each strategy consisted of two parameters, p^C and p^D. These give the probability of cooperating after a C (cooperate) or D (defect) by the other player. In these simulations Generous-TFT ($p^C = 1$ and $p^D = 1/3$ given the usual payoffs) appeared to be a stable end state, as almost any starting condition converged to it provided the run was long enough.

Later Nowak and Sigmund [10] extended this treatment to "two-step memory" strategies, in which players base their decision on the other player's action as well as on their own previous action. They found a different strategy dominating the long-term behavior. They called this strategy "Pavlov" or "win-stay, lose-shift"(WSLS) because if it receives a good payoff (either T or R) it repeats its previous action (D in the former and C in the latter). Conversely if it receives a low payoff (P or S) it prefers to change its behavior next time.

A general conclusion from these models is that strategies which support cooperation dominate in the long-run prospect. This proposition is supported by several theoretical results. Fudenberg and Maskin [4] and Binmore and Samuelson [3] consider different variants of supergames with time-averaged payoffs, bounded rationality and symmetry of players and establish that the stable behavior is connected with the "utilitarian" payoff vectors which maximize the sum of player's payoffs. The notion of a utilitarian outcome generalizes the concept of cooperative behavior for symmetric two-player games. Thus, all mentioned papers give the impression that evolution of behavior in repeated games leads to cooperation.

The main reason for this contrast to our conclusions is the bounded rationality of players in the mentioned models. The strategy combinations specified in Sections 3 and 4 become the dominance solutions only if the behavior of the players is flexible with respect to the time until "the day of reckoning" T. If we consider players with bounded rationality, according to Kalai and Stanford [6], then m^* is not a dominance solution for sufficiently large T. But the specified flexibility seems to be typical for human behavior.

A reasonable question about the proposed scheme of behavior manipulation is if the active players can withstand it in the case where the imposed behavior is unprofitable for them, as in the example above. Of course, if some of them are sufficiently intelligent and have enough free money then they can create a "counter-manipulator" supporting cooperative behavior. One thing we would like to stress is that the cooperation (as well as other nice things) needs some regulation to support it.

REFERENCES

[1] Axelrod R., *The Evolution of Cooperation*, Basic Books, New York, 1984.

[2] Benoit G.-P. and Krishna V., Finitely repeated games, *Econometrica*, 53, 905–922, 1984.

[3] Binmore K.J. and Samuelson L., Evolutionary stability in repeated games played by finite automata, *J. of Econ. Theory*, 57, 278–305, 1992.

[4] Fudenberg D. and Maskin E., Evolution and cooperation in noisy repeated games, *New Developments in Economic Studies*, 38, 1–12, 1990.

[5] Fudenberg D. and Maskin E., The Folk theorem in repeated game with discounting and with incomplete information, *Econometrica*, 54, 533–554, 1986.

[6] Kalai E. and Stanford W., Finite rationality and interpersonal complexity in repeated games, *Econometrica* 56, 397–410, 1988.

[7] Moulin H., Dominance-solvability and Cournot-stability, *Mathematical Social Sciences*, 7, 83–102, 1984.

[8] Nachbar J.H., "Evolutionary" selection dynamics in games: Convergence and limit properties, *Int. J. of Game Theory*, 4, 25–55, 1990.

[9] Nowak M.A. and Sigmund K., The evolution in stochastic strategies in the prisoner's dilemma, *Acta Appl. Math.*, 20, 247–265, 1990.

[10] Nowak M.A. and Sigmund K., Tit for tat in heterogeneous populations, *Nature*, 355, 250–253, 1992.

[11] Radner R., Collusive behavior in non-cooperative epsilon-equilibria of oligopolies with long but finite lives, *J. of Econ. Theory*, 22, 136–154, 1980.

[12] Selten R., Re-examination of the perfectness concept for finite points in extensive games, *Int. J. of Game Theory*, 4, 25–55, 1975.

[13] Van Damme E., *Stability and perfection of Nash equilibria*, Springer-Verlag, Berlin, 1987.

[14] Vasin A.A., Methods of game theory in studies of collective behavioral dynamics, *Moscow University Computational Mathematics and Cybernetics Bulletin*, N 3, 51–55, 1989.

[15] Vasin A.A., Evolutionary model of behavior in a supergame, *Moscow University Computational Mathematics and Cybernetics Bulletin*, N 3, 54–60, 1990.

[16] Vasin A.A., New version of the Folk theorem, in M. Breton and G. Zaccour (eds.), *6th International Symposium on Dynamic Games and Applications*, Preprint volume, St. Jovite, Canada, 1994.

[17] Vasin A.A., The Folk theorem for dominance solutions, *Int. J. of Game Theory*, 28, 15–24, 1999.

Stackelberg Problems: Subgame Perfect Equilibria via Tikhonov Regularization

Jacqueline Morgan
Università di Napoli "Federico II"
Dipartimento di Matematica e Statistica
Via Cintia, 80126 Napoli, Italy
morgan@unina.it

Fioravante Patrone
Università di Genova
Dipartimento di Matematica
Via Dodecaneso 35, 16146 Genova, Italy
patrone@dima.unige.it

Abstract

In this chapter we consider a two-stage game with one leader and one (or more) followers and we investigate the behavior of a Tikhonov regularization when the best reply for the follower(s) is not uniquely determined. More precisely, we show, under mild assumptions in the case of one follower and sufficiently mild in the case of two followers, that a convergent sequence of solutions to regularized two-stage games generates a subgame perfect equilibrium (SPE) of the original game, providing a constructive way to approach an SPE in a continuous setting. Various elementary examples show that our results cannot be strengthened up to guaranteeing convergence to a strong or a weak Stackelberg equilibrium and that the method cannot be extended to all of the cases in which two followers play a mixed extension of a finite game.

1 Introduction

A standard way to interpret a Stackelberg equilibrium for strategic form games is to see it as a subgame perfect equilibrium (SPE for short: [22]) of a perfect information game in extensive form where the leader moves first.

This interpretation is quite interesting, and is unambiguous when we have uniqueness for the best reply of the follower. Also, in such a case, the assumptions needed to guarantee existence of the solution to the Stackelberg problem and for its numerical computation are reasonably mild (see [16]).

In the cases in which the best reply for the follower is not uniquely determined, some problems may arise, however. One source of difficulty is that the leader cannot predict the follower's choice simply on the basis of his rational behavior. So, the choice of the best strategy from the leader's point of view can become quite problematic. Two approaches, introduced and called the generalized Stackelberg strategies in [6], are the most common: to assume a "pessimistic" attitude for the leader (which leads to the weak Stackelberg problem), or an optimistic one (leading to the strong Stackelberg problem). See also [1,2], to which the terminology is due.

These approaches assume that the follower will choose the worst (respectively: best) action for the leader, in case of indifference. They define unambiguously the best choice for the leader in simple settings (e.g., in the case of finite sets of actions available to the players). Of course, even in this very simple setting, the predicted choice for the leader can be different (see Example 2.1).

However, as soon as one leaves this simple case, some problems remain.

A difficulty is due to the fact that, even under quite reasonable compactness and continuity assumptions, existence (and continuity) of a solution to the weak Stackelberg problem cannot be guaranteed (see [1], Remark 3, p. 178, for the lack of existence) and for the strong Stackelberg problem, even if existence holds, it does not display continuous dependence on the data of the problem ([9], Remark 2.1). These issues have been considered and the difficulties partially overcome using approximate solutions (see [8,9,11]). This approach, which can be considered as very natural from a numerical viewpoint, does not solve the uniqueness problem for the best reply, however. So, difficulties remain from the point of view of the numerical solution of the problem. In fact, the algorithms that have been proposed generally suffer from the already-mentioned lack of continuity for the exact solutions to these Stackelberg problems.

To overcome the numerical difficulties due to the nonuniqueness of the best reply, regularization methods have been suggested to tackle this problem, like Tikhonov regularization [12], the approaches followed by Dempe [5] to solve the strong Stackelberg problem and the approaches of Molodtsov [15], Soholovic [23] and Loridan and Morgan [12,13].

The nonexistence of a solution to the weak Stackelberg problem imposes the following question:

> If the solutions of regularization methods converge somewhere when the regularization parameter goes to zero, to what limit solution do they converge?

It is easy to give examples showing that the resulting limit point can be a weak Stackelberg, a strong one, or neither. A partial answer to this question was given in [12] for Tikhonov regularization: these authors defined the "lower Stackelberg equilibrium pair" and showed that any limit point obtained via Tikhonov regularization, when the regularization parameter goes to zero, is a lower Stackelberg equilibrium pair. Here we shall prove that for Tikhonov regularization the limit points obtained admit a nice interpretation: they are the SPE of the two-stage game referred to above. Note that the Tikhonov regularization method applies to a wider class of problems than Dempe's and Molodtsov's approach, since it does not require any kind of strong convexity assumption for the leader's functional in order to obtain the uniqueness of the best reply. Thus, it offers more options to the numerical solutions of Stackelberg problems.

Notice that the result of more than theoretical interest since it provides first of all a constructive way to get (approach) an SPE in a continuous setting. Of course, any solution to the strong Stackelberg problem is also an SPE: but in the case of nonuniqueness for the best reply, we recall that the strong Stackelberg problem does not have a nice behavior from the point of view of stability, thus provoking difficulties for its numerical solution.

We shall give versions of our results for both the unconstrained and the constrained cases.

2 Problem Setting

Let U and V be two topological spaces, X, Y be two nonempty subsets respectively of U and V and l and f be two functions from $U \times V$ to $\mathbf{R} \cup \{+\infty\}$. First, we consider the following problem, called the weak Stackelberg problem:

$$(w\text{-}S)\begin{cases}\text{Min}_{x\in X}\ \sup_{y\in M_2(x)}\ l(x,y)\\ \text{where}\ \ M_2(x)\quad \text{is the set of optimal solutions to the problem}\\ P(x):\text{Min}_{y\in Y}\ f(x,y)\end{cases}$$

$v = \inf_{x\in X}\ \sup_{y\in M_2(x)}\ l(x,y)$ is called the value of problem $(w\text{-}S)$.

The problem $(w\text{-}S)$ may have no solutions even for nice functions l and f, so, for all $\epsilon > 0$, the following regularized problem:

$$(w\text{-}S(\epsilon))\begin{cases}\text{Min}_{x\in X}\ \sup_{y\in M_2(x,\epsilon)}\ l(x,y)\\ \text{where}\ \ M_2(x,\epsilon)\ \text{is the set of}\ \epsilon\text{-solutions to}\\ P(x):\text{Min}_{y\in Y}\ f(x,y)\end{cases}$$

has been considered to obtain sufficient conditions ensuring existence and stability of the solutions to the regularized problem under data perturbations (Lignola and Morgan [8,9], Loridan and Morgan [11]). In our opinion these theoretical results offer an insight into the inherent difficulties of the problem and can explain the lack of nonheuristic numerical methods in the continuous case.

Fortunately, the following problem, called the strong Stackelberg problem, appears to be better handled:

$$(s\text{-}S)\begin{cases}\text{Min}_{x\in X}\ \inf_{y\in M_2(x)}\ l(x,y)\\ \text{where}\ \ M_2(x)\ \ \text{is the set of optimal solutions to}\\ P(x):\text{Min}_{y\in Y}\ f(x,y)\end{cases}$$

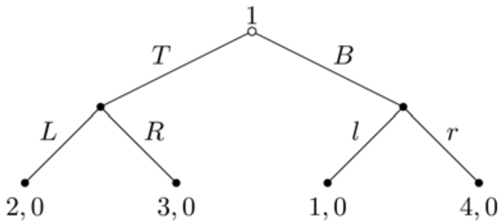


Figure 1: An example of the weak and strong Stackelbeg approaches

In fact, for such a problem under inequality constraints, in addition to existence results and a few stability results [9] there are numerous papers on necessary and sufficient conditions and numerical methods. But the method described in Molodtsov [15], which approaches the weak problem (w-S) by a sequence of strong Stackelberg problems, appeared to be a step towards the numerical resolution of the problem (w-S). Further results in this direction have been given in Loridan–Morgan [13].

As far as the SPE is concerned, we recall that it is a pair (x_0, y_0), where: $x_0 \in X$ and $y_0 : X \to Y$ satisfy the following conditions:

SPE1: $y_0(x) \in M_2(x) \quad \forall x \in X$
SPE2: $l(x_0, y_0(x_0)) = \inf_{x \in X} l(x, y_0(x))$

We recall that, under the assumption of uniqueness for the best reply of the follower, an SPE identifies a Stackelberg equilibrium for the game (X, Y, l, f): given (x_0, y_0) as before, the pair $(x_0, y_0(x_0)) \in X \times Y$ is a Stackelberg equilibrium.

In the following examples the players are assumed to maximize their payoffs.

Example 2.1. It is easy to provide an example in which the predictions offered by the weak and strong Stackelberg approaches differ. Just consider the game in Figure 1, described in extensive form.

The weak Stackelberg approach drives the leader's choice to T, with an expected result for player 1 of 2, while the strong Stackelberg approach leads to the choice of B, with an expected reward of 4.

Notice that this game has four Nash equilibria, all of which are also SPEs: (T, Ll), (B, Lr), (T, Rl), (B, Rr).

3 Results for the Case of One Follower

We shall assume the following:
X is a sequentially compact subset of U and Y is a compact, convex and nonempty subset of a finite-dimensional euclidean space V.

We are given:

$$l, f : X \times Y \to \mathbf{R} \cup \{+\infty\}$$

$$\alpha \in]0, +\infty[.$$

Let $M_2(x)$ be the set of solutions to the minimum problem

$$P(x) : \operatorname*{Min}_{y \in Y} f(x, y)$$

and $\hat{y}(x)$ be the minimum norm solution to $P(x)$. We consider the following regularized second-level problem, assumed to have a unique solution denoted as $\bar{y}_\alpha(x)$:

$$P_\alpha(x) : \qquad \operatorname*{Min}_{y \in Y} f(x, y) + \alpha \|y\|^2$$

A solution to the regularized Stackelberg problem:

$$(S_\alpha) \quad \operatorname*{Min}_{x \in X} f(x, \bar{y}_\alpha(x))$$

is denoted as $\bar{x}_\alpha$.

We recall that an SPE is a pair (x_0, y_0), where: $x_0 \in X$ and $y_0 : X \to Y$ satisfy the following conditions:

SPE1: $\quad y_0(x) \in M_2(x) \quad \forall x \in X$
SPE2: $\quad f(x_0, y_0(x_0)) = \inf_{x \in X} f(x, y_0(x))$

We shall make the following assumptions:

(H1) l is sequentially lower semicontinuous on $X \times Y$.
(H2) $l(x, \cdot)$ is sequentially upper semicontinuous on Y for all $x \in X$.
(H3) f is sequentially lower semicontinuous on $X \times Y$.
(H4) for all $(x, y) \in X \times Y$ and all sequences $x_n \to x$, there exists a sequence $(y_n)_n \in Y$ subject to (s.t.) $\limsup f(x_n, y_n) \leq f(x, y)$.
(H5) $f(x, \cdot)$ is convex with nonempty domain, for all $x \in X$.

As shown by the following example, assumption (H4) is weaker than the upper semicontinuity of the function $x \to f(x, y)$ for all $y \in Y$.

Example 3.1. Let $X = Y = [0, 1]$; $f(x, y) = -y^2 + (1 + x)y - x$ if $x \neq 0$ and $f(x, y) = 0$ if $x = 0$.

We shall prove the following theorem.

Theorem 3.1. *Assume that assumptions (H1) through (H5) are satisfied. Let $\alpha_n \downarrow 0^+$ and let*

$$\begin{cases} \bar{x}_n \text{ be a solution to } (S_{\alpha_n}) \\ \bar{y}_n \text{ be the solution } \bar{y}_{\alpha_n}(\bar{x}_n) \text{ to } P_{\alpha_n}(\bar{x}_n) \end{cases}$$

If $(\bar{x}_n, \bar{y}_n) \to (\bar{x}, \bar{y}) \in X \times Y$, then $(\bar{x}, \bar{y})$ generates an SPE $(\bar{x}, \tilde{y}(\cdot))$, where $\tilde{y}(\cdot)$ is defined as

$$\begin{cases} \tilde{y}(\bar{x}) & = & \bar{y} \\ \tilde{y}(x) & = & \hat{y}(x) \quad \forall x \in X, x \neq \bar{x} \end{cases}$$

(remember that $\hat{y}(x)$ is the element of minimum norm in $M_2(x)$).

Proof. Under assumptions (H1), (H3), (H4) and (H5) the problem (S_α) has a solution (see [16], Corollary 5.1).

Assume that $(\bar{x}_n, \bar{y}_n) \to (\bar{x}, \bar{y})$. We have to show that SPE1 and SPE2 are satisfied.

For SPE1, to prove that $\bar{y} \in M_2(\bar{x})$ it is enough to notice that under assumptions (H3) and (H4) the multifunction M_2 is sequentially closed at $\bar{x}$ [16]. That $\tilde{y}(x) = \hat{y}(x) \in M_2(x)$, when $x \neq \bar{x}$, is obvious by definition (simply notice that $M_2(x)$ is closed, convex and nonempty, so that it has a minimum norm element).

For SPE2, we have to prove that $l(\bar{x}, \bar{y}) = l(\bar{x}, \tilde{y}(\bar{x})) \leq l(x, \tilde{y}(x)) \quad \forall x \in X$, that is: $l(\bar{x}, \bar{y}) \leq l(x, \hat{y}(x)) \quad \forall x \in X, x \neq \bar{x}$.

We have that

$$(\bar{x}_n, \bar{y}_n) \text{ by definition satisfies } l(\bar{x}_n, \bar{y}_n) \leq l(x, y_{\alpha_n}(x)) \quad \forall x \in X;$$
$$y_{\alpha_n}(x) \to \hat{y}(x) \text{ for } \alpha_n \downarrow 0^+.$$

Hence,

$$\begin{aligned} l(\bar{x}, \bar{y}) &\leq \liminf_{n\to\infty} l(\bar{x}_n, \bar{y}_n) && \text{for (H1)} \\ &\leq \limsup_{n\to\infty} l(x, y_{\alpha_n}(x)) \\ &\leq l(x, \hat{y}(x)). && \text{for (H2)} \end{aligned}$$

□

We now provide a few elementary examples to show that our result cannot be strengthened up to guaranteeing convergence to a strong or weak Stackelberg equilibrium.

Example 3.2. We are given $l, f : [-1/2, 1/2] \times [-1, 1] \to \mathbb{R}$, with $l(x, y) = -(x + y)$, $f(x, y) = xy$.

The strong Stackelberg problem has the unique solution $(0, 1)$, whereas the weak Stackelberg problem does not have a solution. The pair $(\bar{x}_n, \bar{y}_n)$, the solution of the regularized problem, converges to $(0, 1)$, the solution of the strong Stackelberg problem. Of course, according to our result, the pair $(0, 1)$ can be extended to an SPE.

Example 3.3. Let us modify f of the previous example as follows:

$$f(x, y) := \begin{cases} (x + 1/4)y & \text{if } x \in [-1/2, -1/4] \\ 0 & \text{if } x \in [-1/4, 1/4] \\ (x - 1/4)y & \text{if } x \in [1/4, 1/2] \end{cases}.$$

The solution of the regularized problem converges to $(-1/4, 1)$, which does not coincide with the solution for the strong Stackelberg problem, which is $(1/4, 1)$.

Example 3.4. Let's modify further the example. Assume that $X = [-2, 2]$ and let

$$f(x,y) := \begin{cases} (x+7/4)y & \text{if } x \in [-2,-7/4] \\ 0 & \text{if } x \in [-7/4, 7/4] \\ (x-7/4)y & \text{if } x \in [7/4, 2] \end{cases}.$$

In such a case, the solution for the regularized problem converges to $(7/4, 0)$, which is different from the strong Stackelberg equilibrium $(2, -1)$ and from the weak Stackelberg equilibrium $(7/4, 1)$.

Assume now that, for all $x \in X, P(x)$ and $P_\alpha(x)$ we have the following constrained minimum problems:

$$P(x): \quad \operatorname*{Min}_{y \in K(x)} f(x,y);$$
$$P_\alpha(x): \quad \operatorname*{Min}_{y \in K(x)} f(x,y) + \alpha \|y\|^2,$$

where K is a set-valued function from X to V.

Then, we have the following theorem.

Theorem 3.2. *Assume that assumptions (H1),(H2), (H5) and the following are satisfied:*

(H6) *f is sequentially continuous on $X \times Y$.*
(H7) *K is a sequentially closed and lower semicontinuous set-valued function.*

Then, the results in Theorem 3.1 are true for constrained problems.

Proof. Assumptions (H6), (H7) and the Berge theorem guarantee that the set $M_2(x)$ is sequentially closed and so, by assumptions (H1) and (H5), the problem (S_α) has a solution.

To obtain the result, it is sufficient to proceed as in Theorem 3.1. □

4 The Case of Two Followers

For simplicity, we limit ourselves to the case of two followers playing a simultaneous move game, but the results can be easily extended to the case of more followers.

Assume that X is a sequentially compact subset of a topological space and Y_i, for $i = 1, 2$, is a compact, convex and nonempty subset of a finite-dimensional euclidean space V_i. Let $Y = Y_1 \times Y_2$. Let l, f_1, f_2 be functions from $X \times Y$ to $\mathbb{R} \cup \{+\infty\}$, $\alpha \in]0, +\infty[$ and $N(x)$ be the set of parametric Nash equilibria [20] for the game in normal form $\Gamma(x) = (Y_1, Y_2, f_1(x, \cdot, \cdot), f_2(x, \cdot, \cdot))$.

Concerning sufficient conditions for the nonemptyness of $N(x)$, see [20,3,7]. For stability results see, for example, [4,14,17,18].

As in the case of one follower, we are interested in an SPE, that is, a pair (x_0, y_0) where $x_0 \in X$ and $y_0 : X \to Y = Y_1 \times Y_2$ satisfy the conditions

(SPE1) $\quad y_0(x) \in N(x) \quad \forall x \in X$

(SPE2) $\quad l(x_0, y_0(x)) = \inf_{x \in X} l(x, y_0(x))$.

We now consider the following regularized normal form game:

$$\Gamma_\alpha(x) = (Y_1, Y_2, J_1^\alpha(x, \cdot, \cdot), J_2^\alpha(x, \cdot, \cdot)),$$

where $J_i^\alpha(x, y_1, y_2) = J_i(x, y_1, y_2) + \alpha||y_i||^2$ for $i = 1, 2$.

When the parametric nonzero sum game $\Gamma_\alpha(x)$ has a unique solution, denoted by $\bar{y}_\alpha(x) = (\bar{y}_{1,\alpha}(x), \bar{y}_{2,\alpha}(x))$, a solution to the regularized two-level problem

(SN_α): $\text{Min}_{x \in X}\, l(x, \bar{y}_\alpha(x))$

is denoted by $\bar{x}_\alpha$ (as in the case of one follower).

Unfortunately, the Tikhonov regularization does not always guarantee uniqueness of the Nash equilibrium for the regularized problem, as shown by the following example.

Example 4.1. The method that we use does not work in the case in which we have (one leader and) two followers who play a strategic game. The reason for this failure is that Tikhonov regularization does not guarantee uniqueness of the Nash equilibrium for the regularized problem. Let's consider the following game:

	L	R
T	$-1, -1$	$0, 0$
B	$0, 0$	$-1, -1$

Assume that the followers play its mixed extension (for consistency, we assume that players minimize their payoff). The payoff for player I is as follows: $f_2^I(p, q) = -[pq + (1-p)(1-q)]$. Tikhonov regularization for player I amounts to considering the payoff $f_{2\alpha}^I(p, q) = f_2^I(p, q) + \alpha p^2$. An analogous result applies for player II. It is easy to check that both the given game and its regularization have three Nash equilibria in mixed strategies: $(p_1, q_1) = (1, 1)$; $(p_2, q_2) = (0, 0)$; $(p_3, q_3) = ([(1+\alpha)/2(1-\alpha^2)], [(1+\alpha)/2(1-\alpha^2)])$. Thus, not only does Tikhonov regularization not guarantee uniqueness of the Nash equilibrium, but the two Nash equilibria in pure strategies are not affected in any way. Building on this, one easily gets a Stackelberg problem without solution: since the Tikhonov regularization method does not eliminate the uniqueness problem at the followers' level, the regularized problem can still be without solution.

Nevertheless, when the functions f_i are continuously differentiable with respect to y_i we can obtain, under sufficiently mild assumptions, uniqueness of the regularized Nash equilibrium and we can extend the results of Section 3 to the case of more followers.

We assume in the following:

(A1) l is sequentially lower semicontinuous on $X \times Y_1 \times Y_2$.
(A2) $l(x, \cdot, \cdot)$ is upper semicontinuous on $Y_1 \times Y_2$, for all $x \in X$.
(A3)

- for all $(x, y_1, y_2) \in Y_1 \times Y_2$ and all sequences $(x_n, y_{2,n})$ in $X \times Y_2$ convergent to (x, y_2), there exists $(\tilde{y}_{1,n})_n$ in Y_1 such that

$$\limsup_{n\to\infty} f_1(x_n, \tilde{y}_{1,n}, y_{2,n}) \leq f_1(x, y_1, y_2).$$

- for all $(x, y_1, y_2) \in X \times Y_1 \times Y_2$ and all sequences $(x_n, y_{1,n})$ in $X \times Y_1$ convergent to (x, y_1), there exists $(\tilde{y}_{2,n})_n$ in Y_2 such that

$$\limsup_{n\to\infty} f_2(x_n, y_{1,n}, \tilde{y}_{2,n}) \leq f_2(x, y_1, y_2).$$

(A4) f_1 and f_2 are sequentially lower semicontinuous on $X \times Y_1 \times Y_2$.
(A5)

- for all $(x, y_2) \in X \times Y_2$ the function $f_1(x, \cdot, y_2)$ is convex and $f_1(x, \cdot, \cdot)$ is continuously differentiable with respect to y_1 on $Y_1 \times Y_2$,
- for all $(x, y_1) \in X \times Y_1$ the function $f_2(x, y_1, \cdot)$ is convex and $f_2(x, \cdot, \cdot)$ is continuously differentiable with respect to y_2 on $Y_1 \times Y_2$.

(A6) For all $(x, y, z) \in X \times Y \times Y$ we have

$$\left\langle \frac{\partial f_1}{\partial y_1}(x, y_1, y_2) - \frac{\partial f_1}{\partial y_1}(x, z_1, z_2), y_1 - z_1 \right\rangle_1 + \left\langle \frac{\partial f_2}{\partial y_2}(x, y_1, y_2) - \frac{\partial f_2}{\partial y_2}(x, z_1, z_2), y_2 - z_2 \right\rangle_2 \geq 0,$$

where $\langle \cdot, \cdot \rangle_i$ denotes the scalar product on V_i, for $i = 1, 2$.

Remark 4.1. Assumption (A6) is weaker than the "diagonally strict convexity" condition introduced by Rosen [21] which guarantees the uniqueness of the Nash equilibrium.

It may be easily checked that the payoffs for the mixed extension of the game used in the last example do not satisfy condition (A6).

Theorem 4.1. *Assume that (A3) through (A6) are satisfied and $\alpha_n \downarrow 0^+$. Then for all $n \in \mathbb{N}$ and all $x \in X$, $\Gamma_{\alpha_n}(x)$ has a unique Nash equilibrium, denoted by $y_{\alpha_n}(x) = (y_{1,\alpha_n}(x), y_{2,\alpha_n}(x))$ and $y_{\alpha_n} \to \hat{y}(x)$ for $\alpha_n \downarrow 0^+$, where $\hat{y}(x) = (\hat{y}_1(x), \hat{y}_2(x))$ is the minimum norm Nash equilibrium in $N(x)$, with*

$$||(y_1, y_2)||_Y^2 = ||y_1||_{Y_1}^2 + ||y_2||_{Y_2}^2.$$

Proof. By (A3), (A4) and (A5), there exists at least a Nash equilibrium, for all $x \in X$ ([3] or [10], Theorem 2.1).

Moreover, due to (A6), the game $\Gamma_\alpha(x)$ is diagonally strictly convex [21] and we can deduce that there exists a unique Nash equilibrium.

In fact, let $h_\alpha(x,y) = ((\partial f_1^\alpha/\partial y_1)(x,y_1,y_2), (\partial f_2^\alpha/\partial y_2)(x,y_1,y_2))$. Then,

$$\begin{aligned}
&\langle h_\alpha(x,z), y-z\rangle_Y + \langle h_\alpha(x,y), z-y\rangle_Y \\
&\quad = \left\langle \frac{\partial f_1}{\partial y_1}(x,z_1,z_2), y_1 - z_1\rangle_{Y_1} + 2\alpha\langle z_1, y_1 - z_1 \right\rangle_{Y_1} \\
&\qquad + \left\langle \frac{\partial f_2}{\partial y_2}(x,z_1,z_2), y_2 - z_2\rangle_{Y_2} + 2\alpha\langle z_2, y_2 - z_2 \right\rangle_{Y_2} \\
&\qquad + \left\langle \frac{\partial f_1}{\partial y_1}(x,y_1,y_2), z_1 - y_1\rangle_{Y_1} + 2\alpha\langle y_1, z_1 - y_1 \right\rangle_{Y_1} \\
&\qquad + \left\langle \frac{\partial f_2}{\partial y_2}(x,y_1,y_2), z_2 - y_2\rangle_{Y_2} + 2\alpha\langle y_2, z_2 - y_2 \right\rangle_{Y_2} \\
&\quad = \left\langle \frac{\partial f_1}{\partial y_1}(x,z_1,z_2) - \frac{\partial f_1}{\partial y_1}(x,y_1,y_2), y_1 - z_1 \right\rangle_1 \\
&\qquad + \left\langle \frac{\partial f_2}{\partial y_2}(x,z_1,z_2) - \frac{\partial f_2}{\partial y_2}(x,y_1,y_2), y_2 - z_2 \right\rangle_2 \\
&\qquad - 2\alpha||y_1 - z_1||^2 - 2\alpha||y_2 - z_2||^2 < 0
\end{aligned}$$

for all $y \in Y_1 \times Y_2$ and $z \in Y_1 \times Y_2$ such that $y \neq z$.

Now for all $x \in X$ let $y_{\alpha_n}(x)$ be the unique parametric Nash equilibrium of the game $\Gamma_{\alpha_n}(x)$. Due to the assumption (A5), $y_{\alpha_n}(x)$ is the unique solution of the variational inequality $(\mathrm{VI})_{\alpha_n}(x)$ defined by the operator $A_{\alpha_n}(x)$ which associates to every $y = (y_1, y_2) \in Y_1 \times Y_2$ the element $A_{\alpha_n}(x)y$ such that

$$\begin{aligned}
\langle A_{\alpha_n}(x)y, z\rangle_Y = &\left\langle \frac{\partial f_1}{\partial y_1}(x,y_1,y_2), z_1 \right\rangle_{V_1} + 2\alpha_n\langle y_1, z_1\rangle_{V_1} \\
&+ \left\langle \frac{\partial f_2}{\partial y_2}(x,y_1,y_2), z_2 \right\rangle_{V_2} + 2\alpha_n\langle y_2, z_2\rangle_{V_2}.
\end{aligned}$$

In light of Theorem C, page 574, in Mosco [19], for all x, $y_{\alpha_n}(x)$ converges to the solution $\hat{y}(x) \in N(x)$ of the variational inequality

$$\langle \hat{y}(x), z - \hat{y}(x)\rangle \geq 0 \quad \forall\, z \in N(x).$$

that is, $\hat{y}(x)$ is the element of minimal norm of $N(x)$ in $Y_1 \times Y_2$. □

Theorem 4.2. *Assume that (A1) through (A6) are satisfied and $\alpha_n \downarrow 0^+$.*

Let $\bar{x}_n$ be a solution to (SN_{α_n}) and $\bar{y}_n$ be the Nash equilibrium $\bar{y}_{\alpha_n}(\bar{x}_n)$ of the game $\Gamma_{\alpha_n}(\bar{x}_n) = (Y_1, Y_2, f_1^{\alpha_n}(\bar{x}_n,\cdot,\cdot), f_2^{\alpha_n}(\bar{x}_n,\cdot,\cdot))$. If $(\bar{x}_n, \bar{y}_n) \to (\bar{x}, \bar{y}) \in$

$X \times Y_1 \times Y_2$, then $(\bar{x}, \bar{y})$ generates an SPE $(\bar{x}, \tilde{y}(\cdot))$ where $\tilde{y}(\cdot)$ is defined by

$$\begin{aligned} &\tilde{y}(\bar{x}) = \bar{y} \\ &\tilde{y}(x) = \hat{y}(x) \quad \forall x \in X, x \neq \bar{x}, \end{aligned}$$

where $\hat{y}(x)$ is the element of minimum norm in $N(x)$.

Proof. Assume that $(\bar{x}_n, \bar{y}_n) \to (\bar{x}, \bar{y})$.

Due to assumptions (A3) and (A4), the set-valued function N is sequentially closed at $\bar{x}$ (Theorem 3.1 in [10]).

We can now easily conclude as in Theorem 3.1 of Section 3. □

5 Conclusion

Let us conclude with some remarks about the limits of our approach. The regularizations that we propose make an essential use of some convexity assumptions. For this reason, we expect that they cannot be extended in a straightforward way to cases like the three-level problems (incorporating the best reply into the functionals will destroy, in general, their convexity properties).

A similar remark also applies to the case in which we have one leader and two followers, the latter playing a simultaneous move game. In this case we have provided a positive result along the same lines used for the classical bilevel optimization problem. On the other hand, we have shown, through an example, that the method cannot be extended to all the cases where the followers play a mixed extension of a finite game.

REFERENCES

[1] Basar T. and Olsder G.J., *Dynamic Noncooperative Game Theory*, Academic Press, New York, 1982.

[2] Breton M., Alj A. and Haurie A., Sequential Stackelberg equilibrium in two-person games, *J. Optimization Theory and Applications*, 59, 71–97, 1988.

[3] Cavazzuti E., An extension of Nash's Theorem, Preprint, Dipartimento di Matematica Pura ed Applicata, Università di Modena, Italy, 1996.

[4] Cavazzuti E. and Pacchiarotti N., Convergence of Nash equilibria, *Boll. UMI*, 5B, 247–266, 1986.

[5] Dempe S., On an algorithm solving two-level programming problems with nonunique lower level solutions, *Comp. Opti. Appl.*, 6, 227–249, 1996.

[6] Leitmann G., On generalized Stackelberg strategies, *J. Optimization Theory and Applications*, 26, 637–643, 1978.

[7] Lignola M.B., Ky fan inequalities and Nash equilibrium points without semicontinuity and compactness, *J. Optimization Theory and Applications*, 94, 137–145, 1997.

[8] Lignola M.B. and Morgan J., Topological existence and stability for Stackelberg problems, *J. Optimization Theory and Applications*, 84, 145–169, 1995.

[9] Lignola M.B. and Morgan J., Stability for regularized bilevel programming problem, *J. Optimization Theory and Applications*, 93, 575–596, 1997.

[10] Lignola M.B. and Morgan J., Existence for optimization problems with equilibrium constraints in reflexive Banach spaces, *Optimization in economics, finance and industry* (Verona 2001), Datanova, Milano, 15–35, 2002.

[11] Loridan P. and Morgan J., New results on approximate solutions in two-level optimization, *Optimization*, 20, 819–836, 1989.

[12] Loridan P. and Morgan J., Regularizations for two-level optimization problems, *Lecture Notes in Economic and Math. Systems*, 382, 239–255, 1992.

[13] Loridan P. and Morgan J., Weak via strong Stackelberg problem: new results, *J. Global Optimization*, 8, 263–287, 1996.

[14] Lucchetti R. and Patrone F., Closure and upper semicontinuity results in mathematical programming, Nash and economic equilibria, *Optimization*, 17, 619–628, 1986.

[15] Molodtsov D.A., The solution of a class of non antagonistic games, *USSR Comput. Maths. Math. Phys.*, 16, 1451–1456, 1976.

[16] Morgan J., Constrained well-posed two-level optimization problems, in F. Clarke, V. Demianov, and F. Giannessi (eds.), *Nonsmooth Optimization and Related Topics*, Ettore Majorana International Sciences Series, Plenum Press, New York, 307–326, 1989.

[17] Morgan J. and Raucci R., New convergence results for Nash equilibria, *J. Convex Analysis*, 6, 377–385, 1999.

[18] Morgan J. and Scalzo V., Pseudocontinuity in optimization and non-zero sum games, *J. Optimization Theory and Applications*, 120, 181–197, 2004.

[19] Mosco U., Convergence of convex sets and solutions of variational inequalities, *Adv. Math.*, 3, 510–585, 1969.

[20] Nash J.F. Jr., Equilibrium points in n-person games, *Proc. Nat. Acad. Sc. USA*, 36, 48–49, 1950.

[21] Rosen J.B., Existence and uniqueness of equilibrium points for convex n-person games, *Econometrica*, 33, 520–534, 1965.

[22] Selten R., Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit. Teil I: Bestimmung des dynamischen Preisgleichgewichts; Teil II: Eigenschaften des dynamischen Preisgleichgewichts, *Zeitschrift für die gesamte Staatswissenschaft*, 121, 301–324 and 667–689, 1965.

[23] Soholovic V.F., Unstable extremal problem and geometric properties of Banach spaces, *Dokl.*, 11, 1470–1472, 1970.

Extended Self, Game, and Conflict Resolution*

Noboru Hidano
Tokyo Institute of Technology
Department of Social Engineering
2-12-1, Ookayama, Meguroku, Tokyo, Japan
hidano@wb3.so-net.ne.jp

Shigeo Muto
Tokyo Institute of Technology
Department of Social Engineering
2-12-1, Ookayama, Meguroku, Tokyo, Japan
muto@valdes.titch.ac.jp

Abstract

The purpose of this chapter is twofold: (1) to bring a new concept of extended self in philosophy into the analysis of conflict resolution, and (2) to construct a game theoretic model with selves and extended selves as players which depicts a conflicting situation and to find its resolution. We find that extended selves could be a useful concept to unite players' opposing ideas. We study a two-stage noncooperative game in which players are given by selves or extended selves and find that, for each player, to unite with another player is always a weakly dominant strategy when payoffs are symmetric and costs for unification are negligible.

1 Introduction

Despite the fact that the situation in which different players may unite during a game is discussed at length in game theory literature as a coalition formation problem, few papers analyze situations in which utility functions of players may change through the formation of coalitions. Of course, there is a school of thought that allows for adjustment of individual utility, but not a utility function, through an altruistic mechanism [1]. But our argument is different from the argument of this school. We are interested in changes of a player himself during the play of a game; and our focus is on a game in which players come

*This research was partially supported by the Japan society for the Promotion of Science through the Grants in Aid for Scientific Research (#16310107,2004). The authors wish to acknowledge an anonymous referee for valuable comments. They also wish to thank Mr. Ryo Kawasaki for editing the English.

to share the same utility function when they form a coalition. In an ordinal game theoretic formulation, every player retains an unchanged utility function during the game because such a formulation assumes the basic unit of a player is unchanged. This formulation is deeply rooted in the fundamental assumption of game theory, i.e., methodological individualism.

But is it reasonable to construct a model on this assumption? We can list several examples that may violate this assumption. In the case of a merger of firms, the behavior of the united firm is different from the behavior of the firms before the merger. The reason for this is not only due to quantitative factors such as changes of scale, but also due to changes in the firm's utility function. The culture of the new merged firm is a mixture of those of the previous firms. In international politics, we can see other examples, such as the unification of European Union (EU) countries. The behavior of countries belonging to the EU before the establishment of a common currency and the behavior after the "euro" seem to be different. This is partly because the introduction of a common currency changes the way of thinking of European citizens. As such, forming a coalition will alter the utility function of players. We should reconsider our assumptions about a player in a game when we suspect that utility functions of the players who formed a coalition become more homogeneous.

Thus, the purpose of this paper is twofold: (1) to discuss the concept of a unit of decision making that allows one individual to hold different perspectives in a game, and (2) to formulate a model that incorporates a new idea of self for an individual player, more precisely the model in which

each player alters his or her utility function during the game if he or she unites in purpose with other players.

2 Concepts of Self

In the Western world, the self-concept has been discussed extensively as one of the most important individual philosophical concepts. But in modern economics and other social sciences, self is usually defined in the abstract as an individual who can make decisions independently and whose body is his/her own possession. In the non-Western world, however, there exist different views of the idea of self. A notable example is the philosophy of Chu Hsi, which is considered the most sophisticated and rational school of Confucianism.

In the English school of experientialism, Parfit [10] discussed the self-concept comprehensively from a nonreligious viewpoint. He introduced the idea of whether self-identity could be determined rationally, introducing new criteria: psychological connectedness and physical similarity. He concluded that individuals could not be distinguished from other persons using these criteria and that continuity from one person to another should be the rule. He applied this concept to all human beings. But it is questionable whether we can really consider all human beings as sharing identities with all others since people tend to be self-centered. In ecological philosophy, Naess and Rothenberg [9] advocated the idea of an ecological self in which personal identity should expand to include not only other human beings but also other living creatures. Naess gave the example of a flea, saying that he felt pain when he saw a flea lying in agony in acid. He even thinks that his ecological self can be expanded to nonliving creatures. But again, it is very difficult to distinguish which creatures we should consider as our ecological selves. Iwata [6], a cultural anthropologist, has studied for many years the lives of indigenous tribes in Southeast Asian countries, and Iwata describes the "kami" concept, i.e., a spirit, existing in living creatures and some nonliving creatures Ito [7]. Iwata argues that this concept is fundamental to any religion or religious experience. His idea of self is, then, the coexistence of all beings in a world of synchronism. There may be many arguments and disagreements, both from religious and nonreligious viewpoints, about his idea. But it should be noted that value conflicts cannot be resolved without paying attention to people's views towards the environment and self in religious and tribal society as well. Chu Hsi's ideology argued that the world is based on the activities of Heaven and Earth by the force of ch'i (vital energy). Kuwako [8, p. 153] recently criticized the animistic aspect of ch'i and emphasized the material side rather than the vital side, purporting that Chu Hsi is a more materialistic idea than is often appreciated. The concept of self in Chu Hsi is that after a human body is shaped, the functions of the mind begin to work as the result of reaction to ch'i outside of the body and to the ch'i of the body. These functions stop after death. Since the body itself is created by ch'i, the self is regarded as the continuity of all other existence. Environment and self are considered identical. Iwata and Chu Hsi are two representative views in East and Southeast Asian cultures, but we think they pay little attention to the psychological aspects of human beings in the real world.

3 Extended Self and Unification of Extended Selves

Based on the discussions in the previous section, we propose and define the extended self as a set of objects (basically human beings) which produce an individual physically and are considered psychologically by the individual as his or her parents, as well as objects (human and otherwise) which the individual has and those he will produce, objects with which a person can feel psychological identity after death (Hidano [3]). More precisely, we extend the concept of self in two directions. The first is an extension related to time, namely past and future beings, and the second is an extension related to the objects of identity to nonhuman beings. The first extension is that we may add the concept of "oneself" to one's parents, grandparents, and other ancestors who produced an individual physically, and we also may include children, grandchildren, and other descendants whom one has produced and is able to produce, as a part of oneself. The latter extension also includes nonhuman beings. As far as we can identify ourselves with any beings or existence, here we include the objects of an extended self by which we have been produced and which we have produced ourselves, and which may exist after our death. Even inanimate objects such as art, novels, and information can be candidates for objects of the extended self. Thus, also included are other subjects, which are considered psychologically as our parental subjects. The idea of considering other living beings and even nonliving objects and creatures as beings with spirits is based on animism, which has recently been discussed by Iwata [6]. It should be noted that a person or an object, which is considered as a part of the extended self, is chosen and determined by the self. Thus, even a parent or child cannot be included as a part of an individual's extended self without the approval of the individual. These extensions of the self-concept enable us to reconsider our value structures related to the individual self, and to drastically lessen internal value conflicts by thinking that in the future we (our extended self) will most probably confront situations which we usually can ignore within our individual lives. One example is this: In our children, grandchildren, and other descendants, we will exist as an extended self. Thus, they do not have an existence isolated from our own. We should automatically consider their existence and decrease the risk of affecting them adversely. Our motivation is not based on altruism but on our own self-centeredness. This idea of self-concept can lessen value conflicts between people at this moment as well. For example, consider that Person "1" is confronting value conflicts he or she has with Person "2". In the future the descendent of Person "1" (Person "n") may become a descendent of Person "2" as well. Person "1" coexists with Person "2" in Person "n". The value conflict between Persons "1" and "2" becomes an internal value conflict in one person specifically, Person "n". Person "1" and Person "2" should at least recognize that they exist in one person. This expectation of the future inevitably changes the behavior of two persons at this moment. It should be noted that this concept is easily accepted as a conceptual vehicle by those who do not agree with the direct expansion of their identity towards Naess's ecological self. This idea of the extended self can be applicable to all agents such as organizations (firms, nations, nonprofit organizations, etc.) and even to the environment (forests, animals, landscape, etc.).

4 Game Theoretic Formulations of Conflicts among Extended Selves

4.1 Game Theoretic Model

We will study under what conditions two or more different selves unite. We suppose that they have a common utility function when they unite, which may differ from their original utility functions.

As an introductory step, we consider a case with two selves. Thus, we assume a two-person game consisting of players 1 and 2. Selves are called players in the following. Let S and T be sets of strategies of players 1 and 2, respectively. If they unite, they have a strategy set R given by $S \times T$, the direct product of S and T. Players 1 and 2 have utility functions

$$f : S \times T \to \Re \text{ and } g : S \times T \to \Re,$$

respectively, where $\Re$ is the set of all real numbers. It is supposed that if they unite, they have an identical utility function $h : R = S \times T \to \Re$ that may differ from f and g. Further, when they unite, each player incurs a cost ε.

The game proceeds as follows. First players 1 and 2 simultaneously and independently decide whether "to unite" or "not to unite" denoted by "U" and "N", respectively. Then unless both choose U, they play the strategic form game ({1,2}, $\{S,T\}$, $\{f,g\}$) where players 1 and 2 simultaneously and independently choose their strategies in S and T, respectively. Their aims are to maximize their own utility levels. If both choose U, they unite and take a strategy pair that maximizes h over R. Both players enjoy the utility given by $\max\{h(s,t)|(s,t) \in R\}$.

Our aim is to study subgame perfect equilibria of the game in order to make clear under what conditions different selves unite.

To simplify the discussion, we consider the case of symmetric 2×2 games, i.e., the players' strategy sets S and T are given by $S = \{s_1, s_2\}, T = \{t_1, t_2\}$ and the payoff matrix is

1/2	t_1	t_2
s_1	a, a	b, c
s_2	c, b	d, d

where we assume for simplicity that a, b, c, d are all different and we let $a > d$ without loss of generality. We further suppose that the utility function h is given by a convex combination of f and g, i.e.,

$$h(s,t) = \alpha f(s,t) + (1-\alpha)g(s,t) \quad \forall\ (s,t) \in R = S \times T,$$

where $0 \leq \alpha \leq 1$. The parameter α uniformly distributes over the interval [0,1].

In what follows we will consider two cases with respect to information that players have on the parameter α, i.e., (1) both players know α before they unite, and (2) neither player knows α before they unite. Of course, the united player knows α. Before studying subgame perfect equilibria in the two cases, we examine what will come out in the second stage.

4.2 Equilibria in the Second Stage

We have two cases. Unless both choose "U" (to unite), two players do not unite and play the noncooperative game; if both choose "U", they unite and choose the strategy pair that maximizes the utility function h. Reflecting the fact that players are symmetric with respect to payoffs, we pick symmetric

Nash equilibria. If there exist multiple symmetric Nash equilibria, we choose the Pareto efficient one.

Symmetric, Pareto efficient Nash equilibria and their payoffs are given in the following table.

		Symmetric, Pareto efficient Nash equilibria	Payoffs
1.	$a > b > c > d$	(s_1, t_1)	a
2.	$a > b > d > c$	(s_1, t_1)	a
3.	$a > c > b > d$	(s_1, t_1)	a
4.	$a > c > d > b$	(s_1, t_1)	a
5.	$a > d > b > c$	(s_1, t_1)	a
6.	$a > d > c > b$	(s_1, t_1)	a
7.	$b > a > c > d$	(s_1, t_1)	a
8.	$b > a > d > c$	(s_1, t_1)	a
9.	$b > c > a > d$	$((p, 1-p), (p, 1-p))$	$ap^2 + (b+c)p(1-p) + d(1-p)^2$
10.	$c > a > b > d$	$((p, 1-p), (p, 1-p))$	$ap^2 + (b+c)p(1-p) + d(1-p)^2$
11.	$c > a > d > b$	(s_2, t_2)	d
12.	$c > b > a > d$	$((p, 1-p), (p, 1-p))$	$ap^2 + (b+c)p(1-p) + d(1-p)^2$

where $p = (b-d)/(c-a+b-d)$.

Strategy pairs maximizing h and their payoffs are given as follows.

		Possible strategy pairs maximizing h	Payoffs
1.	$a > b > c > d$	(s_1, t_1)	a
2.	$a > b > d > c$	(s_1, t_1)	a
3.	$a > c > b > d$	(s_1, t_1)	a
4.	$a > c > d > b$	(s_1, t_1)	a
5.	$a > d > b > c$	(s_1, t_1)	a
6.	$a > d > c > b$	(s_1, t_1)	a
7.	$b > a > c > d$	(s_1, t_1), (s_1, t_2), (s_2, t_1)	a, $\alpha b + (1-\alpha)c$, $\alpha c + (1-\alpha)b$
8.	$b > a > d > c$	(s_1, t_1), (s_1, t_2), (s_2, t_1)	a, $\alpha b + (1-\alpha)c$, $\alpha c + (1-\alpha)b$
9.	$b > c > a > d$	(s_1, t_2), (s_2, t_1)	$\alpha b + (1-\alpha)c$, $\alpha c + (1-\alpha)b$
10.	$c > a > b > d$	(s_1, t_1), (s_1, t_2), (s_2, t_1)	a, $\alpha b + (1-\alpha)c$, $\alpha c + (1-\alpha)b$
11.	$c > a > d > b$	(s_1, t_1), (s_1, t_2), (s_2, t_1)	a, $\alpha b + (1-\alpha)c$, $\alpha c + (1-\alpha)b$
12.	$c > b > a > d$	(s_1, t_2), (s_2, t_1)	$\alpha b + (1-\alpha)c$, $\alpha c + (1-\alpha)b$

In cases 7–12, the maximum payoff is attained by one of these strategy pairs depending on values of α.

We next find subgame perfect equilibria for the whole game.

4.3 Subgame Perfect Equilibria

As mentioned before, the following two cases are examined, i.e., (1) both players know α before they unite, and (2) neither player knows α before they unite.

I. Both players know α before they unite

In this case, when players decide whether to unite or not, they know which strategy pair the united player will take.

Thus, in cases 7 and 8 above, if $a \geq (b+c)/2$ (and thus $b-a \leq a-c$), the united player will take (s_2, t_1) when $0 \leq \alpha \leq (b-a)/(b-c)$, (s_1, t_1) when $(b-a)/(b-c) \leq \alpha \leq (a-c)/(b-c)$, and (s_1, t_2) when $(a-c)/(b-c) \leq \alpha \leq 1$; because the maximum payoffs $\alpha c + (1-\alpha)b$, a, and $\alpha b + (1-\alpha)c$ are attained by these strategy pairs. If $a < (b+c)/2$ (and thus $a-c < b-a$), then the united player will take (s_2, t_1) when $0 \leq \alpha \leq 1/2$ and (s_1, t_2) when $1/2 \leq \alpha \leq 1$. The maximum payoffs are $\alpha c + (1-\alpha)b$ and $\alpha b + (1-\alpha)c$.

In cases 9 and 12, the united player will take (s_2, t_1) when $0 \leq \alpha \leq 1/2$ and (s_1, t_2) when $1/2 \leq \alpha \leq 1$. The maximum payoffs are $\alpha c + (1-\alpha)b$ and $\alpha b + (1-\alpha)c$.

In cases 10 and 11, if $a \geq (b+c)/2$ (and thus $c-a \leq a-b$), the united player will take (s_1, t_2) when $0 \leq \alpha \leq (c-a)/(c-b)$, (s_1, t_1) when $(c-a)/(c-b) \leq \alpha \leq (a-b)/(c-b)$, and (s_2, t_1) when $(a-b)/(c-b) \leq \alpha \leq 1$. The maximum payoffs are $\alpha b + (1-\alpha)c$, a, and $\alpha c + (1-\alpha)b$, respectively. If $a < (b+c)/2$ (and thus $c-a < a-b$), then the united player will take (s_1, t_2) when $0 \leq \alpha \leq 1/2$ and (s_2, t_1) when $1/2 \leq \alpha \leq 1$. The maximum payoffs are $\alpha b + (1-\alpha)c$ and $\alpha c + (1-\alpha)b$.

Comparing these maximum payoffs with Nash equilibrium payoffs, we obtain the following subgame perfect equilibrium outcomes.

(1) Cases 1–6: Payoffs are the same in the two situations. Thus, taking into account the cost ε of unification, players do not unite in equilibrium. Players take s_1 and t_1 independently and each player gains payoff a.

(2) Cases 7 and 8: Suppose $a \geq (b+c)/2$ (and thus $b-a \leq a-c$). Then the maximal payoffs when players unite are given by $\alpha c + (1-\alpha)b$, a, and $\alpha b + (1-\alpha)c$ when $0 \leq \alpha < (b-a)/(b-c)$, $(b-a)/(b-c) \leq \alpha < (a-c)/(b-c)$, and $(a-c)/(b-c) \leq 1$, respectively. Thus, when

$$\begin{aligned} &\alpha c + (1-\alpha)b - \varepsilon > a, \text{ i.e., } \alpha < (b-a-\varepsilon)/(b-c) \\ \text{or} \quad &\alpha b + (1-\alpha)c - \varepsilon > a, \text{ i.e., } \alpha > (a-c+\varepsilon)/(b-c), \end{aligned}$$

players unite, and taking (s_2, t_1) they gain $\alpha c + (1-\alpha)b - \varepsilon$ (in the former case) and taking (s_1, t_2) they gain $\alpha b + (1-\alpha)c - \varepsilon$ (in the latter case).

When

$$(b-a-\varepsilon)/(b-c) < \alpha < (a-c+\varepsilon)/(b-c),$$

players do not unite and behave independently; each of them gains a taking strategies s_1 and t_1. If $\alpha = (a-c-\varepsilon)/(b-c)$ or $\alpha = (b-a-\varepsilon)/(b-c)$, then players may unite or may not unite.

In the case of $a < (b+c)/2$, similar results hold.

(3) Cases 9 and 12: In these cases, a simple calculation shows that the expected payoff to each player in the Nash equilibrium is $(bc-ad)/(c-a+b-d)$.

Therefore, if

$$(b+c)/2-\varepsilon > (bc-ad)/(c-a+b-d)$$

or

$$\varepsilon < ((b-a)(b-d)+(c-a)(c-d))/2(c-a+b-d),$$

two players unite.

Suppose

$$\varepsilon \geq ((b-a)(b-d)+(c-a)(c-d))/2(c-a+b-d).$$

Then in case 9, when

$$\begin{aligned} &\alpha b+(1-\alpha)c-\varepsilon > (bc-ad)/(c-a+b-d), \text{ i.e.,} \\ &\qquad \alpha > (d-c)(c-a)/(c-a+b-d)(b-c)+\varepsilon/(b-c), \\ \text{or}\quad &\alpha c+(1-\alpha)b-\varepsilon > (bc-ad)/(c-a+b-d), \text{ i.e.,} \\ &\qquad \alpha < (b-a)(b-d)/(c-a+b-d)(b-c)-\varepsilon/(b-c), \end{aligned}$$

they unite; and when

$$\begin{aligned} &(b-a)(b-d)/(c-a+b-d)(b-c)-\varepsilon/(b-c) < \alpha \\ &\quad < (d-c)(c-a)/(c-a+b-d)(b-c)+\varepsilon/(b-c), \end{aligned}$$

they do not unite; and when

$$\begin{aligned} &\alpha = (d-c)(c-a)/(c-a+b-d)(b-c)+\varepsilon/(b-c) \\ \text{or}\quad &\alpha = (b-a)(b-d)/(c-a+b-d)(b-c)-\varepsilon/(b-c), \end{aligned}$$

they may or may not unite.

In case 12, when

$$\begin{aligned} &\alpha < (c-d)(c-a)/(c-a+b-d)(c-b)-\varepsilon/(c-b) \\ \text{or}\quad &\alpha > -(b-a)(b-d)/(c-a+b-d)(c-b)+\varepsilon/(c-b), \end{aligned}$$

they unite.

If the reverse inequalities hold in the cases above, players do not unite and behave independently; and if the equalities hold, players may unite or may not unite.

(4) Case 10: Suppose $a \geq (b+c)/2$. Then if

$$a-\varepsilon > (bc-ad)/(c-a+b-d), \text{ i.e., } \varepsilon < (c-a)(a-b)/(c-a+b-d),$$

the players unite.

Suppose

$$\varepsilon \geq (c-a)(a-b)/(c-a+b-d).$$

Then if

$$\begin{aligned} &\alpha < (c-d)(c-a)/(c-a+b-d)(c-b) - \varepsilon/(c-b) \\ \text{or}\quad &\alpha > (a-b)(b-d)/(c-a+b-d)(c-b) + \varepsilon/(c-b), \end{aligned}$$

they unite.

If the reverse inequalities hold in the cases above, players do not unite and behave independently; and if the equalities hold, players may or may not unite.

In the case of $a < (b+c)/2$, similar results hold.

(5) Case 11: If $a-\varepsilon > d$, i.e., $\varepsilon < a-d$, then players unite. Suppose $\varepsilon \geq a-d$. Then if

$$\begin{aligned} &\alpha b + (1-\alpha)c - \varepsilon > d, \text{ i.e., } \alpha < (c-d-\varepsilon)/(c-b) \\ \text{or}\quad &\alpha c + (1-\alpha)b - \varepsilon > d, \text{ i.e., } \alpha > (d-b+\varepsilon)/(c-b), \end{aligned}$$

they unite.

If the reverse inequalities hold in the cases above, players do not unite and behave independently; and if the equalities hold, players may unite or may not unite.

Now we examine the second case, in which neither player knows α before they unite.

II. Neither player knows α before they unite

(1) Cases 1–6: Exactly the same as in I.

(2) Cases 7 and 8: Suppose $a \geq (b+c)/2$. The maximum payoffs are given by $\alpha c + (1-\alpha)b$, a, and $\alpha b + (1-\alpha)c$ when $0 \leq \alpha < (b-a)/(b-c)$, $(b-a)/(b-c) \leq \alpha < (a-c)/(b-c)$, and $(a-c)/(b-c) \leq 1$, respectively. Thus, the expected maximum payoff is given by

$$\int_0^{b-a/b-c} [\alpha c + (1-\alpha)b]\, d\alpha + \int_{b-c/b-a}^{a-c/b-c} a\, d\alpha + \int_{a-c/b-c}^{1} [\alpha b + (1-\alpha)c]\, d\alpha$$
$$= \frac{b^2 + a^2 - ab - ac}{b-c}.$$

Hence, if $(b^2 + a^2 - ab - ac)/(b-c) - \varepsilon > a$, i.e., $\varepsilon < (b-a)^2/(b-c)$, then two players unite. If the reverse inequality holds, players do not unite; and if the equality holds, players may unite or may not unite.

Suppose next that $a < (b+c)/2$. In this case, the expected maximum payoff is

$$\int_0^{1/2} [\alpha c + (1-\alpha)b]\, d\alpha + \int_{1/2}^{1} [\alpha b + (1-\alpha)c]\, d\alpha = \frac{3b+c}{4}.$$

Hence, if $\varepsilon < (3b + c)/4 - a$, then two players unite. If the reverse inequality holds, players do not unite; and if an equality holds, players may or may not unite.

(3) Cases 9 and 12: The expected maximum payoff is given by $(3b + c)/4$. Hence, if $\varepsilon < (3b+c)/4-(bc-ad)/(c-a+b-d)$, then two players unite. If the reverse inequality holds, players do not unite; and if an equality holds, players may unite or may not unite.

(4) Case 10: Suppose $a \geq (b + c)/2$. Then the expected maximum payoff is $(b^2 + a^2 - ab - ac)/(b - c)$. Hence, if $(b^2 + a^2 - ab - ac)/(b - c) - \varepsilon > (bc-ad)/(c-a+b-d)$, i.e., $\varepsilon < (bc-ad)/(c-a+b-d)-(b^2+a^2-ab-ac)/(b-c)$, then two players unite. If the reverse inequality holds, players do not unite; and if an equality holds, players may or may not unite. Suppose next that $a < (b + c)/2$. The expected maximum payoff is $(3b + c)/4$. Hence, if $\varepsilon < (3b + c)/4 - (bc - ad)/(c - a + b - d)$, then two players unite. If the reverse inequality holds, players do not unite; and if equality holds, players may unite or may not unite.

(5) Case 11: Suppose $a \geq (b + c)/2$. If $(b^2 + a^2 - ab - ac)/(b - c) - \varepsilon > d$, i.e., $\varepsilon < (b^2 + a^2 - ab - ac)/(b - c) - d$, then two players unite. If the reverse inequality holds, players do not unite; and if equality holds, players may unite or may not unite. Suppose $a \leq (b + c)/2$. Then if $\varepsilon < (3b + c)/4 - d$, then two players unite. If the reverse inequality holds, players do not unite; and if equality holds, players may unite or may not unite.

It should be noted that a straightforward calculation shows that in the above inequalities with respect to ε the right-hand side terms are all positive.

4.4 Summary of the Results

The results obtained in the previous subsection are summarized as follows.

If the payoff a is greater than off-diagonal payoffs b and c, then players never have an incentive to unite. This holds whether the weight α is revealed before unification or not.

Suppose a is less than off-diagonal payoffs b or c. If the weight α is revealed in advance, then players are more likely to behave independently when they are symmetric with respect to the weight α (i.e., α is close to 1/2). In other words, players are more likely to unite when they are non-symmetric with respect to α. If the weight is revealed after unification, then players always unite when the cost for unification ε is not very large.

If ε is zero, the unification is a weakly dominant strategy for both players independently from the payoffs.

On the basis of the observation, we may claim that (1) if players' similarity is high, then they are likely to behave independently, (2) if players are symmetric with respect to the united player's utility, then they are likely to behave independently, (3) if the information on the united player's utility is accurate, then they are likely to behave independently, and (4) if the cost of unification is zero, they should always unite.

4.5 Future Extension

The model on which we have worked is rather simple; there are only two players and the game is played only once. We need further extensions of the model. The following are possible ways of extending it.
(1) Introduce three or more players.
(2) Repeat the game finitely or infinitely many times.
(3) Introduce the possibility that coalitions may break up if the game is repeated.
(4) Construct players' utility functions based on socio-economic situations behind the model.
(5) Find a more reasonable way to define a utility function of a united player.

In addition, we should study games with nonsymmetric players to obtain more fruitful results.

5 Conclusions

In this chapter, we have explained how a concept of extended self can assist us in solving current value conflicts, i.e., conflicts arising from different value judgements among people. Most value conflicts seem to be based on very short-sighted conceptions of people, which seldom consider their future. This study tries to discuss possible results when we introduce the concept of extended self into conflict situations. It shows that, in the symmetric payoff case, players will follow a strategy of coalition even though the coalition includes an agent who is currently opposed to them. We think that the extended self or the extended organization should be considered now in the present, not as if it belonged only to our grandchildren or a descendant organization. Considering of extended selves right now can help us greatly in solving value conflicts in such areas as global warming, other ecological issues, and conflicts among nations. We would point out that the European Union is a notable example of a united entity. After a tremendous number of conflicts among nations, they were able to form a union. Thus, we hope that game theory and social sciences will devote much attention to the notion of extended self.

REFERENCES

[1] Brainov S., Altruistic cooperation between self-interested agents, in W. Wahlster (ed.), *ECAI'96 12th European Conference on Artificial Intelligence*, 519–523, John Wiley & Sons, Ltd., Chichester, 1996.

[2] Hidano N., *Regional Design from the Perception of the Environment and the Family*, 2nd edition, Hidano Laboratory, Tokyo Institute of Technology, 1995.

[3] Hidano N., Extended self and value conflict resolution, Paper presented at *Frontiers 1: Fundamental Issues of Ecological Economics*, European Society for Ecological Economics, Cambridge University, 2001.

[4] Hidano N., Extended self, game, and value conflict resolution, Paper presented at the *15th VALDES Game Theory Seminar*, Tokyo Institute of Technology, 2001.

[5] Hidano N. and Muto S., Extended self, game, and conflict resolution in ecological systems, *Proceedings of 7th Biennial Conference of International Society for Ecological Economics*, Tunisia, 2002.

[6] Iwata K., *The Works of Keiji Iwata*, Vol. 1–8, Kodansha, Tokyo, 1995.

[7] Ito M., Evolution of the Concept of Kami, http://www.kokugakuin.ac.jp/ijcc/wp/cpjr/kami/ito.html, 2001.

[8] Kuwako T., The philosophy of environmental correlation in Chu Hsi, in M.E. Tucker and J. Berthrong (eds.), *Confucianism and Ecology*, Harvard University Press, Cambridge, MA, 1998.

[9] Naess A. and Rothenberg D., *Ecology, Community and Lifestyle: Outline of an Ecosophy*, Cambridge University Press, London, 1989.

[10] Parfit D., *Reasons and Persons*, Oxford University Press, London, 1984.

Game of Timing in Gas Pipeline Projects Competition: Simulation Software and Generalized Equilibrium Solutions*

Arkadii Kryazhimskii
Moscow State University, Russia
kryazhim@mi.ras.ru

Oleg Nikonov
Urals State Technical University, Russia
aspr@mail.ustu.ru

Yaroslav Minullin
International Institute for Applied Systems Analysis
Laxenburg, Austria
minullin@iiasa.ac.at

Abstract

Many models of energy market development and decision-making processes take into account the competition between energy suppliers, and the theory of games is an appropriate tool to study these problems.

This chapter is devoted to numerical analysis and modification of the game-theoretical gas market model developed by Klaassen, Kryazhimskii, and Tarasyev. We describe a software G-TIME elaborated for this purpose and the results of a simulation and sensitivity analysis on the data of the Turkish gas market. The last section deals with the notion of a generalized Nash equilibrium, which seems to be useful for taking risk and uncertainty into account. The research is based on approaches and methods developed in [1–10].

1 Model Description

Most models of gas pipeline projects competition deal with investment plans of large natural gas exporters to a specified gas market. Several papers are devoted to the market of Western Europe, and some papers concern the Turkish and

*Partly supported by RFBR Grants N 05-01-08034-ofi_p and N 06-01-00483a, and N 05-02-02118a.

Asian gas markets. A game-theoretical model of competition between two gas pipeline projects called *the game of timing* has been proposed in [10].

The pipelines are expected to operate in the same market. Players 1 and 2 are associated with the investors/managers of projects 1 and 2, respectively. Assuming that the starting time for making investments is 0, we consider the "virtual" positive commercialization times of projects 1 and 2 (i.e., the final times of the construction of the pipelines), t_1 and t_2. Given a commercialization time, t_i, player i $(i = 1, 2)$ can estimate the cost, $C_i(t_i)$, for finalizing project i at time t_i.

At any time $t > 0$, the gas price and cost for extraction, transportation, distribution and transit fees determine a *benefit rate* for each player. When one of the players solely occupies the market he gets *an upper benefit rate*, b_{i1}. When another player enters the market both of them get *a lower benefit rate*, b_{i2}, which is lower than the upper benefit rate since the appearance of another competitor decreases the market price:

$$b_{i1}(t) > b_{i2}(t).$$

We stress the dependence of benefit rates on competitive commercialization time and write

$$b_1(t|t_2) = \begin{cases} b_{11}(t) \text{ if } & t < t_2, \\ b_{12}(t) \text{ if } & t \geq t_2 \end{cases}.$$

Similarly, a commercialization time t_1 of project 1 determines *the benefit rate* of player 2 as

$$b_2(t|t_1) = \begin{cases} b_{21}(t) \text{ if } & t < t_1, \\ b_{22}(t) \text{ if } & t \geq t_1 \end{cases}.$$

The total benefit for each player is determined by the following equalities:

$$B_1(t_1, t_2) = \int_{t1}^{\infty} b_1(t|t_2)dt, B_2(t_1, t_2) = \int_{t2}^{\infty} b_2(t|t_1)dt,$$

and the total profit as

$$P_i(t_1, t_2) = B_i(t_1, t_2) - C_i(t_i).$$

We assume that the functions $b_{ij}(t)$ $(1, j = 1, 2)$ are continuous and monotonically decreasing and the above integrals are finite. Figures 1 and 2 illustrate a typical behavior of the graphs of the functions introduced above.

According to the standard terminology of the game theory, a strategy t_1^* of player 1 is said to be the best response of player 1 to a strategy t_2 of player 2 if t_1 maximizes the payoff to player 1, $P_1(t_1, t_2)$, over the set of all strategies of player 1, t_1:

$$P_1(t_1^*, t_2) = \max_{t_1 > 0} P_1(t_1, t_2).$$

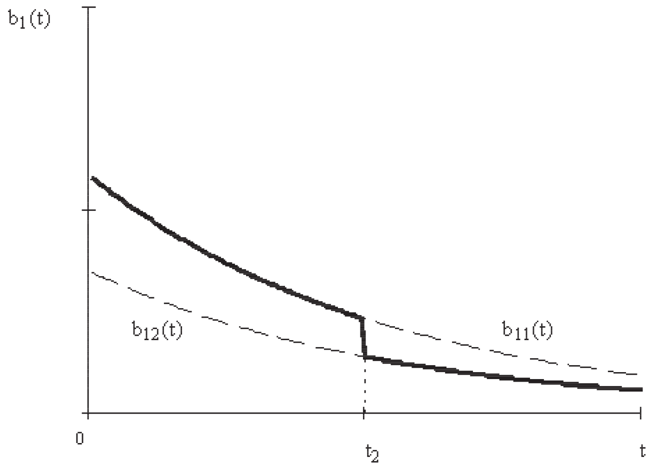


Figure 1: Benefit rate of player 1, $b_1(t)$.

Similarly, a strategy t_2^* of player 2 is said to be the best response of player 2 to a strategy t_1 of player 1 if t_2 maximizes the payoff to player 2, $P_2(t_1, t_2)$, over the set of all strategies of player 2, t_2:

$$P_2(t_1, t_2^*) = \max_{t_2 > 0} P_2(t_1, t_2).$$

The pair (t_1^*, t_2^*) is said to be a Nash equilibrium in the game if both of the preceding conditions are satisfied.

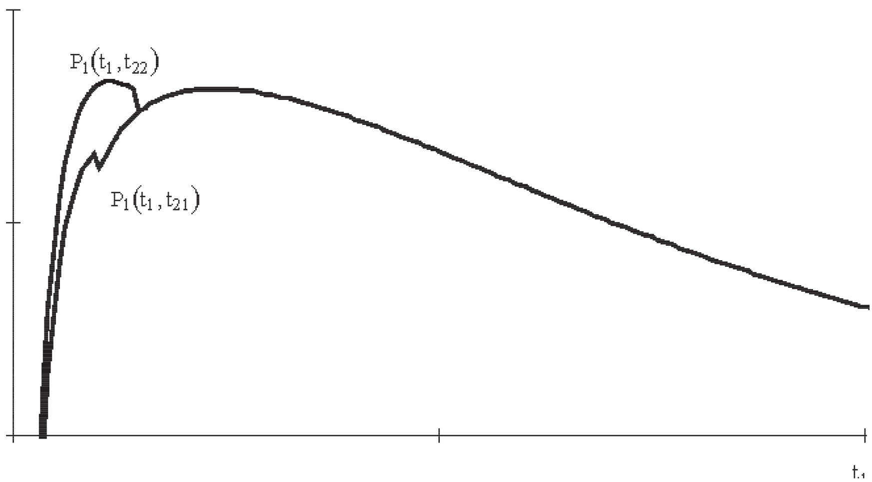


Figure 2: Examples of total profit curves comparison for player 1: $t_{22} > t_{21}$.

Thus, we are modeling the decision-making process related to constructing the new pipelines for the developing energy market as a competition between two energy suppliers formalized in the form of a noncooperative two-person game. The solutions that are to be found correspond to Nash equilibrium pairs t_1^*, t_2^* which define the rational commercialization times for the projects.

We call this game *the game of timing.* The information structure of such a game is quite simple: each player (investor) knows the main characteristics of both projects. That is enough to construct the Nash equilibrium point (if it exists), and its components are treated as the rational choice for the participants.

2 Theoretical Study of the Model

A detailed study of the game of timing is given in the cited paper [10]. Here we present the main assumptions and sketch the proofs of existence of Nash equilibrium solutions.

The assumptions are the following. We assume the functions $C_i(t)$ to be smooth, and the functions $a_i(t) = -dC_i(t)/dt$ to be positive and monotonically decreasing. It is also assumed that for each player i, the graph of the rate of cost reduction, $a_i(t)$, intersects the graph of the upper benefit rate, $b_{i1}(t)$, from above at the unique point $t_i^- > 0$, called the *fast choice*, and stays below it afterwards; similarly, the graph of $a_i(t)$ intersects the graph of $b_{i2}(t)$ from above at the unique point $t_i^+ > 0$, called the *slow choice*, and stays below it afterwards:

$$\begin{gathered} a_i(t) > b_{i1}(t) \qquad \text{for } 0 < t < t_i^-, \\ a_i(t_i^-) = b_{i1}(t_i^-), \quad a_i(t) < b_{i1}(t) \qquad \text{for } t > t_i^-; \end{gathered} \tag{1}$$

$$\begin{gathered} a_i(t) > b_{i2}(t) \qquad \text{for } 0 < t < t_i^+, \\ a_i(t_i^+) = b_{i2}(t_i^+), \quad a_i(t) < b_{i2}(t) \qquad \text{for } t > t_i^+. \end{gathered} \tag{2}$$

Note that since $a_i(t)$ is monotonically decreasing and $b_{i1}(t) > b_{i2}(t)$, we have $t_i^- < t_i^+$.

A graphical illustration is given in Figure 3.

Despite the fact that any concavity property of $P_1(t_1, t_2)$ and $P_2(t_1, t_2)$ does not follow from these conditions, it can be proved that Nash equilibrium points do exist and may be either unique or consist of two pairs.

This is concluded by the following propositions.

Proposition 2.1. *For every $t_1 > 0$ the payoff to player 1, $P_1(t_1, t_2)$, increases in t_2; moreover, given a $t_{21} > 0$ and $t_{22} > t_{21}$, one has $P_1(t_1, t_{22}) = P_1(t_1, t_{21})$ for $t_1 \geq t_{22}$ and $P_1(t_1, t_{22}) > P_1(t_1, t_{21})$ for $t_1 < t_{22}$.*

This property can be derived from the explicit expression for the derivative of the function $P_1(t_1, t_2)$ with respect to t_1 and accepted assumptions.

Analogously, we have a similar property for the second player.

Proposition 2.2. *For every $t_2 > 0$ the payoff to player 2, $P_2(t_1, t_2)$, increases in t_1; moreover, given a $t_{11} > 0$ and $t_{12} > t_{11}$, one has $P_2(t_{11}, t_2) = P_2(t_{12}, t_2)$ for $t_{12} \geq t_{11}$ and $P_2(t_{11}, t_2) > P_2(t_{12}, t_2)$ for $t_2 < t_{12}$.*

A more detailed study of the payoff functions $P_1(t_1, t_2)$ and $P_2(t_1, t_2)$ and their derivatives allows us to describe the best response of one player to any given choice of another. Let us fix, e.g., t_2—a strategy of the second player. Then one can prove the following result.

Proposition 2.3. *In the interval (t_1^-, t_1^+), there exists a unique point $\hat{t}_2$ such that*

$$P_1(t_1^-, \hat{t}_2) = P_1(t_1^+, \hat{t}_2). \tag{3}$$

If $0 < t_2 < \hat{t}_2$, then the unique best response of player 1 to t_2 is t_1^+ (slow choice). For $t_2 = \hat{t}_2$ the set of all best responses is $\{t_1^-, t_1^+\}$. And finally, if $t_2 > \hat{t}_2$, then the unique best response of player 1 to t_2 is t_1^- (fast choice).

This proposition shows a particular role that points t_1^-, t_1^+ play in the game under consideration. Similar arguments are true for player 2.

Proposition 2.4. *In the interval (t_2^-, t_2^+), there exists a unique point $\hat{t}_1$ such that*

$$P_2(\hat{t}_1, t_2^-) = P_2(\hat{t}_1, t_2^+). \tag{4}$$

If $0 < t_1 < \hat{t}_1$, then the unique best response of player 2 to t_1 is t_2^+ (slow choice). For $t_1 = \hat{t}_1$ the set of all best responses is $\{t_2^-, t_2^+\}$. And if $t_1 > \hat{t}_1$, then the unique best response of player 2 to t_1 is t_2^- (fast choice).

We call the points $\hat{t}_1$ and $\hat{t}_2$ the *switch* points of the players.

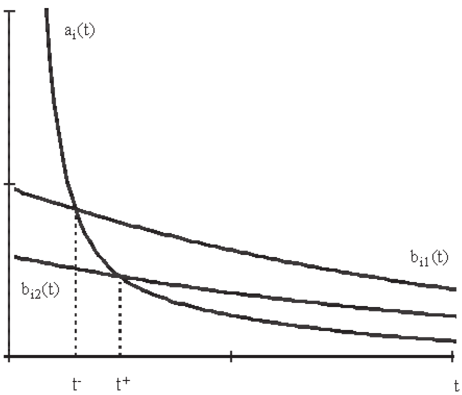


Figure 3: Fast and slow choices of the players.

From the last two propositions one can conclude that in the game of timing the Nash equilibrium solutions do exist and such solutions are concentrated in the set $\{t_1^-, t_1^+\} \times \{t_2^-, t_2^+\}$ of possible combinations of fast and slow choices of the players. The type of equilibrium is completely determined by the reciprocal location of the players' fast and slow choices and the switch points. In particular, the following theorem is true.

Theorem 2.1. *In cases $t_2^- \leq \hat{t}_2 < t_2^+$, $t_1^- < \hat{t}_1 \leq t_1^+$ and $t_2^- < \hat{t}_2 \leq t_2^+$, $t_1^- \leq \hat{t}_1 < t_1^+$ the game of timing has precisely two (fast-slow and slow-fast) Nash equilibria. Otherwise, in any possible situation, it has the unique (fast-slow or slow-fast) Nash equilibrium.*

In [10] one can find the complete classification of the equilibrium points.

We can obtain precise formulas for the introduced functions while using a specific model of price formation mechanism and that of the cost for finalizing the project at a fixed time. In Ref. [10] this has been done for the case when prices are determined by a function of Cobb–Douglas type and costs of construction are obtained as solutions of an optimal investment problem.

The main parameters that should be given for application of this version of the model are listed in Table 1.

The indicated parameters are used in the price formation model taken in the form

$$p(t) = \left(\frac{g(t)}{y(t)}\right)^{\beta}, \tag{5}$$

where $g(t)$ is the consumer's Gross Domestic Product (GDP), and $y(t)$ denotes gas supply which is supposed to be equal to gas demand. The cost for finalizing the project i at time t_i, $C_i(t_i)$, is defined as a solution to the special optimal control problem, which describes the optimal investment process, and is determined by the relation

$$C_i(t_i) = \rho^{\alpha-1} \frac{e^{-\lambda t_i} x_i^{\alpha}}{(1 - e^{-\rho t_i})^{\alpha-1}},$$

where $\alpha = 1/\gamma$, $\rho = (\alpha\sigma + \lambda)/(\alpha - 1)$.

3 Software G-TIME

For detailed research of the model, sensitivity analysis and econometric data application, the software "G-TIME" has been elaborated. It allows the analyst to:

- Perform simulations with various parameters of the model with numerical and graphical results.

Table 1: Parameters for specific price formation model.

Parameter	Description	Units	Range	Used values
λ	Discount coefficient	Share	$0 < \lambda < 1$	0.05–0.4
β	Integral coefficient determined by gas demand price elasticity, gas demand GDP elasticity and scalable parameter	Share	$0 < \beta < 1$	0.3–0.8
γ	Rate of investment return (delay)	Share	$0 < \gamma < 1$	0.5–0.7
σ	Obsolescence coefficient	Share	$\sigma \geq 0$	0–0.5
g_0	Consumer's GDP at time 0	billions US$	$g_0 > 0$	198
c_{10}, c_{20}	Cost for extraction, transportation and distribution and transit fees for players 1 and 2 respectively at time 0	US$ $/1000m^3$	$c_{i0} > 0$	60–80
x_1, x_2	Prescribed commercialization level for players 1 and 2 respectively at time t_i	billions US$	$x_i > 0$	2–6

- Obtain illustrations for the switch times definition process, and for profit surfaces for both players and to analyze the profit functions with a fixed opposite player time coordinate for both players.
- Calibrate the model on the basis of real econometric data.
- Provide sensitivity analysis of the model with respect to all parameters, and to obtain distributions of results corresponding to the parameter changes and related graphical illustrations.

The program has a user-friendly interface, designed to make the simulations process and results observation more convenient. For convenience, everything is compiled within one window so that the user can observe both the initial data and the results and their illustration in one view. G-TIME consists of one main window divided into 4 main parts:

(1) Input parameters section (Figure 4, A);
(2) Numerical results section (Figure 4, B);

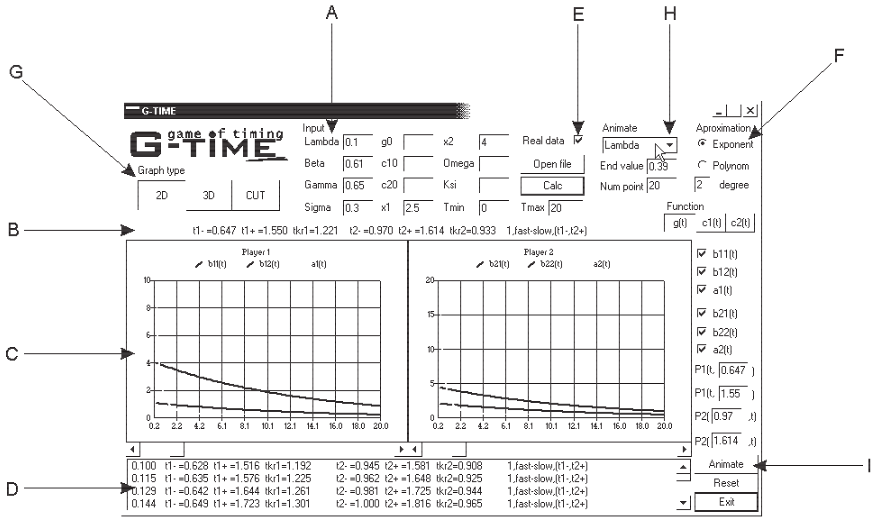


Figure 4: G-TIME: Graphical results illustration frame in 2D mode.

(3) Graphical results illustration frame (Figure 4, C);
(4) Sensitivity analysis results frame (Figure 4, D).

The different parts are described in more detail as follows.

Input parameters section: allows one to input the initial values for simulation: input parameters of the model, $\lambda, \beta, \gamma, \sigma$, commercialization investment levels, x_1 and x_2, and, for a theoretical approach, coefficients of GDP and costs of extraction development. Note the checked "real data" (Figure 4, E)—if it is chosen, the program uses a tab-given description for coefficients of GDP and costs of extraction development. The "open file" function allows the user to input an external file of special format where econometric data for functions are located and to apply an approximation algorithm (Figure 4, F) that puts the functions in the form appropriate for simulating process. The functions are approximated by exponential or polynomial ones. In the bottom-right corner of the section there is a "calculate" button—pressing on it causes the calculations to be performed.

Numerical results section: the values of simulation results—$t_i^+, t_i^-, t_i^\wedge$ and equilibrium type—are displayed.

Graphic results illustration frame: divided into two parts, each for one of two players. It works in three modes (Figure 4, G):

- 2D—in this mode diagrams illustrating rate of cost reduction, upper and lower benefit rate functions and points of their crossing are displayed
- 3D—in this mode the profit surfaces of both players are displayed (Figure 5)

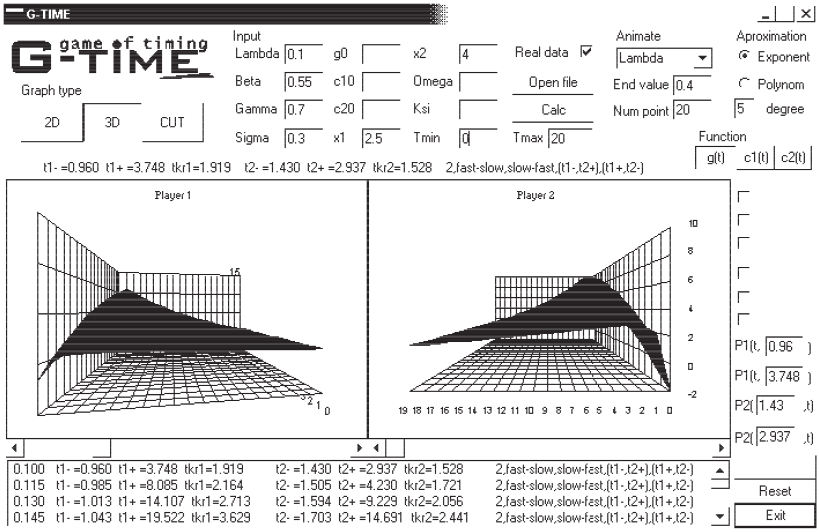


Figure 5: G-TIME: Graphical results illustration frame in 3D mode.

- CUT—in this mode the profit functions with the fixed time coordinate of the opposite player ($P1(t_1, t_2 = \text{const})$ for 1st and $P2(t_1 = \text{const}, t_2)$ for 2nd player) are displayed (Figure 6).

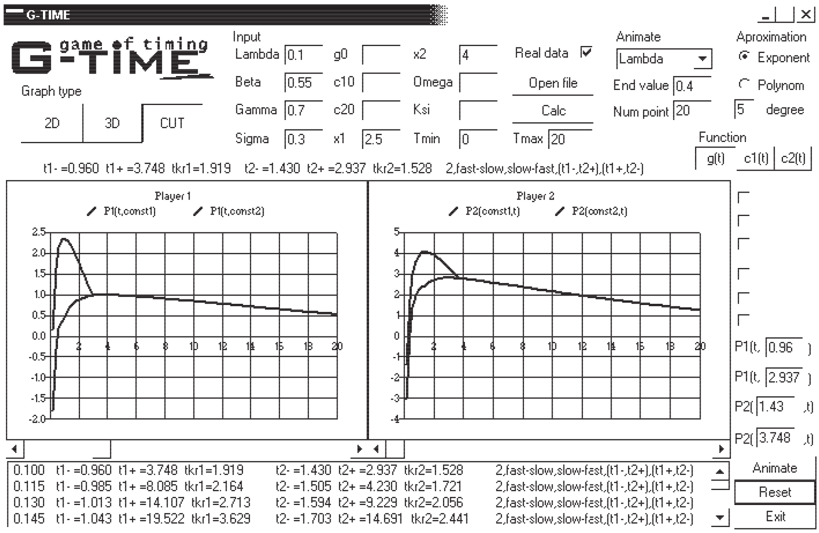


Figure 6: G-TIME: Graphical results illustration frame in CUT mode.

A right-click at any part of the main window leads to a pop-up menu where a full-screen mode preview of the current-mode diagram can be chosen.

Sensitivity analysis results frame: activated when one chooses the specific parameter relative to which one needs to examine the model's stability (robustness) (Figure 4, H), initial and final values for this parameter (the program takes the current value of the parameter which is being defined in A section of Figure 4 as the initial value) and number of points (the program automatically calculates the step). After one presses the "animate" (Figure 4, I) button, the program starts simulations varying the parameter in a given interval and putting the results at each step into the frame (Figure 4, D). In the chart the first column reflects the change of the chosen parameter. While simulating with the changing parameter, the animation of diagrams is displayed in the graphic results illustration frame (only in 2D and CUT modes). A right-click at any part of the main windows leads to a pop-up menu where one can choose to copy the results to a clipboard or to a file (Figure 4, J).

4 Results of Simulations

The routing of oil and gas pipelines in Asia and especially the Caspian region is at the center of the geopolitics of energy. One of the most promising markets in the region is Turkey, not in the least because Turkey constitutes a gateway from Asia to Europe. Official forecasts suggest that Turkey's gas demand might quintuple by 2010. Various countries in the Caspian region are interested in exporting gas to Turkey. Russia's Gazprom started to build the "Blue Stream" pipeline under the Black Sea to expand its current gas deliveries to Turkey, and Turkmenistan is heading for the Trans-Caspian gas pipeline to deliver gas to Turkey. It seems that some of these countries are moving ahead fast to preempt the investment decisions of others, i.e., making it unattractive to build a new transmission pipeline since the market might not be big enough. Currently gas (around 30% of demand in 1999) is being shipped to Turkey in the form of LNG from Algeria and Egypt. The remaining 70% comes from Russia via Bulgaria. For these reasons we focus on Turkey.

We associate a manager of the Trans-Caspian project with player 1 and a manager of the Blue Stream project with player 2.

The main task in calibrating the model described in the first section is to identify the value of β, which is the key parameter in the description of the price formation mechanism, corresponding, in turn, to a demand model. The demand model is taken in the form

$$d(t) = Ag(t)^{E_g} p(t)^{E_p}, \tag{6}$$

where $d(t)$ is demand at time t, $g(t)$ stands for GDP, $p(t)$ is the price at time t, E_g the GDP elasticity of demand, and E_p the price elasticity of demand.

Using the values for E_g and E_p, taking into account the equilibrium condition $y(t) = d(t)$, econometric data and forecasts for $g(t)$ and $d(t)$, one can estimate the value of β.

The software G-TIME allows one to realize the procedure described above and apply the model described in Section 1 or to use the relation (6) directly. In the last case we consider supply of each player as an exogenous variable and use the price modeling relation derived from (6).

There were several goals pursued by the authors while calibrating the model: to calibrate the original model described in [10] and, correspondingly, to identify the β parameter; to calibrate the model with a modified price formation mechanism and, correspondingly, to identify elasticity and scale parameters; and to compensate the inadequacy in the market capacity/agents supplies ratio due to a reduced number of players. For these purposes, the three scenarios listed below have been selected for the calibration.

Scenario 1. Further development of the gas market was modeled on the basis of historical data for the 1988–1998 period by using G-TIME software. The original price formation mechanism has been considered.

Scenario 2. Simulations were made with the modified model considering the price modeling mechanism and forecasts obtained by (6).

Scenario 3. Same as scenario 2, but with different supply modeling. The forecast for the total demand on the gas market for both agents plays a role of an upper bound for the supply. The supply of each agent while occupying the market solely equals the total demand; when the opponent enters the market, the supply of each agent equals his share determinated by his marginal capacity.

The results of identification of the parameters are presented in Table 2.

Values of parameters used in simulations: $\lambda = 0.1, \beta = 0.55$ (only for scenario 1), $\gamma = 0.65, \sigma = 0.3, x_1 = 2.5, x_2 = 4$. Time $t = 0$ corresponds to 2001 (year).

Results of the simulations are presented below.

The optimal values for players' times for entering the market obtained as a result of simulations in accordance with selected scenarios are also presented below.

Table 2: Parameter for the three scenarios.

Scenario	Identified β value	Time period	Coefficients
1	0.55	1988–1998	$A = 7 \times 10^{-9}$, $E_g = 4.2$, $E_p = -0.72$
2	1.95	2000–2020	$A = 0.47$, $E_g = 1.25$, $E_p = -0.7$
3	1.63	2000–2020	$A = 0.38$, $E_g = 1.25$, $E_p = -0.7$

Player 1: Trans-Caspian project. Player 2: Blue Stream project.

Scenario 1: Equilibrium type 2, *fast-slow*, *slow-fast*, $(t_1^-, t_2^+), (t_1^+, t_2^-)$

$$t_1^- = 1.183, \quad t_2^+ = 3.708; \qquad t_1^+ = 4.142, \quad t_2^- = 1.799.$$

This case brought two equilibriums. According to the *slow-fast* choice (*slow-fast* means that Player 1 enters the market later than Player 2), the Blue Stream project should commercialize at the end of 2002 and the Trans-Caspian project at in the beginning of 2005. This is economical evidence: when the market capacity is not big enough for a concurrence and one of the suppliers has already started operation on the market, it is not reasonable for a second supplier to compete for a small share of the market.

Remark 4.1. When the research was still in progress, at the beginning of 2001, PSG International Ltd. (an affiliate of GE Capital and Bechtel Enterprises)—one of the major investors of the Trans-Caspian project—quit the project and left it in a "frozen" state. A year after the research was finished, at the end of 2002, Blue Stream project managers announced the final stage of construction. And at the beginning of 2003 the supply of natural gas to Turkey started. After a year of operation, considering increasing natural gas demand, which corresponds to forecasts used in simulations, interest in the Trans-Caspian project has newly arisen, especially from the Shell Company—also one of the investors of the project.

According to the *fast-slow* choice, the Trans-Caspian project should commercialize in the beginning of 2002. This is almost a reversed situation, when the Trans-Caspian project holds leadership on the market and Blue Stream managers will have to postpone the construction until the "optimal time" when the market capacity will exceed the marginal capacity (or current supply level, if less) of the Trans-Caspian pipeline by a value enough to pay off the Blue Stream project, considering operation and maintenance costs. G-TIME software defines such an "optimal time" as the end of 2004 for the Blue Stream project.

Scenario 2: Equilibrium type 1, *fast-slow*, (t_1^-, t_2^+)

$$t_1^- = 1.132, \quad t_2^+ = 3.441$$

The results obtained during simulations for scenario 2 did not differ significantly from the ones obtained for scenario 1. That proves the validity of the assumption about parameter β as an integral coefficient reflecting the relation between gas demand and GDP elasticity coefficients and demonstrates that β was identified correctly over the indicated time period.

The main difference is in the uniqueness of the equilibrium, which means that each player will have only one strategy for the choice of commercialization time in order to maximize the integral profit over the period of operation. For the Trans-Caspian project such a strategy is the beginning of 2002, for the Blue Stream project it is the beginning of 2004.

Scenario 3: Equilibrium type 1, *fast-slow*, (t_1^-, t_2^+)

$$t_1^- = 1.069, \quad t_2^+ = 6.381$$

In the results of simulations according to scenario 3 one can track an increased gap between "optimal" commercialization times. This can be explained by the specificity of demand modeling for each player. Since a fraction of the market allocated to each agent was considered to be proportional to a marginal capacity of the correspondent pipeline, and the cumulative supplies of both pipelines were equal to the natural gas demand on the market, there was more time needed for the Blue Stream project, whose "optimal" commercialization time is the beginning of 2007, to get its fraction and to enter the market. The optimal timing for the Trans-Caspian project, as in scenario 2, is the beginning of 2002.

5 Generalized Nash equilibrium

In previous sections we considered the Nash equilibria for two participants (players) who operate on a specific gas market. The profit functions $P1(t_1, t_2)$ and $P2(t_1, t_2)$ for the 1st and 2nd players respectively were defined as deterministic functions depending on commercialization times t_1 and t_2 for the corresponding projects. In fact, the profit functions $P1(t_1, t_2)$ and $P2(t_1, t_2)$ depend on a set of parameters: discount factor λ, elasticity coefficients and so on. Some of these parameters can be considered as deterministic ones, but others are really uncertain and not known in advance. Assume that these uncertain parameters can be modeled using probabilistic techniques. In what follows we suppose them to be stochastic variables with given distribution functions. Then, at least theoretically, one can assume that the profit functions $P_1(t_1, t_2)$ and $P_2(t_1, t_2)$ for any t_1 and t_2 are also stochastic variables with corresponding distribution functions. Putting aside for the moment the problem of determining these distributions, consider the two characteristics of these stochastic variables. Namely, denote by $\mu_i(t_1, t_2)$ the mathematical expectation (mean value) of the variable $P_i(t_1, t_2)$ and by $\sigma_i(t_1, t_2)$ the corresponding standard deviation.

Thus,

$$\begin{aligned}
\mu_i(t_1, t_2) &= \mathbf{E}[P_i(t_1, t_2)], \\
\sigma_i^2(t_1, t_2) &= \mathbf{E}[(P_i(t_1, t_2) - \mu_i(t_1, t_2))^2], \qquad i = 1, 2.
\end{aligned}$$

According to the methodology of *the mean-variance* analysis one can associate the value μ_i with an expected profit for the corresponding player, and the value

σ_i with the evaluation of risk, which is to be minimized. Using this approach we will define a generalized best reply t_1^* of the first player to a fixed strategy $t_2^\wedge$ of the second one as a strategy, which in some sense simultaneously maximizes the mean value $\mu_1(t_1, t_2^\wedge)$ and minimizes $\sigma_1(t_1, t_2^\wedge)$.

Definition 5.1. A strategy t_1^* of the first player is said to be the G-best reply to the strategy $t_2^\wedge$ of the second player if there is no strategy t_1 for which the following inequalities hold simultaneously:

$$\mu_1(t_1, t_2^\wedge) \geq \mu_1(t_1^*, t_2^\wedge);$$
$$\sigma_1(t_1, t_2^\wedge) \leq \sigma_1(t_1^*, t_2^\wedge),$$

where at least one inequality is strict.

This is nothing else than the well-known definition of the Pareto optimal solution for a two-criteria optimization problem:

$$\mu_1(t_1, t_2^\wedge) \underset{t_1}{\to} \max$$

$$\sigma_1(t_1, t_2^\wedge) \underset{t_1}{\to} \min$$

The set of the points t_1^* which satisfy the above definition is called GBR1(t_2).

The same arguments allow us to define the set of best replies t_2^* of player 2 to a given $t_1^\wedge$ of player 1.

Definition 5.2. A strategy t_2^* of the second player is said to be the G-best reply to the strategy $t_1^\wedge$ of the first player if there is no strategy t_2 for which the following inequalities are simultaneously true:

$$\mu_2(t_1^\wedge, t_2) \geq \mu_2(t_1^\wedge, t_2^*);$$
$$\sigma_2(t_1^\wedge, t_2) \leq \sigma_2(t_1^\wedge, t_2^*);$$

where at least one inequality is strict.

And again we introduce a set of points t_2^*, each of which satisfies the definition, denoting it by GPR2(t_1).

Definition 5.3. A pair $\{t_1^*, t_2^*\}$ is said to be a generalized equilibrium if

$$t_1^* \in GBR1(t_2^*)$$

and

$$t_2^* \in GBR2(t_1^*).$$

We will denote by GEP the set of all the points $\{t_1^*, t_2^*\}$ which satisfy the above definition.

Remark 5.1. In a particular case, when the distributions for both functions $P_1(t_1, t_2)$ and $P_2(t_1, t_2)$ are concentrated in unique points for every admissible t_1 and t_2 (it means that both functions are determined), then the set GEP coincides with the Nash equilibrium points.

A numerical algorithm has been developed to obtain the GEP points in the framework of the computer model.

6 Conclusions

Approaches to energy infrastructure modeling based on game-theoretical methods have been developed, resulting in a new prototype computer model G-TIME, which combines game-theoretical and probabilistic (statistical) approaches to reflect uncertainty. With the help of this software a model-based analysis of Turkey's gas market has been carried out. The sensitivity analysis of the models showed that the results of simulations are very sensitive to changes of some inputs. A standard way to overcome the difficulties connected with risk and uncertainty caused by stochastic disturbances is the mean-variance approach. A methodological background related to a generalized Nash equilibrium was developed, which allows one to combine the mean-variance approach and the notion of Nash equilibrium.

REFERENCES

[1] Arrow K.J. and Kurz M., *Public Investment, the Rate of Return and Optimal Fiscal Policy*, Johns Hopkins University Press, Baltimore, MD, 1970.

[2] Barzel Y., Optimal timing of innovations, *Rev. Econ. and Statist.*, 50, 348–355, 1968.

[3] Basar T. and Olsder G.J., *Dynamic Noncooperative Game Theory*, Academic Press, London, 1982.

[4] Friedman D., Evolutionary Games in Economics, *Econometrica*, 59, 637–666, 1991.

[5] Gusev M.I. and Kurzhanskii A.B., Equilibrium situations in multi-criterion game problems, *Dokl. Akad. Nauk*, 229, 6, 1295–1298, 1976.

[6] Hofbauer J. and Sigmund K., *The Theory of Evolution and Dynamic Systems*, Cambridge University Press, Cambridge, 1988.

[7] Intriligator M., *Mathematical Optimization and Economic Theory*, Prentice-Hall, New York, 1971.

[8] Klaassen G., Roehrl R.A. and Tarasyev A.M., The Great Caspian Pipeline Game, *Proc. IIASA Workshops*, May 2001.

[9] Klaassen G., McDonald A. and Zhao J., The future of gas infrastructures in Eurasia, *Energy Policy*, 29, 399–413, 2001.

[10] Klaassen G., Kryazhimskii A. and Tarasyev A., Multiequilibrium game of timing and competition of gas pipeline projects, *JOTA*, 120, 1, 147–179, 2004.

The Effects of Incomplete Information in Stochastic Common-Stock Harvesting Games*

Robert McKelvey
Department of Mathematical Sciences
University of Montana
Missoula, Montana 59812, U.S.A.
RWMcKelvey@msn.edu

Peter V. Golubtsov
Physics Department
Moscow State Lomonosov University
Moscow, 119899, Russia
P_V_G@mail.ru

Abstract

Here the dynamic fishery harvesting game is generalized to a stochastic environment in order to examine the implications of incomplete and asymmetric information. The main emphasis is on a split stream version of the game: At the beginning of each harvest season the initial fish stock (or "recruitment") divides into two streams, each one accessible to harvest by just one of the two competing fishing fleets. The fleets simultaneously harvest down their streams, achieving net seasonal payoffs for the catch. After harvest, the residual sub-stocks reunite to form the broodstock for the subsequent generation. The strength of this subsequent generation is determined by a specified "stock-recruitment relation," and the cycle repeats. In this cyclic process, both natural environmental factors (stream-split proportions and stock-recruitment relation) and economic factors (harvest costs and benefits) will incorporate Markovian stochastic elements. At the beginning of each season, both fleets know the current recruitment and also have some (generally incomplete or delayed, and often asymmetric) knowledge of the current values of the stochastic elements. The knowledge structure of each specific game version is held in common by the competitors. In the dynamic game each fleet sets its harvest policy with the objective of maximizing the expected discounted sum of seasonal payoffs, and conditional on the extent of its current knowledge and of the anticipated policy of its competitor.

The implications of alternative knowledge structures are explored, through dynamic programming and simulation. Both information structures and the stochastic characteristics of bioeconomic parameters are varied continuously to explore their interplay. The asymmetric trade-offs among them are examined. The focus is on demonstrating the often unexpected, and sometimes counter-intuitive, effects that knowledge enrichment may have in these incomplete-information, common-property games.

Key words. Stochastic harvesting game, Asymmetric information, Fish wars, Bioeconomic competition models, Stochastic dynamic programming.

1 Introduction

This chapter studies various types of stochastic dynamical harvesting games. Following a long tradition [1, 2], the specific context is taken to be a marine fishery.

Two independently operated fleets competitively harvest a fish stock. The harvesting occurs annually, and each fleet chooses its harvest policy in such a way as to optimize discounted long-run net returns—given the expected harvest policies of its competitor. Thus, the projected outcome of the game is a dynamic Nash equilibrium.

The classical version of the game has the two fleets harvesting simultaneously from a common pool. But here we also consider situations where the fleets may have access to different parts of the fish stock at different times and places during the season. For example, the fish stock may be harvested sequentially by the fleets, as it travels a migration route back from its adult feeding range to its spawning grounds. Or the migrating stock may split into separate streams as it returns, each stream being accessible to harvest by only one of the fleets. Much of our focus here will be on the latter case, but our formulation of the game in dynamic programming format will remain general, and thus applicable to other possible situations.

Our primary focus is on fishery harvesting games which are played out in a stochastic environment, and in which the fleets must formulate and implement their policies under circumstances of incomplete information. The stochastic character of these games will be expressed through a random Markov parameter (which can be multidimensional).

We shall explore a wide range of information structures in our games. In particular, players may possess various levels of knowledge about the realization of a particular random parameter, e.g., current or delayed information, or even no information other than its Markov transition probability distribution. Alternatively, they may have only partial information of a parameter value, information obtained from imperfect observation. Moreover, the information structure may be asymmetric, e.g., one player may possess full current information while the other has only delayed or imperfect information. However, in all cases the information structure of the game is common knowledge [3].

*This material is based on work supported by the U.S. National Science Foundation under Grant No. 9708475.

For a related earlier study of incomplete-information stochastic harvesting games, see [4] and [5].

In such games it may well happen that a refinement of information may have unexpected, seemingly perverse, effects on the outcome of the game. For example, simulations show that (unlike in a single-player game) enhanced information may be harmful for both players. It may only lead them to harvest more aggressively, and thus may result in destructive stock reduction which will adversely affect future returns.

Other counter-intuitive situations may result when the information structure is asymmetric. Usually the player who knows more will do better than his competitor, but it will sometimes be to his advantage to reveal his private information, with the result that both will then do better. Other, seemingly counter-intuitive, simulation results are summarized in Section 2.4.

These situations run counter to those observed when players cooperate, by sharing information and adopting a common policy, in order to maximize their joint return. In this case not only do they both benefit from cooperation, but the benefits never are decreased by gaining additional information.

The structure of this chapter is the following: after the description of basic harvesting games and a brief problem statement we first present a collection of striking simulation results. All the technical details of deriving Nash equilibrium harvesting policies are given in the second half of the paper.

These involve reducing optimal reaction policies to the corresponding dynamical programming algorithms, and generalizing the classical optimal control technique, see, e.g., [6].

2 Informal Problem Description and Simulations for Split Stream Harvesting

2.1 Split Stream Harvesting

In the split stock harvesting model we assume that each player harvests in his own stream and that the random split of stock between streams may be unknown or imperfectly known to the players. The split stock harvesting game is illustrated by the following diagram:

$$\begin{tikzcd}
 & R^\alpha \arrow[r] & S^\alpha \arrow[dr] & & \\
R \arrow[ur] \arrow[dr] & & & S \arrow[r] & R_+ \\
 & R^\beta \arrow[r] & S^\beta \arrow[ur] & &
\end{tikzcd}$$

Here R is the current year's harvestable stock level, or "recruitment," and R^α and R^β are partial recruitments, in the separate streams, accessible to players α

and β, respectively. Thus,

$$R^\alpha = \theta^\alpha R, \qquad R^\beta = \theta^\beta R,$$

where

$$\theta^\alpha = \theta, \qquad \theta^\beta = 1 - \theta,$$

and θ is a random split factor. The residual substream stock or "escapement," following harvest, is denoted by S^α or S^β, respectively. These are determined by

$$S^\alpha = \sigma^\alpha(R^\alpha, p^\alpha), \qquad S^\beta = \sigma^\beta(R^\beta, p^\beta).$$

Here p^α and p^β are players' harvesting policies for this season. Typically, we shall define policies as escapement fractions so that $S^\alpha = p^\alpha R^\alpha$ and $S^\beta = p^\beta R^\beta$. Finally, the substream escapements combine to form the current year's total escapement

$$S = S^\alpha + S^\beta,$$

which is the broodstock, for determining the following year's recruitment R_+, through the "stock-recruitment relation" (or "growth function"):

$$R_+ = F(S, \varphi),$$

where φ is a random disturbance of the growth function.

We will assume that θ and φ are random variables and denote a whole set of random variables by a single symbol ν. Each player's policy depends on the mutually known information structure of the game, and on the specific information that a player has when he makes his annual harvest decisions. We will always assume that both players know the current total recruitment R, and also that each one has some information ξ^α and ξ^β about current and past random disturbances. Thus,

$$p^\alpha = P^\alpha(R, \xi^\alpha), \qquad p^\beta = P^\beta(R, \xi^\beta),$$

where P^α and P^β are the players' decision strategies.

The situation would change in fundamental ways if the players did not know the total recruitment, but each knew only that portion recruited to his own stream. In this case the policies would depend on the respective stream recruitments, i.e.,

$$p^\alpha = P^\alpha(R^\alpha, \xi^\alpha), \qquad p^\beta = P^\beta(R^\beta, \xi^\beta).$$

We will report on this alternative, less informationally rich situation in a subsequent publication.

The extent of players' knowledge about the random split fraction θ or φ may vary. In all cases we assume that both players know at least the stochastic properties of random parameters (e.g., transition probability distribution for a Markov process). In addition a player may have additional information: e.g., may have full current knowledge (this season's value), or only delayed knowledge (the previous season's value), or may know the result of imperfect observation of a current parameter value. Alternatively, he may possess no additional knowledge at all.

Furthermore, the structure of knowledge may be asymmetric; that is, the players may have differing levels of knowledge.

In each season a player gets a net return (annual payoff) v^α or v^β, which is a function of his stream's recruitment and his own policy, i.e.,

$$v^\alpha = v^\alpha(R^\alpha, p^\alpha), \qquad v^\beta = v^\beta(R^\beta, p^\beta).$$

The player's payoff in the dynamic game is taken to be a discounted sum of his seasonal returns over the time span of the game.

2.2 The Infinite Horizon Harvesting Problem

In all of our simulations we shall consider the infinite time-horizon risk-neutral harvesting game, but only as the limit of a sequence of long finite time-horizon problems.

Our numerical procedures (see Sections 3.2–3.3 and 4.1 for more details) are based on a dynamic programming formulation of the game-theoretic equilibrium solution of each finite horizon game in such a convergent sequence. This approach implies an exclusion of any infinite horizon equilibria other than those arising through limits of such sequential processes.

Of course, the numerical procedure can only approximate the limit of equilibria of harvesting games with a long finite horizon—but this is, for practical applications, what is of interest. It is well known that some infinite horizon problems can have a great many equilibria that are not of this sort.

Denote a decision policy at time τ by P^α_τ and a sequence of decision functions P^α_τ from the moment t until T by

$$\mathbf{P}^\alpha_t = \langle P^\alpha_t, P^\alpha_{t+1}, \ldots, P^\alpha_T \rangle = \langle P^\alpha_t, \mathbf{P}^\alpha_{t+1} \rangle.$$

Thus, the complete policy sequence for the player α from the season $t = 0$ until the season $t = T$ is the sequence $\mathbf{P}^\alpha_0$.

Each player's objective in the game with finite time-horizon T is to choose an optimal policy to maximize the expected discounted sum of his annual payoffs, given the policy of his competitor. Thus, player α will choose $\mathbf{P}^\alpha_0$ conditional on $\mathbf{P}^\beta_0$ to maximize

$$U^\alpha_0(R_0, \mathbf{P}^\alpha_0, \mathbf{P}^\beta_0) = \mathsf{E}\left[\left(\sum_{t=0}^{T-1} \gamma^t_\alpha v^\alpha(R^\alpha_t, p^\alpha_t, \nu_t)\right) + \gamma^T_\alpha v^\alpha(R^\alpha_T, \widetilde{p}^\alpha_T, \nu_T)\right].$$

Here differing policies $\widetilde{p}^\alpha_T, \widetilde{p}^\beta_T$ are specified at terminal time T, to reflect differing objectives for the status of the fish stock at time $T+1$. The typical options are either to specify directly the substream escapements at time T or to assign those escapements a positive "scrap value."

To incorporate "scrap value" into the problem statement we alter the objective function adding the (discounted) term $\mathsf{E}\gamma_\alpha^{T+1} Z_\alpha(R_{T+1})$, where Z_α is monotone increasing. The chosen policy will maximize the sum, trading off the payoff for harvested fish through time T against the value of the stock retained at time $T+1$.

Thus we have a two-player game, which consists of finding a Nash equilibrium

$$\begin{cases} \max\limits_{\mathbf{P}_0^\alpha} U^\alpha(R, \mathbf{P}^\alpha, \mathbf{P}^\beta), \\ \max\limits_{\mathbf{P}_0^\beta} U^\beta(R, \mathbf{P}^\alpha, \mathbf{P}^\beta) \end{cases}$$

for all possible R and for all policy functions $\mathbf{P}_0^\alpha$, $\mathbf{P}_0^\beta$. Then, if the pair $\widehat{\mathbf{P}}_0^\alpha$, $\widehat{\mathbf{P}}_0^\beta$ constitute a Nash equilibrium,

$$U^\alpha(R, \widehat{\mathbf{P}}_0^\alpha, \widehat{\mathbf{P}}_0^\beta) \geqslant U^\alpha(R, \mathbf{P}_0^\alpha, \widehat{\mathbf{P}}_0^\beta),$$
$$U^\beta(R, \widehat{\mathbf{P}}_0^\alpha, \widehat{\mathbf{P}}_0^\beta) \geqslant U^\beta(R, \widehat{\mathbf{P}}_0^\alpha, \mathbf{P}_0^\beta).$$

We find, at least in the risk-neutral games which we simulate, that each of these finite time-horizon games has a unique solution, in a pair of optimal policies, which will be non-stationary and will depend on the particular terminal-time policies adopted. Furthermore, in the limit of the infinite horizon game the sequence of finite time-horizon Nash equilibria policy pairs converges to a common stationary policy, which optimizes the limit objectives

$$U^\alpha(R_0, \mathbf{P}_0^\alpha, \mathbf{P}_0^\beta) = \mathsf{E}\sum_{t=0}^{\infty} \gamma_\alpha^t v^\alpha(R_t^\alpha, p_t^\alpha, \nu_t).$$

The Nash equilibrium formulation of the modified game, in which each player has only information R^α or R^β about recruitment in his own stream, may be formulated in a similar fashion.

2.3 Results of Simulation

In this section we present the results of simulations for various game parameters and for various information structures. All the simulations are based on dynamic programming algorithms for the corresponding harvesting problems, as described in Sections 3.2 and 3.3.

The knowledge that each player has at time t always includes the current total recruitment R_t. In addition he will have some information about the stochastic parameter θ, which determines the split of flow between the two streams. Knowledge of θ includes, at minimum, the probability transition matrix for this Markovian random sequence. In these simulations, the stock-recruitment parameters are deterministic.

A player may have only this minimum knowledge θ or in addition may know, at time t, historical values of θ up through the previous season's value θ_{t-1}. He even may know the current value of θ_t.

Alternatively, a player may obtain partial information of the flow-split from the result of imperfect observation of the split fraction θ, through a measurement parameter ξ. Measurement of ξ_t specifies only a known conditional probability distribution of the true value θ_t. When studying games with measurement information we can augment measurement precision continuously, thus revealing the influence of the degree of knowledge on the outcome of competition.

The knowledge state may be symmetric, with both players having the same information, or may be asymmetric, e.g., one player may have current knowledge of θ while the other has only delayed knowledge, or one may have measurement information while the other knows only the Markov transition function.

The other items that may be incorporated in simulations include various classes of annual return function, incorporating adjustable parameters; e.g., the cost of unit harvesting effort may be varied. (See Section 4.2 for details.) Finally, various classes of stock-recruitment growth functions, also incorporating adjustable parameters, will be considered. (See Section 4.3.)

In the simulations displayed below we show time-averaged values (of, e.g., expected annual payoff, or seasonal escapement), averaged over sufficiently long periods (typically 2000 time steps) to achieve stability.

2.3.1 Influence of Harvesting Cost on Competition

Here we examine the influence, on the outcome of the split stream harvesting game, of alternative information structures, with a fixed (i.e., deterministic) growth function $F(S)$ and a range of levels of the constant unit harvesting cost parameter c (see Section 4.2).

In the first set of simulations (Figure 1) we assume that $F(S)$ is of "compensatory" type, i.e., is monotone increasing and concave, with a positive fix-point at the "carrying capacity" $S = K$, where $F(K) = K$. Specifically, we utilize the "cubic growth function" (see Section 4.3 and Figure 15 bottom), normalizing the carrying capacity to $K = 1$. The graphs displayed show predicted game outcomes for five distinct circumstances: competitive harvesting with complete symmetric current information about θ ("Cur"); competitive harvesting with asymmetric (current vs. minimal) information, where the first player ("Cur-Min 1") has current information, while the second one ("Cur-Min 2") has only minimal knowledge; cooperative harvesting with current information ("Cur Coop"); competitive harvesting with minimal symmetric information ("Min"); and finally cooperative harvesting with minimum information ("Min Coop").

The graphs in Figure 1 show, respectively, the average annual steady-state payoffs to the players and the corresponding stock escapement levels. Unless

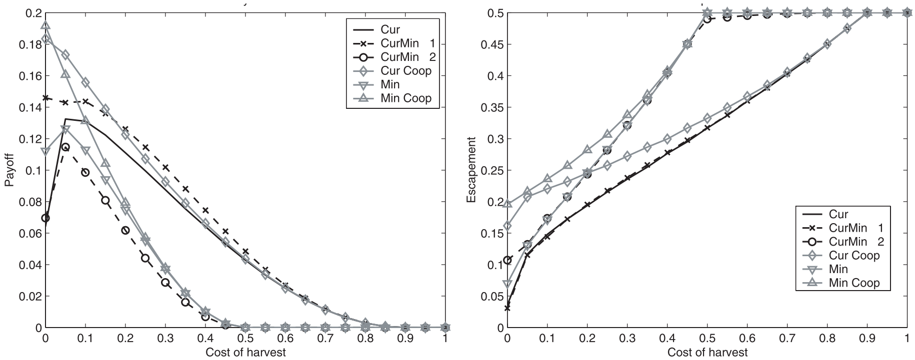


Figure 1: Influence of harvesting cost for different types of knowledge. "Compensatory" cubic growth function, see Figure 15 bottom.

harvest costs are very low ($c < 0.1$), the result of increased harvesting costs is to decrease individual payoffs and to increase separate-stream escapements.

Furthermore, provided the players possess identical knowledge, their individual payoffs increase as their information increases and, at a given information level, are greater when the players cooperate than when they compete.

In the symmetric knowledge cases (solid lines) both players benefit from additional information, but the first player would still prefer the asymmetric case ("Cur-Min 1").

Finally, when the competing players' information levels differ, then player 1 ("Cur-Min 1"), with more complete information, will do better than player 2 ("Cur-Min 2"), and better than he would have done even had he cooperated fully ("Cur Coop"), assuming that the cooperating players' goal was to set their individual harvests to maximize the sum of their payoffs.

All of these results seem natural, and may be thought of as displaying a certain baseline condition to be expected from compensatory growth.

The anomalous behavior in the competition models, seen when costs are very low, is real: It is the result of competition driving stock escapements down into a region of very slow growth, so that payoffs drop. With cooperation this low escapement hazard is avoided.

These negative effects of competition become much more pronounced when the stock-recruitment relation is depensatory, especially when it exhibits critical depensation. This is demonstrated in Figures 2 and 3.

The second set of figures (Figure 2) result when one assumes, instead of compensatory growth, that the growth function is assumed to display "non-critical depensation." That is, it remains monotone increasing, but is convex for small S, but with $S < F(S)$ remaining true on all of $0 < S < K$. The effect is that the recruitment gain remains small on an extended interval of low S.

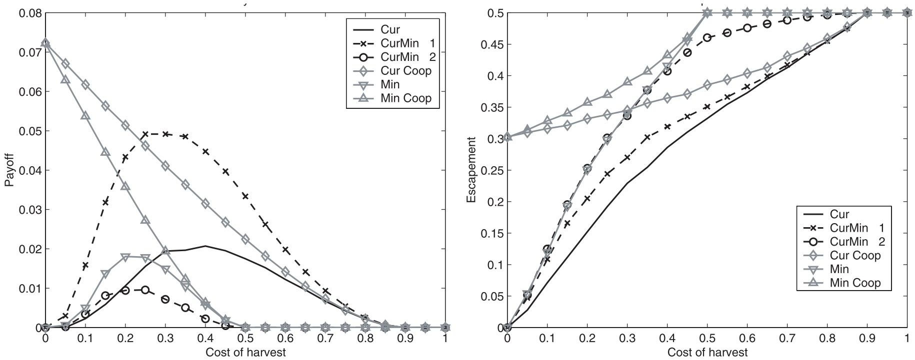


Figure 2: Influence of harvesting cost for different types of knowledge. Growth function with "non-critical depensation," see Figure 15 middle.

In Figure 3, where the growth function exhibits critical depensation, the effects are even more pronounced. In this case, $F(S) < S$ on an interval $0 < S < S_0 < K$ and should escapement drop into this region, stock extinction would become inevitable.

The most obvious difference from the compensatory case, when the players compete, is the reduction of mean payoff as well as expected escapement levels when harvesting costs are low and thus harvest effort is large. This effect is not seen when the players cooperate and hence are able to avoid the critical region.

It is seen again in Figure 3 (left) that an information advantage in an asymmetric knowledge case is highly beneficial for the first player ("Cur-Min 1"). Moreover, he would not wish to share his additional knowledge with his competitor and thereby switch to the symmetric complete knowledge case ("Cur"). Typically, the second player ("Cur-Min 2") in the asymmetric case would pre-

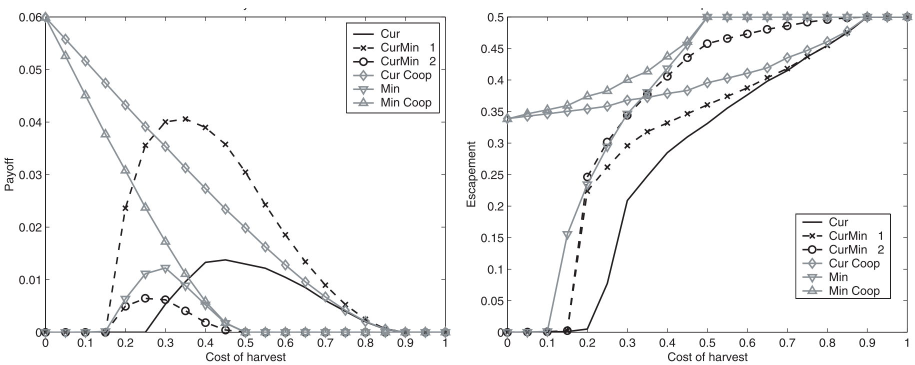


Figure 3: Influence of harvesting cost for different types of knowledge. Growth function with "critical depensation," see Figure 15 top.

fer to get current knowledge—but not for low cost (< 0.3) where asymmetrical knowledge is more beneficial (even for him) than symmetrical complete current knowledge ("Cur"). Furthermore, at low enough costs (< 0.35), minimal information ("Min") is more beneficial than complete current knowledge ("Cur"). This is because, in the absence of precise knowledge, harvest levels must be compromises, and hence extremely low escapements are avoided.

As noted above, competition becomes especially destructive, for low cost of harvest, in the critical depensatory case where overharvest can completely destroy the stock. However, if the players cooperate their return is significantly higher, especially when the cost of harvest is low for complete information ("Cur Coop") and for minimal information ("Min Coop"). Cooperatively, in contrast to the competitive case they are able to hold expected escapements at relatively high levels.

2.3.2 Harvesting with Information from Imperfect Observations

In this set of simulations, information about current θ is obtained from imperfect measurements, so that its impacts can be compared along a continuum. The measurement accuracy is a variable parameter, increasing from 0 (no information) to 1 (complete information).

It is assumed once again that stock recruitment conforms to the "Cubic" growth function (see Section 4.3), and thus may display compensation, noncritical depensation or critical depensation, depending on the growth parameter value. But now the unit harvesting cost c is fixed throughout.

Figures 4–6 show game outcomes for three types of games: competitive harvesting with imperfect measurement information about θ ("Meas"); competitive harvesting with asymmetric (measurement vs. minimal) information, where the first player ("Meas-Min 1") obtains measurement information, while the second one ("Meas-Min 2") has only minimal knowledge; and cooperative harvesting with measurement information ("Meas Coop").

In the first set of simulations (Figure 4) stock recruitment is compensatory, and the demonstrated effect of enhanced information is quite intuitive.

As always, when the players cooperate ("Meas Coop") their total return will be maximized, and the greater their (shared) information, the greater their payoff will be—at the expense of a diminished fish-stock escapement.

In a competitive game, where the players retain the same level of knowledge ("Meas"), both players still benefit from additional information, though to a lesser degree than with cooperation, and the stock escapement is depressed even more than with cooperation.

Finally, under asymmetric-knowledge competition (dashed lines), the information-advantaged first player ("Meas-Min 1") will do even better (at high precision) than he would by cooperating and sharing knowledge, while the second player ("Meas-Min 2") will do much worse. As the first player's knowledge is enhanced, the total stock escapement level in the two streams will drop

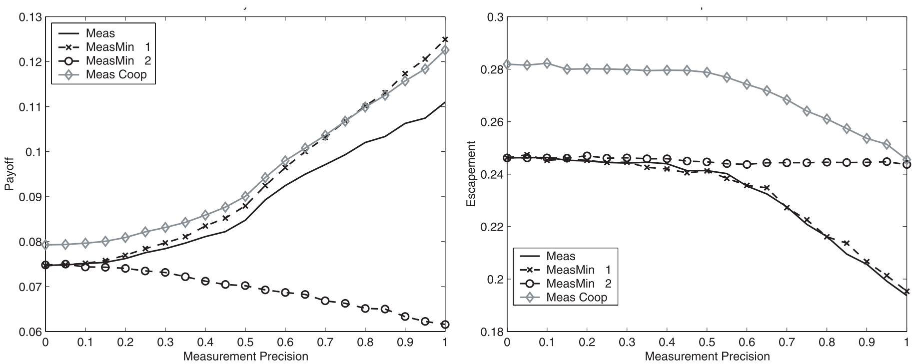


Figure 4: Compensatory harvesting with information from imperfect observation.

as precision increases—though not as rapidly as with cooperation, but to a much lower level, and in fact, the stock escapement in the first player's stream ("Meas-Min 1") will drop, while the stock escapement in the second player's stream ("Meas-Min 2") will remain almost unchanged.

In the second and third sets of simulations (Figures 5 and 6) stock-recruitment displays, with respectively, (non-critical) depensation (Figure 5; see also Figure 15 middle), and critical depensation (Figure 6, see also Figure 15 top) are shown.

Note that if the players cooperate ("Meas Coop"), the usual pattern holds that enhanced information yields better payoffs. As before, this result is correlated with a drop in escapement levels, but escapement must, with high probability, be held above a critical level (here, around 0.35). The sum of payoffs here necessarily will be superior to those in competitive situations, with or without informational asymmetry.

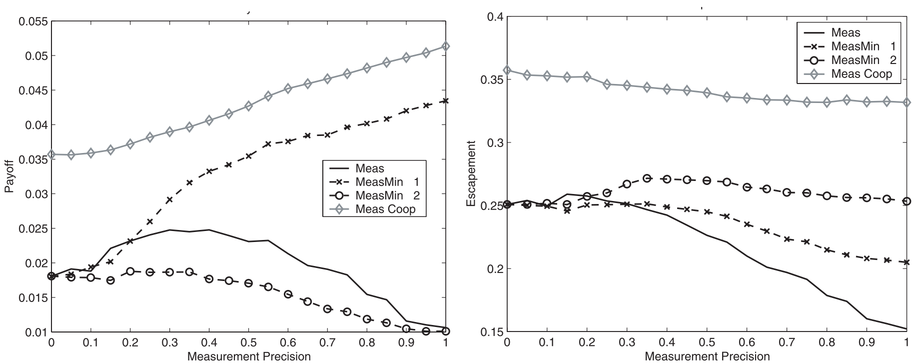


Figure 5: Depensatory harvesting with information from imperfect observation ("non-critical depensation").

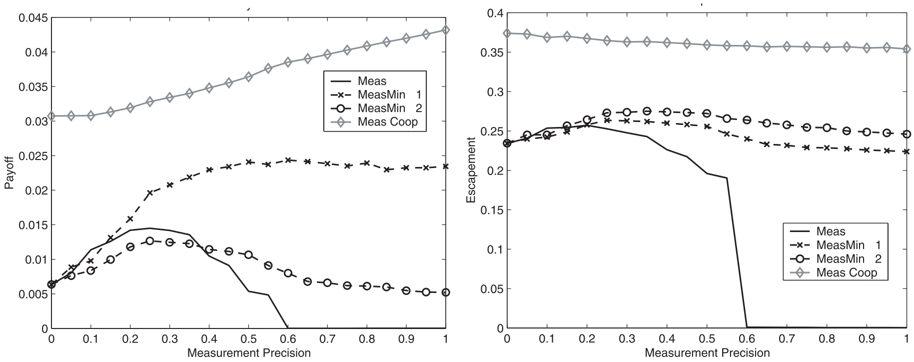


Figure 6: Depensatory harvesting with information from imperfect observation ("critical depensation")

It is seen, especially for the "critical depensation" case (Figure 6), in competitive games with symmetric ("Meas") and asymmetric ("Meas-Min") information structures, that below a certain level additional information is beneficial to both players, even for the player who does not possess the additional information ("Meas-Min 2"). However, further increasing the knowledge level degrades the situation dramatically, presumably by making harvesting policies more aggressive.

In addition, for low measurement accuracy the situation when the second player also has access to the measurement information is better for him then when he does not ("Meas" vs. "Meas-Min 2"). However, for high enough accuracy levels (> 0.38) his additional knowledge does not benefit him.

At the same time cooperative management ("Meas Coop") provides a much higher return, which is constantly growing with the increase of measurement accuracy.

Reduction of payoff at high precision of measurements is connected with significant reduction of escapement. It is interesting that in the cooperation case ("Meas Coop") escapement also decreases with the increase of knowledge, but very slightly, and always remains at a very high (compared to competition cases) level.

2.3.3 Harvesting under Various Degrees of Compensation or Depensation

In this set of simulations we show the interplay between the knowledge structure and the level of depensation by continuously varying the degree of depensation or compensation of the cubic growth function. The stock-recruitment parameter b is fixed ($b = 0$), while A, which is equal to the derivative of the cubic growth function at $S = 0$, varies from 0.5 to 1 (critical depensation, Figure 7) or from 1 to 3 (non-critical depensation and compensation, Figure 8).

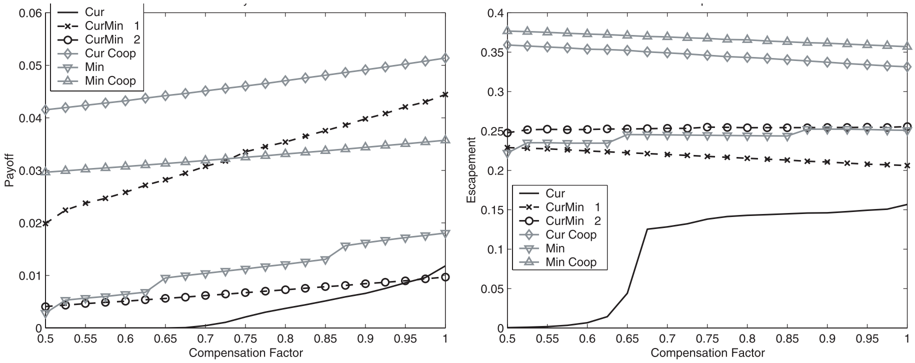


Figure 7: Harvesting at various degrees of depensation.

The graphs displayed show game outcomes for five types of games: competitive with complete symmetric information ("Cur"), competitive with asymmetric (current vs. minimal) information ("Cur-Min 1" and "Cur-Min 2"), cooperative with current information ("Cur Coop"), competitive with minimal information ("Min"), and finally cooperative with minimal information ("Min Coop").

Since with the increase of A the cubic growth function uniformly increases it is natural that payoffs also increase for all types of knowledge. However, the rate of growth of payoff for different games is quite different. It is seen that while for low A additional information leads to very low payoffs (and even zero payoff for complete current symmetric knowledge "Cur"), at a high compensation factor it becomes highly beneficial.

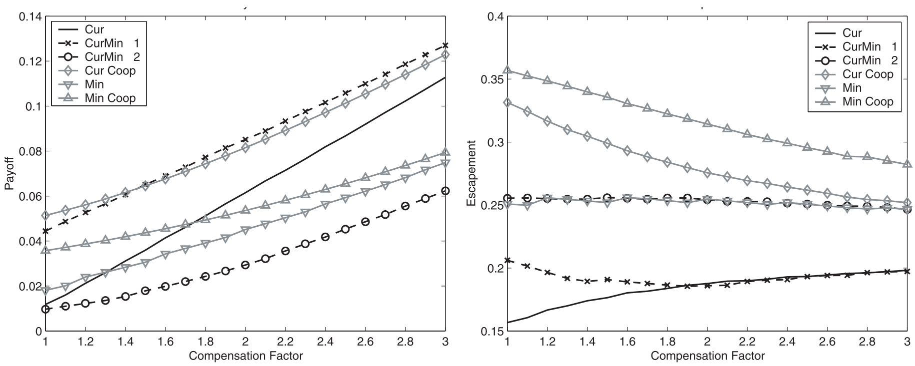


Figure 8: Harvesting at various degrees of compensation.

2.3.4 Balancing Asymmetric Information against Asymmetric Environmental Conditions

Here nature slightly favors the second player (Figure 9). Specifically, θ takes the values 0.1 and 0.8 with equal probabilities (so the first player's fraction of total recruitment fluctuates between 0.1 or 0.8 of the whole recruitment, while the second player receives the fraction 0.9 or 0.2). Thus, the mean recruitment for the first player is lower than for the second one.

The graphs show game outcomes for three types of games: competitive with imperfect measurement information ("Meas 1" and "Meas 2"), competitive with asymmetric (measurement vs. minimal) information ("Meas-Min 1" and "Meas-Min 2"), and finally cooperative with measurement information without side payments ("Coop 1" and "Coop 2") and with equal sharing ("Coop").

As one would expect, when the players possess identical information, player 2 ("Meas 1") will always do better than player 1 ("Meas 2"). But when player 1 ("Meas-Min 1") has a strong informational advantage (here when MP > 0.3), this can overbalance player 2's ("Meas-Min 2") environmental advantages.

On the other hand, the sum of the players' payoffs will be greatest with cooperation: players sharing information, and with the common objective of maximizing the sum of their returns. In this case, and because of the environmental asymmetry, the direct harvest returns in the two substreams ("Coop 1" and "Coop 1") will not be equal. This can be considered as a pure "goodwill" cooperation. Alternatively, the two players can negotiate a different split of this total return, an outcome achievable through negotiation of a compensating "side-payment," from one player to the other. The case of equal sharing ("Coop") is shown on the graph as well.

We will get qualitatively similar results if we consider an asymmetry, not in nature but in the economic environment. Figure 10 corresponds to the case where the only asymmetry is cost of harvest (see Section 4.2). It is equal to

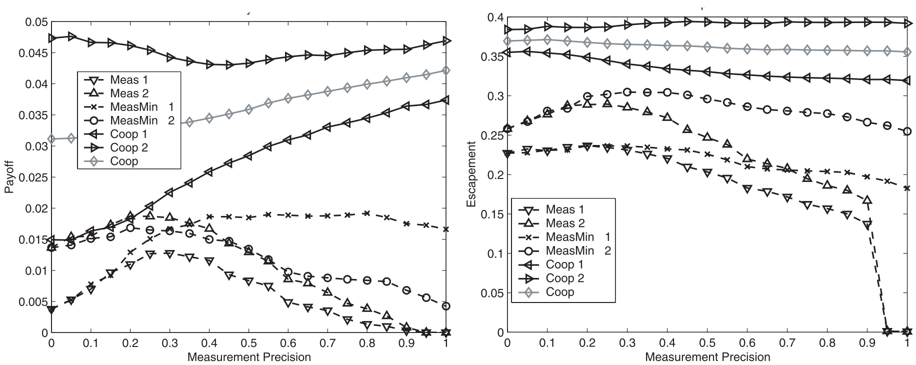


Figure 9: Influence of measurement precision in asymmetric natural environment case.

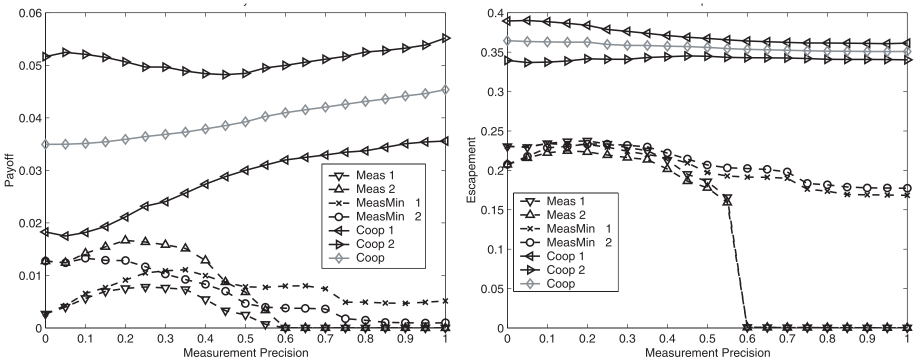


Figure 10: Influence of measurement precision in asymmetric economic environment case.

0.2 for the first player and 0.15 for the second one. Thus, economic asymmetry again slightly favors the second player.

2.3.5 The Influence of Environmental Variability

In this set of simulations (Figure 11) we change the variance of the stock-split factor θ. Specifically, θ randomly takes two values: θ_1 and $\theta_2 = 1 - \theta_1$, where θ_1 may be any fraction between 0 and 0.5. So, the standard deviation s_θ of θ may take any value between 0 and $s_{\theta\,\max} = 0.5$. We define "variability" of θ as variability $= s_\theta / s_{\theta\,\max}$. Thus, for $\theta_1 = 0.5$ there is no variability ($\theta = 0.5$ always, variability $= 0$), while for $\theta_1 = 0$ variability is highest (θ jumps randomly between 0 and 1, variability $= 1$). All the simulations are performed for the same depensatory growth function (specified by $A = 0.6$, $b = 0$, see Figure 15 top).

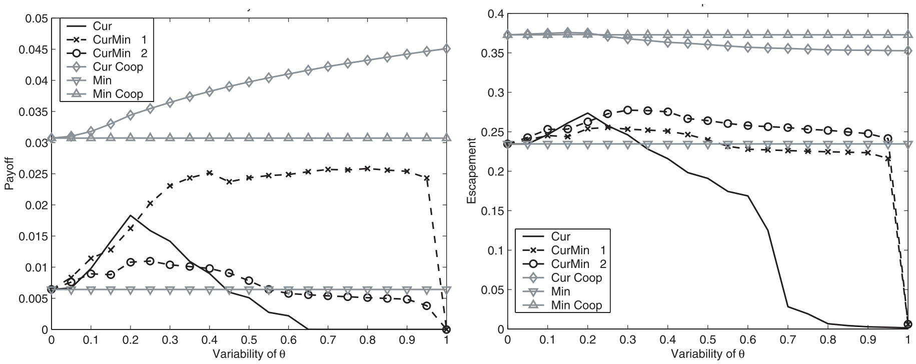


Figure 11: Influence of the level of environmental variability.

The graphs show game outcomes for five types of games: competitive with complete symmetric information ("Cur"), competitive with asymmetric (current vs. minimal) information ("Cur-Min 1" and "Cur-Min 2"), cooperative with current information ("Cur Coop"), competitive with minimal information ("Min"), and cooperative with minimal information ("Min Coop").

For cooperative harvesting with complete knowledge ("Cur Coop"), the increase of the payoff with an increase of θ variability is quite natural. Indeed, with high variability of θ almost the entire fish stock goes into just one of two streams, and this leads to reduction of harvesting cost per unit (cf. Section 4.2). Because this cooperative game is fully symmetric, the average annual payoffs to the two players will be identical.

It appears that the increase of payoff in competitive games with an increase of variability from 0 to 0.2 may have the same explanation. However, at higher variability values the effect of competition (especially for complete knowledge, "Cur") becomes dominant. Indeed, if all the stock is in one stream, the corresponding fleet can harvest almost all of it at relatively low cost.

2.3.6 From Competition to Cooperation

In Section 3.4 we describe a simple way to introduce "cooperation" into the competitive harvesting game. In Figure 12 we can see that the average immediate payoff constantly increases with the increase of "degree of cooperation" for both complete ("Cur") and minimum ("Min") types of knowledge.

Here a zero degree of cooperation means purely competitive behavior, with each player maximizing his own discounted payoff, while degree 1 cooperation means both players maximize the total discounted payoff. For intermediate values of cooperation, each player will maximize a convex linear combination of his own and his competitor's discounted payoffs.

Simulation is performed for a depensatory (cubic, see Section 4.3) growth function ($A = 0.6$, $b = 0$, see Figure 15 top). It is clearly seen on Figure 12 that

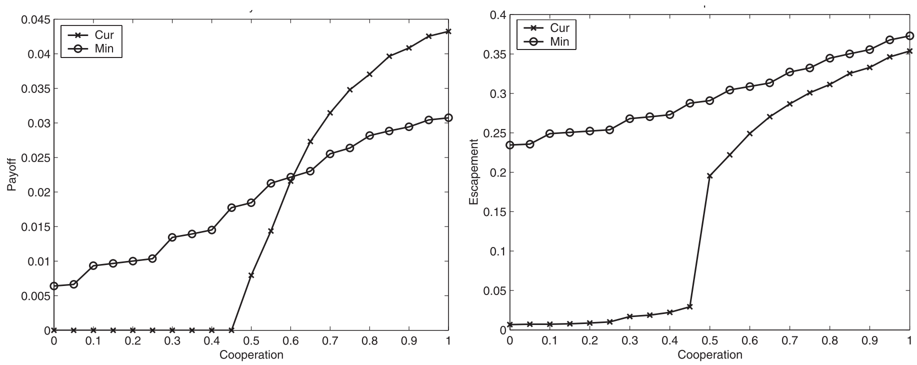


Figure 12: Increasing degree of cooperation. Current vs. minimum knowledge.

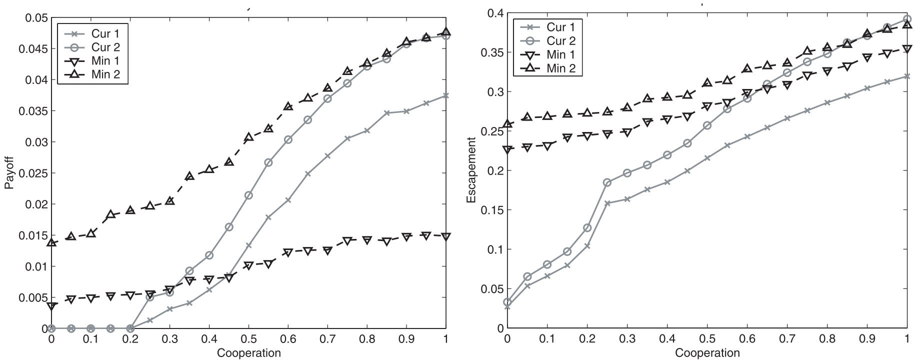


Figure 13: Increasing degree of cooperation in asymmetric environment case. Current and minimum knowledge. Critical depensation case.

additional information ("Cur") is beneficial only when there is a high degree of cooperation. At low degrees of cooperation, and especially in the case of pure competition, additional knowledge leads to a critical drop of average escapement and to zero average payoff.

Of course, total cooperation implies sharing all private knowledge, as well as taking into account both players' conflicting objectives.

Figure 13 displays a case of asymmetric nature conditions, where players get different average payoffs. Here, θ takes the values 0.1 and 0.8 with equal probabilities, so nature slightly favors the second player.

Note that, at all levels of cooperation, the difference between average payoffs is much higher for minimal ("Min 1" and "Min 2") than for complete ("Cur 1" and "Cur 2") knowledge. In this simulation (critical depensation, $A = 0.6$) minimum knowledge is always more beneficial for the second player ("Min 2") than complete knowledge ("Cur 2"). However, this is not invariably true: In simulations corresponding to a higher compensation factor A (see Figure 14 for $A = 3$), the curve "Cur 2" (and even "Cur 1") rises above "Min 2."

2.4 Summary of simulation results

The simulations demonstrate the trade-offs that harvesters face when there is a potential for gain from risk-taking, but under circumstances of limited information and destructive competition.

One such situation arises when a low harvesting cost index c encourages large harvests, but depensatory biological growth carries a risk that any overharvest may trap the stock permanently in the range of lowest biological productivity and highest harvesting costs.

Supposing full symmetry between players, it is found (e.g., Figures 2–8) that cooperative management is able to take advantage of low c to achieve high returns, and that the returns improve significantly with enhanced knowledge of

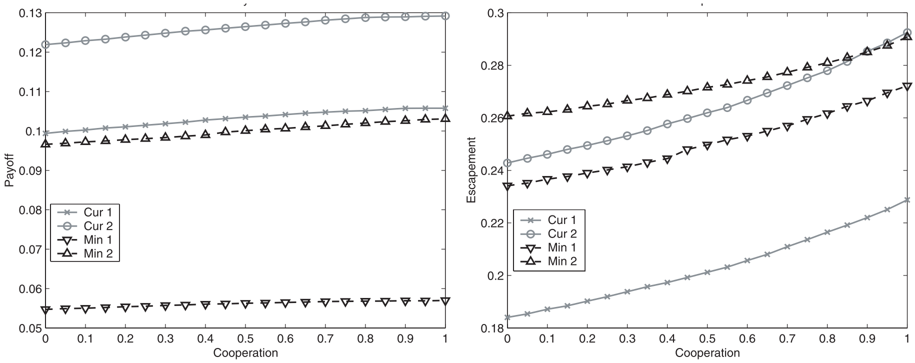


Figure 14: Increasing degree of cooperation in asymmetric environment case. Current and minimum knowledge. Compensatory growth function.

the current stock-split fraction θ. But a competitive harvest, under those same conditions of low c and high information precision, will result in a depleted stock and much lower returns—lower even than when information precision is minimal (e.g., Figure 6 (also Figures 2, 3, 5)). Indeed, the best returns under competition are achieved when both the cost index and the precision of information are in the midrange.

The competitive results under information asymmetry also are interesting. Typically, the player with the greater information precision will do better than his opponent, by strategically increasing his harvest modestly without unduly reducing mean subsequent recruitment. It will seldom be advantageous for him to reveal his private information to his opponent. In most circumstances his mean payoff would decrease, should he reveal private information—only mildly when the growth function is compensatory (Figure 4), but quite significantly in the presence of depensation (Figures 5 and 6).

In this asymmetric information case the player with less information typically will gain either if he knows more or his opponent knows less; that is, by a return to the circumstance of symmetric information. However, in rare cases (see Figure 11, $0.225 < \theta < 0.3$) he may prefer unfavorable knowledge asymmetry to either of the two symmetric alternatives.

Another case where added risk carries a potential for higher returns occurs when the stream-distribution factor θ has high variance, as illustrated in Figure 11. Indeed, with cooperation and accurate tracking of θ, the harvest in a given year can then be concentrated within the stream with higher recruitment, and hence the lower costs. And the higher the variance of θ, the higher will be the return (Figure 11).

But with symmetric access, the high precision of knowledge of total recruitment size intensifies destructive competition, and mean returns decline. Analogously to the critical depensation-low c case, payoffs peak with mid-level environmental variance, and at high variance are higher for minimal information than for full current information of θ.

In asymmetric information cases, the player with greater information is much better off, except at very high levels of fluctuation, to withhold his private information, while the player with less information would gain from additional information for small θ-variance, but would lose thereby when variance is large.

3 General Study of Dynamical Fishery Games

3.1 Games with Common Knowledge of Recruitment

The primary focus of this chapter is on dynamic competitive harvesting games where, prior to each harvest season, both players are informed of the current joint recruitment R, as well as some, generally incomplete, information concerning current random system parameters.

We assume that the subsequent season's recruitment R_+ and the players' current seasonal payoffs v^α and v^β all depend on initial recruitment R, on players' policies p^α, p^β, and on random factors ν, i.e.,

$$R_+ = \rho(R, \nu, p^\alpha, p^\beta),$$

$$v^\alpha = v^\alpha(R, \nu, p^\alpha, p^\beta), \qquad v^\beta = v^\beta(R, \nu, p^\alpha, p^\beta).$$

Furthermore, players' policies depend on their common recruitment R and some additional information ξ^α and ξ^β, about random disturbances ν:

$$p^\alpha = P^\alpha(R, \xi^\alpha), \qquad p^\beta = P^\beta(R, \xi^\beta).$$

Thus, at this level of generality the transition of the system from one harvesting stage to the next may be represented schematically as

$$R \xrightarrow{\qquad\qquad \rho(R, \nu, p^\alpha, p^\beta) \qquad\qquad} R_+$$

Many specific structural patterns of fishery can be represented in this way. Here we indicate only a couple of them.

3.1.1 Split Stream Harvesting

The simulations described in Section 2.3 all concern split stream harvesting.

It is easily seen that such split stream harvesting (for the case when players know their common recruitment) can be considered as a particular case of a common-stock harvesting. Indeed, recalling the notation of Section 2.1, in the split stream case $\nu = (\theta, \varphi)$, the function ρ has a form:

$$\rho(R; \theta, \varphi; p^\alpha, p^\beta) = F\left((\theta p^\alpha + (1-\theta)p^\beta)R, \varphi\right)$$

and the functions v^α and v^β are expressed through the corresponding functions $v^\alpha_{\rm spl}$ and $v^\beta_{\rm spl}$ (for the split case) as follows:

$$v^\alpha(R;\theta,\varphi;p^\alpha,p^\beta) = v^\alpha_{\rm spl}(\theta R, p^\alpha), \qquad v^\beta(R;\theta,\varphi;p^\alpha,p^\beta) = v^\beta_{\rm spl}((1-\theta)R, p^\beta).$$

In what follows we assume that random vectors θ_t constitute a Markov chain with known stochastic properties. In particular, if θ_t takes a finite number n of states, its stochastic behavior is completely determined by a certain stochastic $n \times n$ matrix.

3.1.2 Sequential Harvesting

In this case the player α harvests first, then the player β harvests, and then the fish spawns.

$$R = R^\alpha \xrightarrow{\sigma^\alpha(R^\alpha,\nu^\alpha,p^\alpha)} S^\alpha = R^\beta \xrightarrow{\sigma^\beta(R^\beta,\nu^\beta,p^\beta)} S^\beta = S \xrightarrow{F(S,\varphi)} R_+$$

Here each player's escapement depends on his recruitment, his policy, and some random factors at this step,

$$S^\alpha = \sigma^\alpha(R^\alpha,\nu^\alpha,p^\alpha), \qquad S^\beta = \sigma^\beta(R^\beta,\nu^\beta,p^\beta),$$

and it is natural to assume that each player knows his respective recruitment R^α or R^β and has some knowledge ξ^α or ξ^β of the stochastic state parameters ν^α or ν^β, respectively.

Then his policy will be of the form, e.g.,

$$p^\alpha = P^\alpha(R^\alpha,\xi^\alpha), \qquad p^\beta = P^\beta(R^\beta,\xi^\beta),$$

and

$$R_+ = F\Big(\sigma^\beta\left(\sigma^\alpha(R,\nu^\alpha,p^\alpha),\nu^\beta,p^\beta\right),\varphi\Big).$$

In general, this sequential model cannot be thought of as a special case of our general formulation, unless the second player is assumed to also know R. But, if such additional knowledge (and some information about random factors ν^α and ν^β) is assumed, then the sequential harvesting case can be subsumed within the general formulation, with

$$R_+ = \rho(R;\nu^\alpha,\nu^\beta,\varphi;p^\alpha,p^\beta) = F\Big(\sigma^\beta\left(\sigma^\alpha(R,\nu^\alpha,p^\alpha),\nu^\beta,p^\beta\right),\varphi\Big).$$

3.2 Finite Horizon Problem with Simple Types of Knowledge

Let us consider the finite horizon game with the final season T. Here we will not restrict ourselves to stationarity, so the functions v_τ^α and v_τ^β may be different at different moments τ. Denote a decision policy at time τ by P_τ^α and a sequence of decision functions P_τ^α from the moment t until T by

$$\mathbf{P}_t^\alpha = \langle P_t^\alpha, P_{t+1}^\alpha, \ldots, P_T^\alpha \rangle = \langle P_t^\alpha, \mathbf{P}_{t+1}^\alpha \rangle.$$

The knowledge that each player has at time t is the current common recruitment R_t and some information about stochastic parameter ν. Knowledge of ν includes, at minimum, the probability transition matrix for this Markovian random sequence. In this section we consider the simplest cases: a player may have only this minimum knowledge ν or, in addition, may, at time t, know historical values of ν up through the previous season's value ν_{t-1}, or even also know the current value of ν, i.e., ν_t. The knowledge state may be symmetric, with both players having the same information or may be asymmetric, e.g., one player having current knowledge of ν while the other has only delayed knowledge.

3.2.1 Current Information

Assume that the information held by both players at time t includes the current value of the random vector ν_t.

Then, at the moment t, the expected discounted payoff for the player α is

$$V_t^\alpha(R_t, \nu_t, \mathbf{P}_t^\alpha, \mathbf{P}_t^\beta) = \mathsf{E}_{(\nu_{t+1}, \nu_{t+2}, \ldots, \nu_T | \nu_t)} \sum_{\tau=t}^{T} \gamma_\alpha^{\tau-t} v_\tau^\alpha(R_\tau, \nu_\tau, p_\tau^\alpha, p_\tau^\beta) \tag{1}$$

with a similar expression for player β, where

$$p_\tau^\alpha = P_\tau^\alpha(R_\tau, \nu_\tau), \qquad p_\tau^\beta = P_\tau^\beta(R_\tau, \nu_\tau).$$

A pair $\langle \widehat{\mathbf{P}}_t^\alpha, \widehat{\mathbf{P}}_t^\beta \rangle$ provides a Nash equilibrium for a pair $\langle V_t^\alpha, V_t^\beta \rangle$ if for all possible values of R_t and ν_t

$$\begin{cases} V_t^\alpha(R_t, \nu_t, \widehat{\mathbf{P}}_t^\alpha, \widehat{\mathbf{P}}_t^\beta) &= \max\limits_{\mathbf{P}_t^\alpha} V_t^\alpha(R_t, \nu_t, \mathbf{P}_t^\alpha, \widehat{\mathbf{P}}_t^\beta), \\ V_t^\beta(R_t, \nu_t, \widehat{\mathbf{P}}_t^\alpha, \widehat{\mathbf{P}}_t^\beta) &= \max\limits_{\mathbf{P}_t^\beta} V_t^\beta(R_t, \nu_t, \widehat{\mathbf{P}}_t^\alpha, \mathbf{P}_t^\beta). \end{cases}$$

In what follows we will denote the corresponding Nash equilibrium discounted payoffs as

$$\begin{cases} \widehat{V}_t^\alpha(R_t, \nu_t) &= V_t^\alpha(R_t, \nu_t, \widehat{\mathbf{P}}_t^\alpha, \widehat{\mathbf{P}}_t^\beta), \\ \widehat{V}_t^\beta(R_t, \nu_t) &= V_t^\beta(R_t, \nu_t, \widehat{\mathbf{P}}_t^\alpha, \widehat{\mathbf{P}}_t^\beta). \end{cases}$$

Since the random parameter ν is a Markov process, completely determined by its current value and single-stage transition distribution (i.e., the distribution of ν_{t+1} for any given ν_t), the mathematical expectation $\mathsf{E}_{(\nu_{t+1},\nu_{t+2},\ldots,\nu_T|\nu_t)}$ can be presented as a sequence of conditional expectations:

$$\mathsf{E}_{(\nu_{t+1},\nu_{t+2},\ldots,\nu_T|\nu_t)} = \mathsf{E}_{(\nu_{t+1}|\nu_t)}\mathsf{E}_{(\nu_{t+2}|\nu_{t+1})}\cdots\mathsf{E}_{(\nu_T|\nu_{T-1})}.$$

It follows that V_t^α can be expressed through immediate payoff v_t^α and V_{t+1}^α:

$$\begin{aligned} V_t^\alpha(R_t,\nu_t,\mathbf{P}_t^\alpha,\mathbf{P}_t^\beta) &= v_t^\alpha(R_t,\nu_t,p_t^\alpha,p_t^\beta) \\ &\quad + \gamma_\alpha \mathsf{E}_{(\nu_{t+1}|\nu_t)} V_{t+1}^\alpha\Big(\rho(R_t,\nu_t,p_t^\alpha,p_t^\beta),\nu_{t+1},\mathbf{P}_{t+1}^\alpha,\mathbf{P}_{t+1}^\beta\Big). \end{aligned} \tag{2}$$

Note that we can also use expression (2) as an alternative (recursive) definition of the discounted payoff.

This leads to the following dynamic programming solution for this problem. Suppose that $\langle \widehat{\mathbf{P}}_{t+1}^\alpha, \widehat{\mathbf{P}}_{t+1}^\beta\rangle$ are Nash equilibrium policies starting from the moment $t+1$ and $\langle \widehat{V}_{t+1}^\alpha, \widehat{V}_{t+1}^\beta\rangle$ are the corresponding optimal discounted values. Define

$$\widetilde{V}_t^\alpha(R_t,\nu_t,P_t^\alpha,P_t^\beta) = \widetilde{V}_t^\alpha(R_t,\nu_t,\langle P_t^\alpha,\widehat{\mathbf{P}}_{t+1}^\alpha\rangle,\langle P_t^\beta,\widehat{\mathbf{P}}_{t+1}^\beta\rangle),$$

i.e., the discounted payoff corresponding to arbitrary policies P_t^α and P_t^β at time t and optimal "tails" $\widehat{\mathbf{P}}_t^\alpha$ and $\widehat{\mathbf{P}}_t^\beta$, and a similar function for the player β.

Then the optimal policies $\langle \widehat{P}_t^\alpha, \widehat{P}_t^\beta\rangle$ for the time t can be obtained by solving, for all possible values of R_t and ν_t, the Nash equilibrium problem for the functions

$$\begin{cases} \widetilde{V}_t^\alpha(R_t,\nu_t,p_t^\alpha,p_t^\beta) = v_t^\alpha(R_t,\nu_t,p_t^\alpha,p_t^\beta) \\ \qquad\qquad + \gamma_\alpha \mathsf{E}_{(\nu_{t+1}|\nu_t)}\widehat{V}_{t+1}^\alpha(\rho(R_t,p_t^\alpha,p_t^\beta,\nu_t),\nu_{t+1}), \\ \widetilde{V}_t^\beta(R_t,\nu_t,p_t^\alpha,p_t^\beta) = v_t^\beta(R_t,\nu_t,p_t^\alpha,p_t^\beta) \\ \qquad\qquad + \gamma_\beta \mathsf{E}_{(\nu_{t+1}|\nu_t)}\widehat{V}_{t+1}^\beta(\rho(R_t,p_t^\alpha,p_t^\beta,\nu_t),\nu_{t+1}) \end{cases}$$

with respect to $\langle p_t^\alpha, p_t^\beta\rangle$. Specifically, $\widehat{P}_t^\alpha(R_t,\nu_t) = \hat{p}_t^\alpha$ and $\widehat{P}_t^\beta(R_t,\nu_t) = \hat{p}_t^\beta$, where the pair $\langle \hat{p}_t^\alpha, \hat{p}_t^\beta\rangle$ attains Nash equilibrium to these functions for given R_t and ν_t. Thus, the Nash equilibrium policies starting from the moment t and $\langle \widehat{V}_t^\alpha, \widehat{V}_t^\beta\rangle$ can be obtained recursively as

$$\widehat{\mathbf{P}}_t^\alpha = \langle \widehat{P}_t^\alpha, \widehat{\mathbf{P}}_{t+1}^\alpha\rangle, \qquad \widehat{\mathbf{P}}_t^\beta = \langle \widehat{P}_t^\beta, \widehat{\mathbf{P}}_{t+1}^\beta\rangle.$$

3.2.2 Delayed Information

This case looks very similar to the case of current knowledge except for the fact that the policies and discounted values at time t depend not on ν_t but on ν_{t-1}, i.e.,

$$p_t^\alpha = P_t^\alpha(R_t,\nu_{t-1}), \qquad p_t^\beta = P_t^\beta(R_t,\nu_{t-1}),$$

and

$$\begin{aligned}
V_t^\alpha(R_t,\nu_{t-1},\mathbf{P}_t^\alpha,\mathbf{P}_t^\beta) &= \mathsf{E}_{(\nu_t,\nu_{t+1},\ldots,\nu_T|\nu_{t-1})}\sum_{\tau=t}^{T}\gamma_\alpha^{\tau-t}v_\tau^\alpha(R_\tau,\nu_\tau,p_\tau^\alpha,p_\tau^\beta)\\
&= \mathsf{E}_{(\nu_t|\nu_{t-1})}\mathsf{E}_{(\nu_{t+1},\ldots,\nu_T|\nu_t)}\sum_{\tau=t}^{T}\gamma_\alpha^{\tau-t}v_\tau^\alpha(R_\tau,\nu_\tau,p_\tau^\alpha,p_\tau^\beta)\\
&= \mathsf{E}_{(\nu_t|\nu_{t-1})}\dot{V}_t^\alpha(R_t,\nu_t,\mathbf{P}_t^\alpha,\mathbf{P}_t^\beta),
\end{aligned}$$

where $\dot{V}_t^\alpha$ denotes the discounted value function for the "current information" case.

Thus, the case when the both players have delayed information leads to the dynamic programming procedure with the following Nash equilibrium problem at each step:

$$\begin{cases}
\widetilde{V}_t^\alpha(R_t,\nu_{t-1},p_t^\alpha,p_t^\beta) = \mathsf{E}_{(\nu_t|\nu_{t-1})}\Big[v_t^\alpha(R_t,\nu_t,p_t^\alpha,p_t^\beta)\\
\qquad\qquad + \gamma_\alpha\widehat{V}_{t+1}^\alpha(\rho(R_t,p_t^\alpha,p_t^\beta,\nu_t),\nu_t)\Big],\\
\widetilde{V}_t^\beta(R_t,\nu_{t-1},p_t^\alpha,p_t^\beta) = \mathsf{E}_{(\nu_t|\nu_{t-1})}\Big[v_t^\beta(R_t,\nu_t,p_t^\alpha,p_t^\beta)\\
\qquad\qquad + \gamma_\beta\widehat{V}_{t+1}^\beta(\rho(R_t,p_t^\alpha,p_t^\beta,\nu_t),\nu_t)\Big],
\end{cases}$$

where $\widehat{V}_t^\alpha(R_t,\nu_{t-1})$ and $\widehat{V}_t^\beta(R_t,\nu_{t-1})$ are Nash equilibrium values for these functions. Optimal decision policies $\widehat{P}_t^\alpha$ and $\widehat{P}_t^\beta$ are defined as $\widehat{P}_t^\alpha(R_t,\nu_{t-1})=\hat{p}_t^\alpha$ and $\widehat{P}_t^\beta(R_t,\nu_{t-1})=\hat{p}_t^\beta$, where the pair $\langle\hat{p}_t^\alpha,\hat{p}_t^\beta\rangle$ attains Nash equilibrium to these functions for given R_t and ν_{t-1}.

3.2.3 Asymmetric Information: Current vs. Delayed

Now suppose that the first player has current knowledge of ν and the second one has delayed information. Thus, the first player's policy depends on ν_t and ν_{t-1}, while the second player's policy depends only on ν_{t-1}, i.e.,

$$p_t^\alpha = P_t^\alpha(R_t,\nu_{t-1},\nu_t),\qquad p_t^\beta = P_t^\beta(R_t,\nu_{t-1}).$$

In this case at each dynamic programming (DP) step we have an asymmetric Nash equilibrium problem:

$$\begin{cases}
\widetilde{V}_t^\alpha(R_t,\nu_{t-1},\nu_t,p_t^\alpha,p_t^\beta) = v_t^\alpha(R_t,\nu_t,p_t^\alpha,p_t^\beta)\\
\qquad\qquad + \gamma_\alpha\mathsf{E}_{(\nu_{t+1}|\nu_t)}\widehat{V}_{t+1}^\alpha(\rho(R_t,p_t^\alpha,p_t^\beta,\nu_t),\nu_{t+1}),\\
\widetilde{V}_t^\beta(R_t,\nu_{t-1},p_t^\alpha,p_t^\beta) = \mathsf{E}_{(\nu_t|\nu_{t-1})}\Big[v_t^\beta(R_t,\nu_t,p_t^\alpha,p_t^\beta)\\
\qquad\qquad + \gamma_\beta\widehat{V}_{t+1}^\beta(\rho(R_t,p_t^\alpha,p_t^\beta,\nu_t),\nu_t)\Big].
\end{cases}$$

Here the first player utilizes the knowledge of ν_{t-1} to compute the second player's policy at time t. However, while the second player can calculate the first player's policy, he cannot know his opponent's actual response (since he does not know ν_t). Instead he can only assign to it a probability distribution.

Strictly speaking, the solution of this Nash equilibrium problem is a pair of functions $\widehat{P}_t^{\alpha}(R_t, \nu_{t-1}, \nu_t)$ and $\widehat{P}_t^{\beta}(R_t, \nu_{t-1})$ among all the functions $P_t^{\alpha}(R_t, \nu_{t-1}, \nu_t)$ and $P_t^{\beta}(R_t, \nu_{t-1})$ that attain Nash equilibrium to the following pair of functions for all the values R_t, ν_{t-1}, and ν_t:

$$\begin{cases} \widetilde{V}_t^{\alpha}(R_t, \nu_{t-1}, \nu_t, P_t^{\alpha}, P_t^{\beta}) & = & v_t^{\alpha}(R_t, \nu_t, P_t^{\alpha}(R_t, \nu_{t-1}, \nu_t), P_t^{\beta}(R_t, \nu_{t-1})) \\ & & + \gamma_{\alpha} \mathsf{E}_{(\nu_{t+1}|\nu_t)} \widehat{V}_{t+1}^{\alpha} \\ & & \Big(\rho(R_t, \nu_t, P_t^{\alpha}(R_t, \nu_{t-1}, \nu_t), \\ & & P_t^{\beta}(R_t, \nu_{t-1})), \nu_{t+1}\Big), \\ \widetilde{V}_t^{\beta}(R_t, \nu_{t-1}, P_t^{\alpha}, P_t^{\beta}) & = & \mathsf{E}_{(\nu_t|\nu_{t-1})}\Big[v_t^{\beta}\Big(R_t, \nu_t, P_t^{\alpha}(R_t, \nu_{t-1}, \nu_t), \\ & & P_t^{\beta}(R_t, \nu_{t-1})\Big) + \gamma_{\beta} \widehat{V}_{t+1}^{\beta} \\ & & \Big(\rho(R_t, \nu_t, P_t^{\alpha}(R_t, \nu_{t-1}, \nu_t), \\ & & P_t^{\beta}(R_t, \nu_{t-1})), \nu_t\Big)\Big]. \end{cases}$$

Note, that each player can adjust his policy "pointwise" i.e., for all the possible values of his policy arguments, provided the other player's policy is fixed. This, in fact, can be used for computing optimal policies iteratively. For some α-policy we can find an optimum response β-policy. Then we fix this β-policy and find the corresponding optimum α-policy, and so on.

3.2.4 Minimal Knowledge

Let us consider the situation where the players know nothing of the specific realization of ν, but know only its transition probability matrix, and hence its stationary asymptotic distribution. In this case policies and discounted values depend on R but do not depend on the realization of ν, i.e.

$$V_t^{\alpha}(R_t, \mathbf{P}_t^{\alpha}, \mathbf{P}_t^{\beta}) = \mathsf{E}_{\nu_t} \mathsf{E}_{(\nu_{t+1}, \nu_{t+2}, \ldots, \nu_T|\nu_t)} \sum_{\tau=t}^{T} \gamma_{\alpha}^{\tau-t} v_{\tau}^{\alpha}(R_{\tau}, \nu_{\tau}, p_{\tau}^{\alpha}, p_{\tau}^{\beta})$$

$$= \mathsf{E}_{\nu_t} \dot{V}_t^{\alpha}(R_t, \nu_t, \mathbf{P}_t^{\alpha}, \mathbf{P}_t^{\beta}). \tag{3}$$

Here, the first (unconditional) expectation E_{ν_t} is taken over the "known" distribution of ν, i.e., its limit distribution. The limit distribution can be obtained by taking the infinite power of the stochastic matrix that describes Markov process. The infinite power of the stochastic matrix has (under very general conditions) equal columns representing the limit (stationary) distribution of ν. Let us denote ν^* a random variable that has this limit distribution.

Also, $\dot{V}_t^\alpha$ denotes the discounted value function for the complete information case. Finally,

$$p_\tau^\alpha = P_\tau^\alpha(R_\tau), \qquad p_\tau^\beta = P_\tau^\beta(R_\tau).$$

So, in the case where both players possess only minimal information about ν, we can obtain optimal policies $\mathbf{P}_t^\alpha(R_t)$ and $\mathbf{P}_t^\beta(R_t)$ by finding the Nash equilibrium for the functions

$$\begin{cases} \widetilde{V}_t^\alpha(R_t, p_t^\alpha, p_t^\beta) &= \mathsf{E}_{\nu^*}\left[v_t^\alpha(R_t, \nu^*, p_t^\alpha, p_t^\beta) + \gamma_\alpha \widehat{V}_{t+1}^\alpha(\rho(R_t, \nu^*, p_t^\alpha, p_t^\beta))\right] \\ \widetilde{V}_t^\beta(R_t, p_t^\alpha, p_t^\beta) &= \mathsf{E}_{\nu^*}\left[v_t^\beta(R_t, \nu^*, p_t^\alpha, p_t^\beta) + \gamma_\beta \widehat{V}_{t+1}^\beta(\rho(R_t, \nu^*, p_t^\alpha, p_t^\beta))\right] \end{cases}$$

with respect to $\langle p_t^\alpha, p_t^\beta \rangle$. Now $\widehat{P}_t^\alpha(R_t) = \hat{p}_t^\alpha$ and $\widehat{P}_t^\beta(R_t) = \hat{p}_t^\beta$, where the pair $\langle \hat{p}_t^\alpha, \hat{p}_t^\beta \rangle$ attains Nash equilibrium to these functions for given R_t, and $\langle \widehat{V}_{t+1}^\alpha, \widehat{V}_{t+1}^\beta \rangle$ are again the corresponding optimal discounted values.

3.2.5 Asymmetric Cases

As in Section 3.2.3 we can study other asymmetric situations. For example, suppose that the first player has current information ν_t and his discounted payoff is determined by (1) and the second player does not have any information and his discounted value is given by (3). Now

$$p_t^\alpha = P_t^\alpha(R_t, \nu_t), \qquad p_t^\beta = P_t^\beta(R_t)$$

and these policies can be found recursively by solving the following sequence of Nash equilibrium problems:

$$\begin{cases} \widetilde{V}_t^\alpha(R_t, \nu_t, P_t^\alpha, P_t^\beta) &= v_t^\alpha(R_t, \nu_t, P_t^\alpha(R_t, \nu_t), P_t^\beta(R_t)) \\ & \quad + \gamma_\alpha \mathsf{E}_{(\nu_{t+1}|\nu_t)} \widehat{V}_{t+1}^\alpha(\rho(R_t, \nu_t, P_t^\alpha(R_t, \nu_t), \\ & \quad P_t^\beta(R_t)), \nu_{t+1}), \\ \widetilde{V}_t^\beta(R_t, P_t^\alpha, P_t^\beta) &= \mathsf{E}_{(\nu_t)}\Big[v_t^\beta(R_t, \nu_t, P_t^\alpha(R_t, \nu_t), P_t^\beta(R_t)) \\ & \quad + \gamma_\beta \widehat{V}_{t+1}^\beta(\rho(R_t, \nu_t, P_t^\alpha(R_t, \nu_t), P_t^\beta(R_t)))\Big]. \end{cases}$$

3.3 Information Obtained from Imperfect Observation

Our assumption in Section 3.2 that the players possess precise knowledge of the realization of a stochastic parameter ν_t clearly is an idealization. Typically, its value cannot be determined precisely but only with a certain error. In this section we will introduce the notion of measurement error in observation of stochastic parameters. This leads to a certain generalization of our previous "incomplete information" case and introduces a continuum in the precision of knowledge.

An imperfect observation of ν_t can be characterized through a transition probability from the space of states of the parameter ν_t to the space of states of the observation ξ_t. In a case where these spaces are finite, say the number of states for ν_t and ξ_t are n and m, respectively, then a measurement is completely determined by an $m \times n$ stochastic matrix. The i-th column of this matrix represents the conditional distribution of the observation ξ_t when ν_t is in the ith state.

We assume that the information that a player has at time t consists of the current common recruitment R_t and the result of an imperfect measurement ξ_t of the current parameter ν_t. Different players may have results of distinct measurements ξ^α and ξ^β or these measurements may be the same. Now a player's policies at time t depend on R_t and on ξ_t^α (for player α), i.e.,

$$p_t^\alpha = P_t^\alpha(R_t, \xi_t^\alpha), \qquad p_t^\beta = P_t^\beta(R_t, \xi_t^\beta).$$

If players have the same measurement information, i.e., $\xi_t^\alpha = \xi_t^\beta$, we will denote it ξ_t. The random sequence ν may be Markovian or, as a special case, independent and identically distributed (i.i.d.).

3.3.1 Common Observations of I.I.D. Parameters

In this subsection we state the optimum harvesting problem as a problem of optimizing discounted payoffs *conditional* on measurement results ξ_t:

$$V_t^\alpha(R_t, \xi_t, \mathbf{P}_t^\alpha, \mathbf{P}_t^\beta) = \mathsf{E}_{(\nu_t, \nu_{t+1}, \xi_{t+1}, \dots | \xi_t)} \sum_{\tau=t}^{T} \gamma_\alpha^{\tau-t} v_\tau^\alpha(R_\tau, \nu_\tau, p_\tau^\alpha, p_\tau^\beta). \qquad (4)$$

Rewrite V_t^α in recursive form:

$$\begin{aligned}
V_t^\alpha(R_t, \xi_t, \mathbf{P}_t^\alpha, \mathbf{P}_t^\beta) &= \mathsf{E}_{(\nu_t|\xi_t)} \mathsf{E}_{(\nu_{t+1}, \xi_{t+1}, \dots)} \sum_{\tau=t}^{T} \gamma_\alpha^{\tau-t} v_\tau^\alpha(R_\tau, \nu_\tau, p_\tau^\alpha, p_\tau^\beta) \\
&= \mathsf{E}_{(\nu_t|\xi_t)} \Big[v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) \\
&\quad + \gamma_\alpha \mathsf{E}_{(\nu_{t+1}, \xi_{t+1})} \sum_{\tau=t+1}^{T} \gamma_\alpha^{\tau-(t+1)} v_\tau^\alpha(R_\tau, \nu_\tau, p_\tau^\alpha, p_\tau^\beta) \Big] \\
&= \mathsf{E}_{(\nu_t|\xi_t)} \Big[v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) + \gamma_\alpha \mathsf{E}_{(\xi_{t+1})} V_{t+1}^\alpha \\
&\quad \Big(\rho_t(R_t, \nu_t, p_t^\alpha, p_t^\beta), \xi_{t+1}, \mathbf{P}_{t+1}^\alpha, \mathbf{P}_{t+1}^\beta \Big) \Big].
\end{aligned}$$

Here we make use of the fact that expectation $\mathsf{E}_{(\nu_{t+1},\xi_{t+1})}$ over the joint distribution of the pair (ν_{t+1},ξ_{t+1}) can be presented as a sequence of expectation operations, i.e.,

$$\mathsf{E}_{(\nu_{t+1},\xi_{t+1})} = \mathsf{E}_{(\xi_{t+1})}\mathsf{E}_{(\nu_{t+1}|\xi_{t+1})}.$$

This leads to the following DP algorithm, which involves computation of a Nash equilibrium pair $\langle p_t^\alpha, p_t^\beta\rangle$ for the following functions:

$$\begin{cases} \widetilde{V}_t^\alpha(R_t,\xi_t,p_t^\alpha,p_t^\beta) &= \mathsf{E}_{(\nu_t|\xi_t)}\Big[v_t^\alpha\left(R_t,\nu_t,p_t^\alpha,p_t^\beta\right) \\ & \qquad + \gamma_\alpha \mathsf{E}_{(\xi_{t+1})}\widehat{V}_{t+1}^\alpha\left(\rho_t(R_t,\nu_t,p_t^\alpha,p_t^\beta),\xi_{t+1}\right)\Big] \\ \widetilde{V}_t^\beta(R_t,\xi_t,p_t^\alpha,p_t^\beta) &= \mathsf{E}_{(\nu_t|\xi_t)}\Big[v_t^\beta\left(R_t,\nu_t,p_t^\alpha,p_t^\beta\right) \\ & \qquad + \gamma_\beta \mathsf{E}_{(\xi_{t+1})}\widehat{V}_{t+1}^\beta\left(\rho_t(R_t,\nu_t,p_t^\alpha,p_t^\beta),\xi_{t+1}\right)\Big]. \end{cases}$$

Both functions $\langle P_t^\alpha, P_t^\beta\rangle$ here can be found pointwise by defining $\widehat{P}_t^\alpha(R_t,\xi_t) = \hat{p}_t^\alpha$ and $\widehat{P}_t^\beta(R_t,\xi_t) = \hat{p}_t^\beta$, where $\langle \hat{p}_t^\alpha, \hat{p}_t^\beta\rangle$ attain Nash equilibrium for the above functions.

More precisely, computation of the optimal policies $\langle \widehat{P}_t^\alpha, \widehat{P}_t^\beta\rangle$ at each step can be performed in two different ways:

(a) Fix R and ξ, find the Nash equilibrium point $\langle \hat{p}^\alpha, \hat{p}^\beta\rangle$ for the functions $\widetilde{V}_t^\alpha(R,\xi,p^\alpha,p^\beta)$ and $\widetilde{V}_t^\beta(R,\xi,p^\alpha,p^\beta)$ and set $\widehat{P}_t^\alpha(R,\xi) = \hat{p}^\alpha$ and $\widehat{P}_t^\beta(R,\xi) = \hat{p}^\beta$. Thus, the problem reduces to a pointwise computation of a Nash equilibrium for all possible values of R and ξ.

Note that the equilibrium pair $\langle \hat{p}^\alpha, \hat{p}^\beta\rangle$ can be found iteratively: For a fixed initial iteration $p_{(1)}^\alpha$ find an optimal response $p_{(1)}^\beta$, i.e.,

$$p_{(1)}^\beta = \arg\max_{p^\beta} \widetilde{V}_t^\beta(R,\xi,p_{(1)}^\alpha,p^\beta),$$

then find an optimal response $p_{(2)}^\alpha$ for $p_{(1)}^\beta$, etc. Typically, the sequence $\langle p_{(i)}^\alpha, p_{(i)}^\beta\rangle$ will converge to a Nash equilibrium point $\langle \hat{p}^\alpha, \hat{p}^\beta\rangle$.

(b) Fix some policy $P_{(1)}^\alpha$ and find an optimal response $P_{(1)}^\beta$, i.e., such a policy, that for all R,ξ the function $P_{(1)}^\beta$ attains the maximum to $\widetilde{V}_t^\beta(R,\xi,P_{(1)}^\alpha(R,\xi),P^\beta(R,\xi))$ with respect to P^β. Then similarly find $P_{(2)}^\alpha$, etc. It is obvious that at each step an optimal response can be found pointwise (separately for each combination R,ξ) for any (!) policy used by another player.

3.3.2 I.I.D. Parameters, Distinct Observations

In a more general situation players may obtain information based on *different* measurements ξ_t^α and ξ_t^β of the random parameter ν_t. Then each player's policy depends on the respective information available to him, and each player maximizes his own conditional discounted payoff:

$$V_t^\alpha(R_t, \xi_t^\alpha, \mathbf{P}_t^\alpha, \mathbf{P}_t^\beta) \;=\; \mathsf{E}_{(\nu_t, \xi_t^\beta, \ldots | \xi_t^\alpha)} \sum_{\tau=t}^{T} \gamma_\alpha^{\tau-t} v_\tau^\alpha(R_\tau, \nu_\tau, p_\tau^\alpha, p_\tau^\beta)$$

and the similar expression for $V_t^\beta(R_t, \xi_t^\beta, \mathbf{P}_t^\alpha, \mathbf{P}_t^\beta)$.

$$\begin{aligned}
V_t^\alpha(R_t, \xi_t^\alpha, \mathbf{P}_t^\alpha, \mathbf{P}_t^\beta) \;&=\; \mathsf{E}_{(\nu_t, \xi_t^\beta, \ldots | \xi_t^\alpha)} \sum_{\tau=t}^{T} \gamma_\alpha^{\tau-t} v_\tau^\alpha\Big(R_\tau, \nu_\tau, P_\tau^\alpha(R_\tau, \xi_\tau^\alpha), \\
&\quad P_\tau^\beta(R_\tau, \xi_\tau^\beta)\Big) \\
&=\; \mathsf{E}_{(\nu_t, \xi_t^\beta | \xi_t^\alpha)} \Big[v_t^\alpha \left(R_t, \nu_t, p_t^\alpha, p_t^\beta \right) + \gamma_\alpha \mathsf{E}_{\xi_{t+1}^\alpha} \\
&\quad V_t^\alpha \left(\rho_t(R_t, \nu_t, p_t^\alpha, p_t^\beta), \xi_{t+1}^\alpha, \mathbf{P}_{t+1}^\alpha, \mathbf{P}_{t+1}^\beta \right) \Big].
\end{aligned}$$

This leads to the following DP algorithm, which involves computation of the Nash equilibrium pair $\langle P_t^\alpha, P_t^\beta \rangle$ for the following functions:

$$\begin{cases}
\widetilde{V}_t^\alpha(R_t, \xi_t^\alpha, P_t^\alpha, P_t^\beta) \;=\; \mathsf{E}_{(\nu_t, \xi_t^\beta | \xi_t^\alpha)} \Big[v_t^\alpha \Big(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), P_t^\beta(R_t, \xi_t^\beta) \Big) \\
\qquad + \gamma_\alpha \mathsf{E}_{\xi_{t+1}^\alpha} \widehat{V}_{t+1}^\alpha \Big(\rho_t(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), \\
\qquad P_t^\beta(R_t, \xi_t^\beta)), \xi_{t+1}^\alpha \Big) \Big] \\
\widetilde{V}_t^\beta(R_t, \xi_t^\beta, P_t^\alpha, P_t^\beta) \;=\; \mathsf{E}_{(\nu_t, \xi_t^\alpha | \xi_t^\beta)} \Big[v_t^\beta \Big(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), P_t^\beta(R_t, \xi_t^\beta) \Big) \\
\qquad + \gamma_\beta \mathsf{E}_{\xi_{t+1}^\beta} \widehat{V}_{t+1}^\beta \Big(\rho_t(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), \\
\qquad P_t^\beta(R_t, \xi_t^\beta)), \xi_{t+1}^\beta \Big) \Big].
\end{cases}$$

3.3.3 Markov Parameter, Common Observations

$$\begin{aligned}
V_t^\alpha(R_t, \xi_t, \mathbf{P}_t^\alpha, \mathbf{P}_t^\beta) \;&=\; \mathsf{E}_{(\nu_t, \nu_{t+1}, \xi_{t+1}, \ldots | \xi_t)} \sum_{\tau=t}^{T} \gamma_\alpha^{\tau-t} v_\tau^\alpha(R_\tau, \nu_\tau, p_\tau^\alpha, p_\tau^\beta) \\
&=\; \mathsf{E}_{(\nu_t, \nu_{t+1}, \xi_{t+1}, \ldots | \xi_t)} \sum_{\tau=t}^{T} \gamma_\alpha^{\tau-t} v_\tau^\alpha(R_\tau, P_\tau^\alpha(R_\tau, \xi_\tau), \\
&\quad P_\tau^\beta(R_\tau, \xi_\tau), \nu_\tau).
\end{aligned} \tag{5}$$

This leads to the following DP algorithm, which involves computation of a Nash equilibrium pair $\langle p_t^\alpha, p_t^\beta \rangle$ for the following functions:

$$\begin{aligned}
\widetilde{V}_t^\alpha(R_t, \xi_t, p_t^\alpha, p_t^\beta) \;&=\; \mathsf{E}_{(\nu_t|\xi_t)}\Big[v_t^\alpha\Big(R_t, \nu_t, p_t^\alpha, p_t^\beta\Big) + \gamma_\alpha \mathsf{E}_{(\nu_{t+1}|\nu_t)} \\
&\qquad \mathsf{E}_{(\xi_{t+1}|\nu_{t+1})}\widehat{V}_{t+1}^\alpha\Big(\rho_t(R_t, \nu_t, p_t^\alpha, p_t^\beta), \xi_{t+1}\Big)\Big] \\
\widetilde{V}_t^\beta(R_t, \xi_t, p_t^\alpha, p_t^\beta) \;&=\; \mathsf{E}_{(\nu_t|\xi_t)}\Big[v_t^\beta\Big(R_t, \nu_t, p_t^\alpha, p_t^\beta\Big) + \gamma_\beta \mathsf{E}_{(\nu_{t+1}|\nu_t)} \\
&\qquad \mathsf{E}_{(\xi_{t+1}|\nu_{t+1})}\widehat{V}_{t+1}^\beta\Big(\rho_t(R_t, \nu_t, p_t^\alpha, p_t^\beta), \xi_{t+1}\Big)\Big].
\end{aligned} \tag{6}$$

Both functions $\langle P_t^\alpha, P_t^\beta \rangle$ here can be found pointwise by defining $\widehat{P}_t^\alpha(R_t, \xi_t) = \hat{p}_t^\alpha$ and $\widehat{P}_t^\beta(R_t, \xi_t) = \hat{p}_t^\beta$, where $\langle \hat{p}_t^\alpha, \hat{p}_t^\beta \rangle$ attain Nash equilibrium to the above functions.

Note that this section generalizes results of Section 3.2.1, in that we are here assuming imperfect current knowledge of ν. With perfect information ($\xi = \nu$) our results here reduce to those in 3.2.1. Analogously, one could analyze imperfect delayed knowledge of ν.

3.3.4 Asymmetry: Imperfect Current Observation vs. Minimal Information

Suppose that one player, say, β does not know the results of measurements of ν. Thus, his policy must be based on the asymptotic distribution of ν and does not depend on ξ. His discounted payoff is defined as an average of the payoff for the "measurement" case, i.e.,

$$\begin{aligned}
V_t^\beta(R_t, \mathbf{P}_t^\alpha, \mathbf{P}_t^\beta) \;&=\; \mathsf{E}_{(\nu_t, \xi_t, \nu_{t+1}, \xi_{t+1}, \dots)} \sum_{\tau=t}^{T} \gamma_\alpha^{\tau-t} v_\tau^\alpha(R_\tau, \nu_\tau, p_\tau^\alpha, p_\tau^\beta) \\
&=\; \mathsf{E}_{\xi_t} \mathsf{E}_{(\nu_t, \nu_{t+1}, \xi_{t+1}, \dots|\xi_t)} \sum_{\tau=t}^{T} \gamma_\alpha^{\tau-t} v_\tau^\alpha(R_\tau, \nu_\tau, p_\tau^\alpha, p_\tau^\beta) \\
&=\; \mathsf{E}_{\xi_t} \dot{V}_t^\beta(R_t, \xi_t, \mathbf{P}_t^\alpha, \mathbf{P}_t^\beta),
\end{aligned}$$

where $\dot{V}_t^\beta$ is the payoff for the "measurement" case defined as in (5). The expectation E_{ξ_t} is an expectation over a "limit" distribution of ξ_t, which can be obtained by taking the limit distribution of ν_t and applying the measurement transition distribution to it. Let us denote a random variable that has this limit distribution by ξ^*.

This leads to the DP solution for the problem with Nash equilibrium functions as in (6) and $\widetilde{V}_t^\beta(R_t, \xi_t, p_t^\alpha, p_t^\beta)$ replaced by $\mathsf{E}_{\xi^*} \widetilde{V}_t^\beta(R_t, \xi^*, p_t^\alpha, p_t^\beta)$ with policies

$$p_t^\alpha = P_t^\alpha(R_t, \xi_t), \qquad p_t^\beta = P_t^\beta(R_t).$$

3.3.5 Markov Parameter, Distinct Observations

$$
\begin{aligned}
V_t^\alpha(R_t, \xi_t^\alpha, \mathbf{P}_t^\alpha, \mathbf{P}_t^\beta) &= \mathsf{E}_{(\nu_t, \xi_t^\beta, \ldots | \xi_t^\alpha)} \sum_{\tau=t}^{T} \gamma_\alpha^{\tau-t} v_\tau^\alpha(R_\tau, \nu_\tau, p_\tau^\alpha, p_\tau^\beta) \\
&= \mathsf{E}_{(\nu_t, \xi_t^\beta, \ldots | \xi_t^\alpha)} \sum_{\tau=t}^{T} \gamma_\alpha^{\tau-t} v_\tau^\alpha\Big(R_\tau, \nu_\tau, P_\tau^\alpha(R_\tau, \xi_\tau^\alpha), \\
&\quad P_\tau^\beta(R_\tau, \xi_\tau^\beta)\Big).
\end{aligned}
$$

This leads to the following DP algorithm, which involves computation of the Nash equilibrium pair $\langle P_t^\alpha, P_t^\beta \rangle$ for the following functions:

$$
\begin{cases}
\widetilde{V}_t^\alpha(R_t, \xi_t^\alpha, P_t^\alpha, P_t^\beta) = \mathsf{E}_{(\nu_t, \xi_t^\beta | \xi_t^\alpha)} \Big[v_t^\alpha \Big(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), P_t^\beta(R_t, \xi_t^\beta) \Big) \\
\qquad \gamma_\alpha \mathsf{E}_{(\nu_{t+1}|\nu_t)} \mathsf{E}_{(\xi_{t+1}^\alpha | \nu_{t+1})} \widehat{V}_{t+1}^\alpha \\
\qquad \Big(\rho_t(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), P_t^\beta(R_t, \xi_t^\beta)), \xi_{t+1}^\alpha \Big) \Big] \\
\widetilde{V}_t^\beta(R_t, \xi_t^\beta, P_t^\alpha, P_t^\beta) = \mathsf{E}_{(\nu_t, \xi_t^\alpha | \xi_t^\beta)} \Big[v_t^\beta \Big(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), P_t^\beta(R_t, \xi_t^\beta) \Big) \\
\qquad \gamma_\beta \mathsf{E}_{(\nu_{t+1}|\nu_t)} \mathsf{E}_{(\xi_{t+1}^\beta | \nu_{t+1})} \widehat{V}_{t+1}^\beta \\
\qquad \Big(\rho_t(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), P_t^\beta(R_t, \xi_t^\beta)), \xi_{t+1}^\beta \Big) \Big].
\end{cases}
$$

3.4 Cooperation by "Taking Care of Each Other"

It is easy to introduce some sort of cooperation (or contradiction) in our model by modifying discounted payoffs V^α and V^β in a simple way, which reflects "care" of one player for the other.

Specifically, player α may take care of β by optimizing a linear combination of payoffs

$$\boldsymbol{V^\alpha} = c_{\alpha\alpha} V^\alpha + c_{\alpha\beta} V^\beta$$

instead of his original payoff V^α. Similarly, β may optimize

$$\boldsymbol{V^\beta} = c_{\beta\alpha} V^\alpha + c_{\beta\beta} V^\beta.$$

Effectively, this describes (in the case $c_{\alpha\alpha} + c_{\beta\alpha} = 1$ and $c_{\alpha\beta} + c_{\beta\beta} = 1$) a game with side payments, when one player knows that he will get a certain fraction of another player's payoff.

If $c_{\alpha\beta} > 0$, player α tries to improve the income of β and if $c_{\alpha\beta} > c_{\alpha\alpha}$, then α cares about β more than about himself. Conversely, $c_{\alpha\beta} < 0$ means that α tries to "disserve" β, possibly in order to exclude him from business.

An interesting particular case is when $c_{\alpha\alpha} = c_{\alpha\beta} = c_{\beta\alpha} = c_{\beta\beta} = 1/2$. In effect this represents a sole operator (monopolist) case.

Note that if discount factors are equal, i.e., $\gamma_\alpha = \gamma_\beta = \gamma$, then introduction of cooperation coefficients in the problem statement influences the DP solution algorithm very slightly. Specifically, the expressions like $v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta)$ are just replaced by $c_{\alpha\alpha} v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) + c_{\alpha\beta} v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta)$. For example, in the "current information" case we will have the following Nash equilibrium problem at each step:

$$\begin{cases} \widetilde{V}_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) &= c_{\alpha\alpha} v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) + c_{\alpha\beta} v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) \\ & \quad + \gamma \mathsf{E}_{(\nu_{t+1}|\nu_t)} \widehat{V}_{t+1}^\alpha(\rho(R_t, p_t^\alpha, p_t^\beta, \nu_t), \nu_{t+1}) \\ \widetilde{V}_t^\beta(R_t, \nu_t, p_t^\alpha, p_t^\beta) &= c_{\beta\alpha} v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) + c_{\beta\beta} v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) \\ & \quad + \gamma \mathsf{E}_{(\nu_{t+1}|\nu_t)} \widehat{V}_{t+1}^\beta(\rho(R_t, p_t^\alpha, p_t^\beta, \nu_t), \nu_{t+1}). \end{cases}$$

3.5 Split Stream Harvesting with Separate Information about Recruitments

Let us consider now the case of split stream harvesting (Section 2.1) when each player knows only the recruitment to his own stream, but the total recruitment $R = R^\alpha + R^\beta$. In this case his policy will depend on his partial recruitment and his knowledge of stochastic parameters, i.e.,

$$p^\alpha = P^\alpha(R^\alpha, \xi^\alpha), \qquad p^\beta = P^\beta(R^\beta, \xi^\beta).$$

It is easily seen that if the current θ is known to the players ($\xi^\alpha = \xi^\beta = \theta$) they can easily "reconstruct" common recruitment and the problem reduces to the problem considered in Section 3.2.1.

So, consider the case when the knowledge of θ is imprecise (obtained from measurements). Define the total payoff at time t recursively as the conditional average of the current payoff plus the discounted and averaged total payoff for the next season, i.e.,

$$V_t^\alpha(R_t^\alpha, \xi_t^\alpha, \mathbf{P}_t^\alpha, \mathbf{P}_t^\beta) = v_t^\alpha\left(R_t^\alpha, p_t^\alpha\right) + \gamma_\alpha \mathsf{E}_{(\theta_t, \xi_t^\beta | \xi_t^\alpha)} \mathsf{E}_{(\theta_{t+1}|\theta_t)} \mathsf{E}_{(\xi_{t+1}^\alpha|\theta_{t+1})} V_t^\alpha\left(R_{t+1}^\alpha, \xi_{t+1}^\alpha, \mathbf{P}_{t+1}^\alpha, \mathbf{P}_{t+1}^\beta\right).$$

Here

$$R_{t+1}^\alpha = \theta_{t+1}^\alpha F\left(\sigma^\alpha(R_t^\alpha, p_t^\alpha) + \sigma^\beta(\tfrac{\theta_t^\beta}{\theta_t^\alpha} R_t^\alpha, p_t^\beta)\right).$$

This leads to the following DP algorithm, which involves computation of a Nash equilibrium pair $\langle P_t^\alpha, P_t^\beta \rangle$ for the following functions:

$$\begin{cases} \widetilde{V}_t^\alpha(R_t^\alpha, \xi_t^\alpha, P_t^\alpha, P_t^\beta) & = & v_t^\alpha\left(R_t^\alpha, p_t^\alpha\right) + \gamma_\alpha \mathsf{E}_{(\nu_t, \xi_t^\beta | \xi_t^\alpha)} \mathsf{E}_{(\nu_{t+1}|\nu_t)} \\ & & \mathsf{E}_{(\xi_{t+1}^\alpha | \nu_{t+1})} \widehat{V}_{t+1}^\alpha\left(R_{t+1}^\alpha, \xi_{t+1}^\alpha\right), \\ \widetilde{V}_t^\beta(R_t^\beta, \xi_t^\beta, P_t^\alpha, P_t^\beta) & = & v_t^\beta(R_t^\beta, p_t^\beta) + \gamma_\beta \mathsf{E}_{(\nu_t, \xi_t^\alpha | \xi_t^\beta)} \mathsf{E}_{(\nu_{t+1}|\nu_t)} \\ & & \mathsf{E}_{(\xi_{t+1}^\beta | \nu_{t+1})} \widehat{V}_{t+1}^\beta(R_{t+1}^\beta, \xi_{t+1}^\beta), \end{cases}$$

where R_{t+1}^α and R_{t+1}^β are defined as

$$\begin{aligned} R_{t+1}^\alpha &= \theta_{t+1}^\alpha F\left(\sigma^\alpha(R_t^\alpha, P_t^\alpha(R_t^\alpha, \xi_t^\alpha)) + \sigma^\beta(\tfrac{\theta_t^\beta}{\theta_t^\alpha} R_t^\alpha, P_t^\beta(\tfrac{\theta_t^\beta}{\theta_t^\alpha} R_t^\alpha, \xi_t^\beta))\right), \\ R_{t+1}^\beta &= \theta_{t+1}^\beta F\left(\sigma^\alpha(\tfrac{\theta_t^\alpha}{\theta_t^\beta} R_t^\beta, P_t^\alpha(\tfrac{\theta_t^\alpha}{\theta_t^\beta} R_t^\beta, \xi_t^\alpha)) + \sigma^\alpha(R_t^\beta, P_t^\beta(R_t^\beta, \xi_t^\beta))\right). \end{aligned}$$

3.6 Sequential Harvesting with Separate Information about Recruitments

In this section we study the case of sequential harvesting (Section 3.1.2) when each player knows only his respective recruitment R^α or R^β. So his policy depends on his respective recruitment and information about random parameters.

For simplicity we assume that the only random parameter is in the growth function φ, i.e.,

$$p^\alpha = P^\alpha(R^\alpha, \xi^\alpha), \qquad p^\beta = P^\beta(R^\beta, \xi^\beta).$$

The player's escapement and immediate payoff depend on his recruitment and his policy,

$$\begin{aligned} S^\alpha = \sigma^\alpha(R^\alpha, p^\alpha), &\qquad S^\beta = \sigma^\beta(R^\beta, p^\beta), \\ v^\alpha = v^\alpha(R^\alpha, p^\alpha), &\qquad v^\beta = v^\beta(R^\beta, p^\beta). \end{aligned}$$

We can again define the total payoff at time t recursively as the conditional average of the current payoff plus the discounted and averaged total payoff for the next season, i.e.,

$$\begin{aligned} V_t^\alpha(R_t^\alpha, \xi_t^\alpha, \mathbf{P}_t^\alpha, \mathbf{P}_t^\beta) = v_t^\alpha\left(R_t^\alpha, p_t^\alpha\right) + \gamma_\alpha \mathsf{E}_{(\varphi_t, \xi_t^\beta | \xi_t^\alpha)} \mathsf{E}_{(\varphi_{t+1}|\varphi_t)} \\ \mathsf{E}_{(\xi_{t+1}^\alpha | \varphi_{t+1})} V_t^\alpha\left(R_{t+1}^\alpha, \xi_{t+1}^\alpha, \mathbf{P}_{t+1}^\alpha, \mathbf{P}_{t+1}^\beta\right), \end{aligned}$$

and a similar expression for player β. Note that the total payoff functions look very much the same, but essentially they are different, since the players' recruitments are very different:

$$\begin{aligned} R_{t+1}^\alpha &= F\left(\sigma^\beta(\sigma^\alpha(R_t^\alpha, p_t^\alpha), p_t^\beta), \varphi_t\right), \\ R_{t+1}^\beta &= \sigma^\alpha\left(F(\sigma^\beta(R_t^\beta, p_t^\beta), \varphi_t), p_{t+1}^\alpha\right). \end{aligned}$$

It is interesting that, due to the sequential nature of the players' actions, the solution of the problem may also be constructed sequentially without involving Nash equilibrium steps.

$$\begin{cases} \widehat{V}_t^\alpha(R_t^\alpha,\xi_t^\alpha) = \max_{p_t^\alpha}\Big[v_t^\alpha\left(R_t^\alpha,p_t^\alpha\right)+\gamma_\alpha \mathsf{E}_{(\nu_t,\xi_t^\beta|\xi_t^\alpha)}\mathsf{E}_{(\nu_{t+1}|\nu_t)}\mathsf{E}_{(\xi_{t+1}^\alpha|\nu_{t+1})} \\ \qquad\qquad \widehat{V}_{t+1}^\alpha\left(F\left(\sigma^\beta(\sigma^\alpha(R_t^\alpha,p_t^\alpha),\widehat{P}_t^\beta(\cdot)),\varphi_t\right),\xi_{t+1}^\alpha\right)\Big], \\ \widehat{V}_t^\beta(R_t^\beta,\xi_t^\beta) = \max_{p_t^\beta}\Big[v_t^\beta(R_t^\beta,p_t^\beta)+\gamma_\beta \mathsf{E}_{(\nu_t,\xi_t^\alpha|\xi_t^\beta)}\mathsf{E}_{(\nu_{t+1}|\nu_t)}\mathsf{E}_{(\xi_{t+1}^\beta|\nu_{t+1})} \\ \qquad\qquad \widehat{V}_{t+1}^\beta\left(\sigma^\alpha\left(F(\sigma^\beta(R_t^\beta,p_t^\beta),\varphi_t),\widehat{P}_{t+1}^\alpha(\cdot)\right),\xi_{t+1}^\beta\right)\Big], \end{cases} \tag{7}$$

where for shortness $\widehat{P}_t^\beta(\cdot)$ and $\widehat{P}_{t+1}^\alpha(\cdot)$ denote the corresponding optimal policies applied to the preceding value as an argument.

Thus,

$$\widehat{P}_t^\beta(R_t^\beta,\xi_t^\beta) = \arg\max_{p_t^\beta}[\cdot]$$

in the second expression of (7) and

$$\widehat{P}_t^\alpha(R_t^\alpha,\xi_t^\alpha) = \arg\max_{p_t^\alpha}[\cdot]$$

in the first expression give us optimal policies, which can be computed step by step, moving backwards from the horizon.

4 Concrete Details of Implementation

In this section we give a brief description of the algorithm and concrete parameters of the simulations demonstrated above.

4.1 Dynamic Programming Realization of the Algorithm

We describe first our basic algorithm, employing what is arguably the most natural finite horizon sequence of approximations to the given infinite horizon game. We then describe a series of variants of these finite time-horizon approximations, intended to uncover any additional natural sequential-limit equilibria of the infinite horizon game. We note that, in virtually all of the cases that we have examined (with exceptions noted below), only a single infinite horizon game-theoretic limiting equilibrium has been found, and that it is of the stationary target-escapement form that is also characteristic of cooperative centrally managed and risk-neutral harvests.

To be specific, let us consider the case of i.i.d. parameters and distinct observations. Let $T \geqslant 0$ denote the finite time-horizon for an approximating harvesting game. At each time stage t let $P_t^\alpha(R_t,\xi_t^\alpha)$ and $P_t^\beta(R_t,\xi_t^\beta)$ be single-stage harvest policies, where R_t is the current recruitment and ξ_t^α, ξ_t^β represent the players' current information on current and past random disturbances. Furthermore, let

$$\mathbf{P}_t^\alpha = \langle P_t^\alpha, P_{t+1}^\alpha, \ldots, P_T^\alpha \rangle$$

denote a full end-game policy for player α, and similarly for β. Then the expected payoffs V_t^α and V_t^β to the players in an end game beginning at stage t can be written in dynamic programming formulation as

$$\begin{cases} V_t^\alpha(R_t, \xi_t^\alpha, \mathbf{P}_t^\alpha, \mathbf{P}_t^\beta) & = \mathsf{E}_{(\nu_t, \xi_t^\beta | \xi_t^\alpha)} \Big[v_t^\alpha \Big(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), P_t^\beta(R_t, \xi_t^\beta) \Big) \\ & \quad + \gamma_\alpha \mathsf{E}_{\xi_{t+1}^\alpha} V_{t+1}^\alpha \Big(\rho_t(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), \\ & \quad P_t^\beta(R_t, \xi_t^\beta), \xi_{t+1}^\alpha, \mathbf{P}_{t+1}^\alpha, \mathbf{P}_{t+1}^\beta \Big) \Big] \\ V_t^\beta(R_t, \xi_t^\beta, \mathbf{P}_t^\alpha, \mathbf{P}_t^\beta) & = \mathsf{E}_{(\nu_t, \xi_t^\alpha | \xi_t^\beta)} \Big[v_t^\beta \Big(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), P_t^\beta(R_t, \xi_t^\beta) \Big) \\ & \quad + \gamma_\beta \mathsf{E}_{\xi_{t+1}^\beta} V_{t+1}^\beta \Big(\rho_t(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), \\ & \quad P_t^\beta(R_t, \xi_t^\beta), \xi_{t+1}^\beta, \mathbf{P}_{t+1}^\alpha, \mathbf{P}_{t+1}^\beta \Big) \Big] \end{cases} \tag{8}$$

in terms of the current single-stage policy pair $\langle P_t^\alpha, P_t^\beta \rangle$ and the subsequent end-game policy pair $\langle \mathbf{P}_{t+1}^\alpha, \mathbf{P}_{t+1}^\beta \rangle$.

Let $\langle \widehat{\mathbf{P}}_t^\alpha, \widehat{\mathbf{P}}_t^\beta \rangle$ denote a game-theoretic Nash equilibrium policy pair for the harvesting end game beginning at time t, and let

$$\widehat{V}_t^\alpha(R_t, \xi_t^\alpha) = V_t^\alpha(R_t, \xi_t^\alpha, \widehat{\mathbf{P}}_t^\alpha, \widehat{\mathbf{P}}_t^\beta), \qquad \widehat{V}_t^\beta(R_t, \xi_t^\beta) = V_t^\beta(R_t, \xi_t^\beta, \widehat{\mathbf{P}}_t^\alpha, \widehat{\mathbf{P}}_t^\beta)$$

denote the corresponding payoffs. Using the dynamic programming format, the optimal period t policy pair, $\langle \widehat{P}_t^\alpha, \widehat{P}_t^\beta \rangle$ can be expressed in terms of the subsequent end-game policy pair, $\langle \widehat{\mathbf{P}}_{t+1}^\alpha, \widehat{\mathbf{P}}_{t+1}^\beta \rangle$ as, respectively,

$$\begin{aligned} \widehat{P}_t^\alpha &= \arg\max_{P_t^\alpha} V_t^\alpha \left(R_t, \xi_t^\alpha, \langle P_t^\alpha, \widehat{\mathbf{P}}_{t+1}^\alpha \rangle, \langle \widehat{P}_t^\beta, \widehat{\mathbf{P}}_{t+1}^\beta \rangle \right), \\ \widehat{P}_t^\beta &= \arg\max_{P_t^\beta} V_t^\beta \left(R_t, \xi_t^\beta, \langle \widehat{P}_t^\alpha, \widehat{\mathbf{P}}_{t+1}^\alpha \rangle, \langle P_t^\beta, \widehat{\mathbf{P}}_{t+1}^\beta \rangle \right). \end{aligned} \tag{9}$$

Our basic algorithm consists of several nested loops. The external loop steps backward over time, beginning with the terminal period $t = T$. In our initial formulation, moving stage-by-stage backward in time via the solutions to (9), the algorithm arrives at an equilibrium end-game policy pair for successively larger time periods t. The algorithm compares the Nash policy pair at stage t with that at stage $t+1$ and, when these are deemed sufficiently close together, regards $\langle P_t^\alpha, P_t^\beta \rangle$ as a stationary policy pair for the infinite horizon game.

At each stage t, and with specified $\langle \widehat{\mathbf{P}}_{t+1}^\alpha, \widehat{\mathbf{P}}_{t+1}^\beta \rangle$, the algorithm's inner loop process carries out a reaction analysis leading to numerical evaluation of the pair of arg max values in (9): The process begins with arbitrary choice of an initial approximate α-fleet stage t policy P_t^α to supplement the known end-game policy pair $\langle \widehat{\mathbf{P}}_t^\alpha, \widehat{\mathbf{P}}_t^\beta \rangle$ and then determines, via the second equation of (9), the optimal β-fleet stage t policy response P_t^β. Next, the first equation in (9) is used to find the best α-fleet stage t response to that β-fleet policy, and so on. Then by dynamic programming based on the second equation of (8), the algorithm determines the optimal response to this policy, and so on until numerical convergence. The two processes must be carried out over a grid of values of R and ξ^α for the first player and R and ξ^β for the second one.

The first variation on this process is to alter the initial specification of period T equilibrium policies by assigning a so-called "scrap value" to the terminal stock left unharvested. In all cases, this alteration will lead to a terminally distinct series of t-stage approximations in policies and payoffs, but by 5 to 10 backward iterations these differences will have disappeared.

Our second algorithmic variation is more elaborate. We note that in all cases the policy sequences in our finite horizon approximations are non-stationary, tending to stationarity only as the horizon becomes more remote. In this second variation, we carried out our reaction analysis beginning always with stationary policies and reacting with the best stationary policies. In this way we searched for possible new equilibria for the infinite horizon problem, equilibria occupying distinct "basins of attraction" from the originally determined equilibrium. But we found nothing new in this way.

In order to find possible alternative Nash equilibria, several types of disturbances were applied to the algorithms. One type of disturbance consists in choosing specific initial iteration policies for the Nash equilibrium iterations at each season t. Specifically, initial policies were chosen from the target escapement class of policies with a target-escapement varied from zero harvesting to full stock harvesting.

In the second type of disturbance a kind of "scrap value" end-point condition was introduced. This kind of disturbance influenced the initial part of the iterations, but after 5–10 seasons the algorithm completely forgot about the "scrap value."

In all simulations for all types of disturbances both algorithms converged to the same pair of strategies (dependent on the game parameters only). So, in the limit these all seem to yield the same solution of the infinite horizon problem—generally a stationary solution of target escapement type. Of course, it does not prove the uniqueness, but it at least demonstrates a kind of global stability of the Nash equilibrium.

4.2 Immediate Payoff Function

We define policy as an escapement fraction, i.e.,

$$S^\alpha = p^\alpha R^\alpha, \qquad S^\beta = p^\beta R^\beta.$$

Then the players' harvests are

$$H^\alpha = (1-p^\alpha)R^\alpha, \qquad H^\beta = (1-p^\beta)R^\beta.$$

In our simulations the cost of harvesting is taken into account. We assume that the cost of harvesting is inversely proportional to the current fish stock in a given stream. Thus, the price of harvesting may be obtained as an integral from the current player's recruitment R down to his escapement $S = pR$:

$$\text{cost} = \int_S^R \frac{c}{x}\,dx = -c\log(p)$$

and the total immediate payoff

$$v(R,p) = H - \text{cost} = H + c\log(p).$$

4.3 Growth Function

In our studies we utilized various types of growth functions. Different types may reflect different specifics of a growth function. Simple growth functions are monotone increasing, concave, and have a compensatory behavior, i.e., $F(S) > S$ (at least for small stock S). However, there may be some deviations of a growth function from this simple type. First, it may not be increasing everywhere, e.g., decreases at high stock level due to overpopulation. Second, it may have depensatory behavior, when $F(S) < S$ for small enough S, i.e., falling below a certain level will lead the stock to complete extinction (even without harvesting). Typically, not everywhere increasing and/or depensatory growth functions lose their concavity for some regions of S.

In order to study depensatory (critical and non-critical) and compensatory cases in a uniform way we consider a growth function, which is specified as a cubic curve

$$F(S) = a_3S^3 + a_2S^2 + a_1S$$

that goes through the points $(0,0)$ (obviously) and $(1,1)$ and satisfies additional conditions

$$F'(0) = A, \qquad F'(1) = b.$$

These conditions uniquely determine a curve:

$$a_1 = A, \quad a_2 = 3 - b - 2A, \quad a_3 = b - 2 + A.$$

By keeping b fixed we can keep the behavior of the growth function almost the same at high S, while by changing A we can change the level of "depensation" (at low A) or compensation (at high A) of the growth function.

We say that there is "depensation" whenever low S gives convex $F(S)$, with a point of inflection over to concavity at higher S. "Critical depensation" requires an interval $(0, S_{\text{crit}})$ in which $F(S) < S$. It is easily seen that there is "critical depensation" whenever $0 < A < 1$ and critical escapement is given by

$$S_{\text{crit}} = -\frac{a_2 + a_3}{a_3} = \frac{A - 1}{A + b - 2}.$$

At the same time the curve has a point of inflection (from convexity to concavity) if $A < 1.5 - b/2$. Thus, we have "depensation" if $0 < A < 1.5 - b/2$, which becomes critical if $0 < A < 1$.

Figure 15 demonstrates three curves for the same $b = 0$ and different values of A: "critical depensation" $A = 0.6$ (top), "non-critical depensation" $A = 1$ (middle), "compensation" $A = 3$ (bottom).

4.4 Parameters of Simulations

All simulations presented in this chapter were performed for the split stream harvesting game. The default parameter values were the following:

- Payoff function (Section 4.2) with cost = 0.2;
- Growth function (Section 4.3) $b = 0$, $A = 0.6$ (critical depensation, Figure 15 top)
- States of θ: $\theta_1 = 0.1$ and $\theta_2 = 0.9$, i.i.d. with equal probabilities ($P(\theta = \theta_1) = P(\theta = \theta_2) = 0.5$);
- Discount factor: $\gamma_\alpha = \gamma_\beta = 0.9$;
- Cooperation weights (Section 3.4): $c_{\alpha\alpha} = 1$, $c_{\alpha\beta} = 0$, $c_{\beta\alpha} = 0$, $c_{\beta\beta} = 1$ (for no cooperation), or $c_{\alpha\alpha} = 0.5$, $c_{\alpha\beta} = 0.5$, $c_{\beta\alpha} = 0.5$, $c_{\beta\beta} = 0.5$ (for complete cooperation).

In order to vary the "completeness" information about θ we use imperfect observations (see Section 3.3). We change the "measurement precision" parameter π, which determines the measurement matrix

$$M = \begin{pmatrix} P(\xi = \xi_1|\theta = \theta_1) & P(\xi = \xi_1|\theta = \theta_2) \\ P(\xi = \xi_2|\theta = \theta_1) & P(\xi = \xi_2|\theta = \theta_2) \end{pmatrix}$$

(i.e., matrix of conditional probabilities of observable ξ for various values of θ) in the following way:

$$M = \begin{pmatrix} (1+\pi)/2 & (1-\pi)/2 \\ (1-\pi)/2 & (1+\pi)/2 \end{pmatrix}.$$

Thus, if $\pi = 1$ (maximum precision), $M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, which corresponds to "identical" measurement, while if $\pi = 0$ (minimum precision), $M = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$, which results in observations "independent" of θ states.

We now list the parameters of the simulations that are different from the default ones:

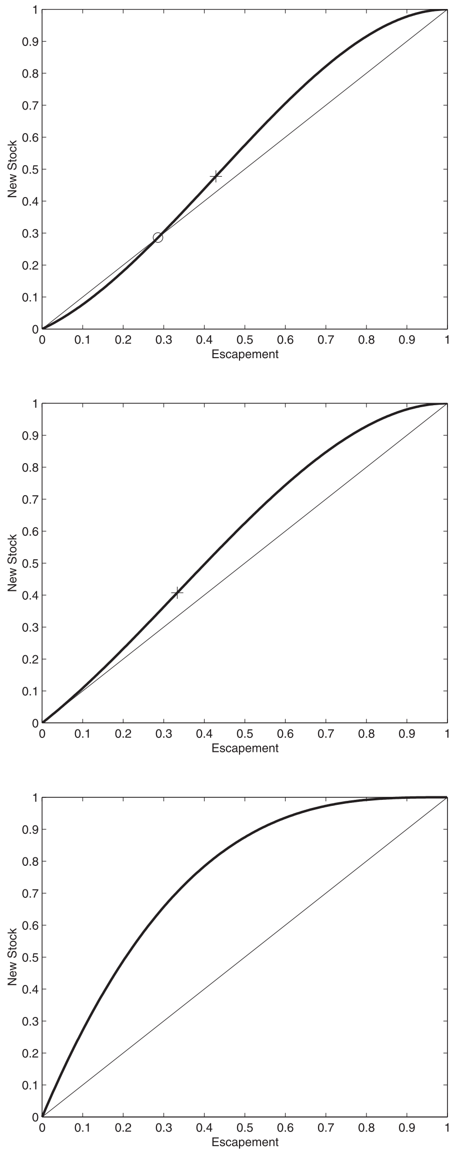


Figure 15: Cubic growth function ($b = 0$): "critical depensation" $A = 0.6$ (top), "non-critical depensation" $A = 1$ (middle), "compensation" $A = 3$ (bottom). Circle shows critical escapement S_{crit} and cross indicates point of inflection.

Figure 1. Influence of harvesting cost for different types of knowledge. "Compensatory" cubic growth function.

Compensation factor: $A = 3.0$ (Figure 15 bottom)

Variable parameter: Cost for both players from 0 to 1 with 0.05 increment.
Figure 2. Influence of harvesting cost for different types of knowledge. Growth function with "non-critical depensation."

Compensation factor: $A = 1.0$ (Figure 15 middle).

Variable parameter: Cost for both players from 0 to 1 with 0.05 increment.
Figure 3. Influence of harvesting cost for different types of knowledge. Growth function with "critical depensation."

Variable parameter: Cost for both players from 0 to 1 with 0.05 increment.
Figure 4. Compensatory harvesting with information from imperfect observation.

Compensation factor: $A = 3.0$ (Figure 15 bottom)

Variable parameter: Measurement precision from 0 to 1 with 0.05 increment.
Figure 5. Depensatory harvesting with information from imperfect observation. "Non-critical depensation."

Compensation factor: $A = 1.0$ (Figure 15 middle)

Variable parameter: Measurement precision from 0 to 1 with 0.05 increment.
Figure 6. Depensatory harvesting with information from imperfect observation. "Critical depensation."

Variable parameter: Measurement precision from 0 to 1 with 0.05 increment.
Figure 7. Harvesting at various degrees of depensation.

Variable parameter: Compensation factor A from 0.5 to 1 with 0.025 increment.

Figure 8. Harvesting at various degrees of compensation.

Variable parameter: Compensation factor A from 1 to 3 with 0.1 increment.
Figure 9. Influence of measurement precision in asymmetric natural environment case.

States of θ: $\theta_1 = 0.1$ and $\theta_2 = 0.8$. Thus, asymmetry in natural environment favors the second player.

Variable parameter: Measurement precision from 0 to 1 with 0.05 increment.
Figure 10. Influence of measurement precision in asymmetric economic environment case.

Costs of harvesting are different: $\text{cost}_1 = 0.2$, $\text{cost}_2 = 0.15$. Again, asymmetry in economic environment favors the second player.

Variable parameter: Measurement precision from 0 to 1 with 0.05 increment.
Figure 11. Influence of the level of environmental variability.

Variable parameter: variability v of θ from 0 to 1 with 0.05 increment. θ takes states $\theta_1 = (1-v)/2$ and $\theta_2 = (1+v)/2$ with equal probabilities.
Figure 12. Increasing degree of cooperation. Current vs. minimum knowledge.

Variable parameter: degree of cooperation c from 0 to 1 with 0.05 increment. Cooperation weights are determined by c as follows: $c_{\alpha\alpha} = 1 - c/2$, $c_{\alpha\beta} = c/2$, $c_{\beta\alpha} = c/2$, $c_{\beta\beta} = 1-c/2$. Thus, when c changes from 0 to 1, cooperation weights change from "no cooperation" to "complete cooperation."

Figure 13. Increasing degree of cooperation in asymmetric environment case. Current and minimum knowledge. Critical depensation case.

States of θ: $\theta_1 = 0.1$ and $\theta_2 = 0.8$. Asymmetry in natural environment favors the second player.

Variable parameter: degree of cooperation c from 0 to 1 with 0.05 increment.

Figure 14. Increasing degree of cooperation in asymmetric environment case. Current and Minimum knowledge. Compensatory growth function.

Compensation factor: $A = 3.0$ (Figure 15 bottom).

States of θ: $\theta_1 = 0.1$ and $\theta_2 = 0.8$. Asymmetry in natural environment favors the second player.

Variable parameter: degree of cooperation c from 0 to 1 with 0.05 increment.

REFERENCES

[1] Clark C.W., Restricted access to common-property fishery resources: a game-theoretic analysis, in P.-T. Liu (ed.), *Dynamic Optimization and Mathematical Economics*, 117–132, Plenum, New York, 1980.

[2] Levhari D. and Mirman L.J., The great fish war: an example using a dynamic Cournot-Nash solution, *Bell Journal of Economics*, 11, 322–344, 1980.

[3] Geanakopolos J., Common knowledge, in R.J. Aumann and S. Hart (eds.), *Handbook of Game Theory with Economic Applications*, Vol. II, Chap. 40, Elsevier, Amsterdam, 1984.

[4] McKelvey R., Miller, K., and Golubtsov, P., Fish wars revisited: Stochastic incomplete-information harvesting game, in J. Wesseler, H.-P. Weikard, and R.D. Weaver (eds.), *Risk and Uncertainty in Environmental and Natural Resource Economics*, 93–112, 2003.

[5] McKelvey R., Golubtsov, P., Miller, K., and Cripe G., Binational management of a transboundary Marine fishery: Modeling the destabilizing impacts of erratic climatic shifts, to appear in R. Hannesson, M. Barange, and S. Herrick, Jr. (eds.), *Climate Change and the Economics of the World's Fisheries: Examples of Small Pelagic Stocks*, Edgard Elgar Press, 2006.

[6] Hernández-Lerma O. and Lasserre J.B., *Discrete-Time Markov Control Processes. Basic Optimality Criteria*, Springer, New York, 1996.

A Two-Level Differential Game of International Emissions Trading*

Alain Haurie
Logilab-HEC, University of Geneva
CH-1211 Geneva 4, Switzerland
alain.haurie@hec.unige.ch

Francesco Moresino
Logilab-HEC, University of Geneva
CH-1211 Geneva 4, Switzerland
francesco.moresino@hec.unige.ch

Laurent Viguier
EPFL REME
BP 16, CH-1015 Lausanne, Switzerland
laurent.viguier@epfl.ch

Abstract

In this chapter we propose a dynamic game-theoretic modeling framework for the international climate change negotiations that should take place at the end of the Kyoto Protocol agreement if the necessity to drastically curb carbon emissions is confirmed. The model is composed of a set of optimal economic growth models corresponding to the different groups of nations that will be parties in the negotiations. Emissions of greenhouse gases (GHGs) are represented as by-products of the economic production process. Two types of capital (clean vs. dirty) can be used to produce the economic good with different emissions effects. The negotiations should determine a set of allowances that define caps on GHG emissions such that a long-term constraint on total emissions is satisfied. At each instant of time, given the emissions caps, an international emissions trading system is organized. In order to be self-enforcing, the emissions caps and the economic growth paths have to satisfy a noncooperative equilibrium condition. We describe this two-level game structure mathematically and give the necessary optimality conditions that must be satisfied by the equilibrium solution under the coupled global emission constraint.

*This work was supported by the NCCR-Climate program of the Swiss NSF and by the GICC, Ministry of Ecology and Sustainable Development, France.

1 Introduction

The aim of this paper is to propose a differential game formalism to represent the negotiation of a self-enforcing agreement on global greenhouse gas (GHG) emissions reduction for a group of nations in different states of development.

The climate change process due to GHG emissions (in particular, CO_2 due to fossil fuel combustion) poses a delicate international negotiation problem. The Kyoto Protocol which has recently entered into force for a group of developed countries is typical of an agreement that stipulates quantities (emission caps) rather than prices (e.g., carbon taxes). It lies with each participating country to agree on a cap for its own emissions and to participate in an international emissions trading system to achieve the global emission reduction aim at minimal global cost. In the current situation, developing countries (DCs) are not part of the agreement, and the United States, Australia and a few other industrialized countries have not signed the protocol. Furthermore, the global aim of the Kyoto Protocol, in terms of global abatement, is modest. Climate research is making progress, and it appears more and more probable that there could soon be a universal scientific consensus on the need to drastically curb GHG emissions in order to avoid a disastrous climate change. This global constraint should be imposed on the total cumulated emissions of all countries, including DCs.[1] The agreement will therefore have to find an appropriate trade-off between the development needs and the limits to emissions imposed upon these countries. Furthermore, to be self-enforcing, an agreement should be a Nash equilibrium, in that each national policy should be the best reply to the policies decided by the other nations, while keeping the global environmental constraint satisfied. This calls naturally for the definition of the agreement as a normalized equilibrium, according to the definition proposed by J.B. Rosen [16], for games where the strategies of players are coupled by a common constraint that must be satisfied.

The use of Rosen's normalized equilibrium concept in the modeling of environmental negotiations has been proposed in [5] and [9], and the mathematical theory of open-loop infinite-horizon differential games with coupled constraints has been fully developed in [3]. Application of the concept to games described by distributed parameter systems has been considered in [2]. Applications of the concept to the modeling of local pollution problems have been proposed in [6] and further explored in [11]. The application we describe herein is original in the sense that the negotiation game has a two-level hierarchical structure. Once the emissions caps are decided for each country, the actual emission reductions will be decided via the implementation of an emissions trading scheme. These competitive markets for carbon emissions ensure an efficient abatement policy worldwide and generate transfer payments that can help the growth of DCs.

[1] Under the business-as-usual scenario, DCs will soon be responsible for more than 50% of the total GHG emissions.

The combination of a general equilibrium model determining the growth paths of different nations and of a noncooperative game structure to decide on the abatement policies of the different (groups of) nations was first proposed in [14]. In [10] the concept of a two-level hierarchical game has been proposed, where nations decide on their allowances while an international emissions trading system is organized. This approach has been extended in [4] to a multi-country general equilibrium framework. A similar two-level hierarchical game structure has been used in [8] to represent the dynamic competition between Russia and China in the exploitation of "hot air" in the Kyoto Protocol. Another two-level hierarchical game based on the computable general equilibrium model (CGEM) GEMINI-E3 has been used in [1] to represent the strategic behavior of European countries in the allocation of allowances for the implementation of the Kyoto Protocol. This chapter presents a generalization of the model proposed in [10] to a dynamic setting and to the concept of normalized equilibrium applied to the negotiation of long-term climate policy. A related game-theoretic approach to emissions reduction and emissions trading is considered in [17,18].

The chapter is organized as follows. In Section 2 we recall the context of international climate negotiations and the concept of self-enforcing agreement. In Section 3 we detail the multi-region optimal economic growth framework and the economic equilibrium conditions on the international emissions trading market. In Section 4 we formulate the upper-level dynamic game to determine the allowances (emission caps) that will lead to the satisfaction of the long-term constraint on cumulative emissions; the definition of a normalized equilibrium is recalled, and the necessary conditions for equilibrium are obtained. In Section 5 these characterizations are given an economic interpretation, and in Section 6 we conclude by announcing the possibility of extending the approach to a model using a fully fledged CGEM, in order to take into account the terms of trade effects.

2 The Context

We consider the following situation. (i) After a comprehensive research effort, the scientific community arrives at a clear understanding of the climate change effect of GHGs; this determines a precise constraint on the long-term accumulation of GHGs in the atmosphere that must be satisfied globally on the planet. (ii) Given these scientific conclusions that cannot be dismissed by any country, an international agreement should be reached on the relative development paths of the different countries and their use of GHGs to foster their development. (iii) GHGs can be used as a by-product in the economic production process, permitting the use of cheaper "dirty" technologies, but their abatement can also be used as a source of income in an international emissions trading system. (iv) The agreement should be self-enforcing. This means that a non-cooperative equilibrium condition must hold under which the development and emissions cap paths for each country are the best reply to the decisions of the other countries.

We now describe the representation of the long-term constraint on cumulative emissions. GHGs are long lived and their effect on climate change is relatively slow compared to the economic dynamics. We therefore propose to represent the limit to growth in GHG accumulation by the following total discounted sum of emissions:

$$\int_0^\infty e^{-\rho t}\bar{e}(t)\,dt \leq \bar{E}, \tag{1}$$

where $\bar{E}$ is a given bound, $\bar{e}(t)$ represents the total emissions at time t and ρ is a pure time preference rate. This extends the representation of the impact of climate changes on the world economies proposed by Labriet and Loulou in [12] to an infinite-horizon setting. In their work, these authors established a direct link between cumulative emissions and damages. In our model, the time horizon being infinite, we propose to represent the damage as a function of the total discounted sum of emissions. To provide some justification for the use of such a constraint, let us assume that the damage at time t is a linear function[2] of the total emissions up to time t,

$$d(t) = \alpha \int_0^t \bar{e}(s)\,ds. \tag{2}$$

Now consider that the planner wants to limit the total discounted damage, represented as

$$D = \int_0^\infty e^{-\rho t} d(t)\,dt = \int_0^\infty e^{-\rho t}\left(\alpha \int_0^t \bar{e}(s)\,ds\right)\,dt. \tag{3}$$

Integrating (3) by parts we obtain

$$D = \left[\alpha \int_0^t \bar{e}(s)\,ds \times \left(-\frac{e^{-\rho t}}{\rho}\right)\right]_0^\infty + \frac{\alpha}{\rho}\int_0^\infty e^{-\rho t}\bar{e}(t)\,dt. \tag{4}$$

The term between the square brackets vanishes and we have

$$D = \frac{\alpha}{\rho}\int_0^\infty e^{-\rho t}\bar{e}(t)\,dt. \tag{5}$$

Therefore, the constraint on the discounted sum of emissions is equivalent to a constraint on the discounted sum of damages.

[2] Damages are usually represented as a nonlinear function of the surface average temperature change which is due to the radiative forcing of GHG concentrations (itself a nonlinear function of these concentrations). In [12] it has been observed that the damage functions used in the literature can be accurately summarized by a linear dependence on the total cumulative emissions.

In summary, the situation that is described here corresponds to a set of growing economies which emit GHGs as a by-product of their economic output and which must find a mutually agreeable emissions abatement plan in order to collectively obtain satisfaction of the long-term constraint (1).

3 The Modeling Framework

In this section we describe the economic model that will serve to define the game problem.

3.1 The Dynamic Multi-Country Economic Growth Model

We consider a set of countries, each described by an economic growth model according to Ramsey [15]. Each country produces a homogenous good that can be either consumed or invested in two types of productive capital. The capital of type 1, denoted K_j^1, is dirty in the sense that it generates a high amount of emissions e_j^1 as a by-product, whereas the capital of type 2, denoted K_j^2, is clean as it generates a low amount of emissions e_j^2 as a by-product. We now summarize this economic growth model.

Players: Set of negotiating countries $j = 1, \dots, m$.

Welfare: Discounted sum of utility derived from consumption for each country

$$\int_0^\infty e^{-\rho t} L_j(t) \log[c_j]\, dt, \quad j = 1, \dots, m,$$

where $c_j(t) = C_j(t)/L_j(t)$ denotes per capita consumption in country j.

Population dynamics: Exogenous population growth

$$\dot{L}_j(t) = g_j(t) L_j(t).$$

Production functions: Output[3] of country j depends on emissions e_j, dirty capital K_j^1, clean capital K_j^2 and labor L_j

$$Y_j = F_j(e_j, K_j^1, K_j^2, L_j), \quad j = 1, \dots, m.$$

[3] We may assume the following form for the production function:

$$F_j(e_j, K_j^1, K_j^2, L_j) = \max_{e_j^1, e_j^2 : e_j = e_j^1 + e_j^2} L_j^\alpha (A_j^1 {e_j^1}^{\beta^1} {K_j^1}^{\gamma^1} + A_j^2 {e_j^2}^{\beta^2} {K_j^2}^{\gamma^2}).$$

In general, we also assume the production function to be concave in all its arguments.

Capital dynamics: The usual accumulation equations with constant depreciation rates,

$$\dot{K}_j^i = I_j^i - \mu_j^i K_j^i, \quad j = 1, \ldots, m, \quad i = 1, 2.$$

Allocation of output: The flexible good can be consumed or invested,

$$C_j \leq Y_j - I_j^1 - I_j^2, \quad j = 1, \ldots, m.$$

3.2 The Emissions Cap Negotiation Game

The negotiation is undertaken under an international aegis and concerns the establishment of GHG emissions caps at each time t and for each country j in order to maintain a sustainable climate. The negotiation boils down to the choice of the strategic emissions caps with a coupled constraint linking these caps over the infinite planning horizon.

Strategic variable: Emission caps schedule for each country,

$$\hat{e}_j(t), \quad j = 1, \ldots, m.$$

We shall denote by $\bar{e}(t) = \sum_{j=1,\ldots,m} \hat{e}_j(t)$ the total emissions level defined at time t by the emissions caps of the m countries.

Coupled constraint: A global long-term bound on total emissions,

$$\int_0^\infty e^{-\rho t} \left(\sum_{j=1,\ldots,m} \hat{e}_j(t) \right) dt = \int_0^\infty e^{-\rho t} \bar{e}(t)\, dt \leq \bar{E}. \tag{6}$$

3.3 The Lower-Level Emissions Trading Equilibrium

At each time t, a market for emissions trading is organized and each country $j = 1, \ldots, m$ determines its output and emissions level by solving the following local optimization problem:

$$\max \quad L_j(t) \log \left[\frac{C_j(t)}{L_j(t)} \right] \tag{7}$$

s.t.

$$Y_j(t) = F_j(e_j(t), K_j^1(t), K_j^2(t), L_j(t)) + \pi(t)(\hat{e}_j(t) - e_j(t)) \tag{8}$$

$$C_j(t) \leq Y_j(t) - I_j^1(t) - I_j^2(t), \tag{9}$$

where the market price $\pi(t) \geq 0$ is such that the global emissions constraint holds and the market clears:

$$\sum_j e_j(t) \leq \sum_j \hat{e}_j(t) = \bar{e}(t) \tag{10}$$

$$\pi(t)\left(\sum_j e_j(t) - \bar{e}(t)\right) = 0. \tag{11}$$

In the preceding expressions (11) is the usual complementarity slackness condition which tells that the price is zero when the constraint is not active.

In this problem the decision variable for each country j is simply the emissions level $e_j(t)$; all the other variables, $L_j(t), K_j^1(t), K_j^2(t), I_j^1(t), I_j^2(t)$ are fixed and the consumption level $C_j(t)$ is a direct consequence of the choice of $e_j(t)$.

This set of local optimization problems in the variables $e_j(t), j = 1, \ldots, m$ is equivalent to solving the following auxiliary optimization problem:[4]

$$\max \quad \sum_{j=1,\ldots,m} F_j(e_j(t), K_j^1(t), K_j^2(t), L_j(t)) \tag{12}$$

$$\text{s.t.}$$

$$\sum_j e_j(t) \leq \bar{e}(t). \tag{13}$$

The global cap at time t thus determines the price of permits $\pi(t)$, defined as the Kuhn–Tucker multiplier associated with the constraint (13) and the respective emissions levels. At the optimum one has equalization of the marginal productivities of emissions:

$$\frac{\partial}{\partial e_j} F_j(e_j(t), K_j^1(t), K_j^2(t), L_j(t)) = \pi(t), \quad j = 1, \ldots, m, \tag{14}$$

$$\sum_j e_j(t) = \bar{e}(t). \tag{15}$$

It will be convenient to denote by $s_j(t) = (K_j^1(t), K_j^2(t), L_j(t))$ the state variable values at time t for country j.

3.3.1 How Post-Trading Emissions Depend on Total Emissions Limit

We adapt to our setting the analysis of [10] to show how emissions levels of each country will be determined in the carbon market at time t. First we differentiate (14) with respect to $\pi(t)$ and (15) with respect to $\bar{e}(t)$ to get

$$\frac{\partial^2}{\partial e_j^2} F_j(e_j(t), s_j(t)) e_j'(\pi(t)) = 1, \quad j = 1, \ldots, m, \tag{16}$$

$$\sum_{j=1}^{m} e_j'(\pi(t)) \pi'(\bar{e}(t)) = 1. \tag{17}$$

[4] It suffices to write the optimality conditions for the two problems to see the equivalence.

This yields, after substitution and rearranging,

$$\pi'(\bar{e}(t)) = \frac{1}{\displaystyle\sum_{j=1}^{m} \frac{1}{\frac{\partial^2}{\partial e_j^2} F_j(e_j(t), s_j(t))}} < 0. \tag{18}$$

Now we differentiate (14) w.r.t. $\bar{e}(t)$ to get

$$\frac{\partial^2}{\partial e_j^2} F_j(e_j(t), s_j(t)) e_j'(\bar{e}(t)) = \pi'(\bar{e}(t)). \tag{19}$$

Substituting $\pi'(\bar{e}(t))$ from (18) yields

$$e_\ell'(\bar{e}(t)) = \frac{1}{\displaystyle\sum_{j=1}^{m} \frac{\frac{\partial^2}{\partial e_\ell^2} F_\ell(e_\ell(t), s_j(t))}{\frac{\partial^2}{\partial e_j^2} F_j(e_j(t), s_j(t))}} \in [0, 1]. \tag{20}$$

It thus appears that, when the total emissions limit $\bar{e}(t)$ increases, the carbon market price decreases and the emissions of each country will increase by a fraction determined by the ratio (20).

3.3.2 The Reduced Welfare Function

In brief, for each country j at time t, given the capital stocks, K_j^1, K_j^2, the population level L_j, the investment rates I_j^1, I_j^2 and the cap levels $\hat{e}_j$ and $\bar{e}$ there is a welfare gain

$$\mathcal{L}_j(L_j, K_j^1, K_j^2, I_j^1, I_j^2, \hat{e}_j, \bar{e}),$$

which is determined by the solution of the local optimization problem (7)–(9). It will be useful for the analysis of the upper-level differential game to express the partial derivative[5] $(\partial/\partial \hat{e}_j)\mathcal{L}_j$. From (7)–(9) we get

$$\begin{aligned}
\frac{\partial}{\partial \hat{e}_j} \mathcal{L}_j &= \frac{1}{c_j} \frac{\partial C_j}{\partial \hat{e}_j} \\
\frac{\partial C_j}{\partial \hat{e}_j} &= \frac{\partial F_j}{\partial e_j} \frac{\partial e_j}{\partial \hat{e}_j} + \pi \left(1 - \frac{\partial e_j}{\partial \hat{e}_j}\right) + \frac{\partial \pi}{\partial \hat{e}_j}(\hat{e}_j - e_j).
\end{aligned}$$

Taking into account that $\partial F_j/\partial e_j = \pi$ at equilibrium on the carbon trading market and that $\partial \pi/\partial \hat{e}_j = \pi'(\bar{e})$, we finally obtain

$$\frac{\partial \mathcal{L}_j}{\partial \hat{e}_j} = \frac{1}{c_j}(\pi + \pi'(\bar{e})(\hat{e}_j - e_j)). \tag{21}$$

[5]To simplify notation we omit the running variables.

4 The Upper-Level Differential Game Problem

We can now reformulate the problem of determining together the capital accumulation paths and the emissions caps for the m countries in the form of a noncooperative differential game with a coupled constraint.

$$\begin{aligned}
\text{Equil.} &= \int_0^\infty e^{-\rho t}\mathcal{L}_j(L_j(t), K_j^1(t), K_j^2(t), I_j^1(t), I_j^2(t), \hat{e}_j(t), \bar{e}(t))\,dt,\\
&\qquad j = 1,\dots,m,\\
&\text{s.t.}\\
\dot{K}_j^i(t) &= I_j^i(t) - \mu_j^i K_j^i(t) \quad i = 1,2, \quad j = 1,\dots,m,\\
\bar{E} &\geq \int_0^\infty e^{-\rho t}\left(\sum_{j=1,\dots,m} \hat{e}_j(t)\right) dt = \int_0^\infty e^{-\rho t}\bar{e}(t)\,dt.
\end{aligned}$$

4.1 Normalized Equilibrium Solutions

We represent the result of the international negotiation on GHG emissions as a *normalized equilibrium* in a dynamic game with a coupled constraint. This concept has been introduced by Rosen [16] to deal with situations where the players are bound by a constraint that links all their strategies together.

Let us recall briefly the definition of an equilibrium under a coupled constraint. Let U_j be the strategy set of player $j = 1,\dots,m$ and let $\mathcal{U} \subset \underline{U} = \Pi_{j=1}^m U_j$ be a proper subset of the cartesian product of strategy sets. Player j has a payoff defined by the function $\Psi_j(u_1, u_2, \dots, u_m)$.

Definition 4.1. The strategy vector $\underline{u}^* \in \mathcal{U}$ is an equilibrium under the coupled constraint $\underline{U}$ if

$$\begin{aligned}
\Psi_j(u_1^*, \dots, u_j, \dots, u_m^*) \leq\ & \Psi_j(u_1^*, \dots, u_j^*, \dots, u_m^*)\\
& \forall u_j \in U_j \text{ s.t. } (u_1^*, \dots, u_j, \dots, u_m^*) \in \mathcal{U}. \qquad (22)
\end{aligned}$$

Let $\alpha = (\alpha_j)_{j=1,\dots,m}$ be a set of positive weights given to the m players. We construct the weighted payoff function

$$\Phi(\underline{u};\alpha) = \sum_{j=1}^m \alpha_j \Psi_j(\underline{u}) \qquad (23)$$

and define the reaction function

$$\Theta(\underline{u}, \underline{v}; \alpha) = \sum_{j=1}^m \alpha_j \Psi_j(u_1, \dots, v_j, \dots, u_m) \qquad (24)$$

with the point-to-set map

$$\Xi(\underline{u};\alpha) = \left\{ \underline{v}^* \in \mathcal{U} : \Theta(\underline{u}, \underline{v}^*; \alpha) = \max_{\underline{v} \in \mathcal{U}} \Theta(\underline{u}, \underline{v}; \alpha) \right\}. \tag{25}$$

The following lemma is proved in [16]

Lemma 4.1. *A fixed point of the point-to-set map $\Xi(\underline{u};\alpha)$, i.e., a strategy vector $\underline{u}^*$ such that $\underline{u}^* \in \Xi(\underline{u}^*;\alpha)$ is an equilibrium under the coupled constraint $\mathcal{U}$.*

Finally, if the constraint $\mathcal{U}$ is defined by a set of inequalities $h(\underline{u}) \geq 0$, where $h(\cdot)$ is a vector-valued function, we can associate with the fixed point condition of Lemma 4.1 a Kuhn–Tucker[6] multiplier $\nu \geq 0$.

Definition 4.2. The strategy vector $\underline{u}^*$ is a normalized equilibrium if the following conditions hold:

$$\begin{aligned}
\Psi_j(\underline{u}^*) + \frac{1}{\alpha_j}\nu^T h(\underline{u}^*) &= \max_{u_j \in U_j} \Psi_j([\underline{u}^{*-j}, u_j]) + \frac{1}{\alpha_j}\nu^T h([\underline{u}^{*-j}, u_j]) \\
& \qquad j = 1, \ldots, m && (26) \\
0 &= \nu^T h(\underline{u}^*) && (27) \\
0 &\leq h(\underline{u}^*). && (28)
\end{aligned}$$

According to this definition, the players share a common Kuhn–Tucker multiplier, but use it with a weighting reflecting their relative importance in sharing the benefits. The conditions (26) define a Nash equilibrium for a game with an extended payoff system built from the common multiplier and the weights given to the different players. The multiplier ν is chosen such that, at equilibrium, the conditions (27), (28) are satisfied.

4.2 First-Order Conditions

The application of Rosen conditions to open-loop differential games has been explored in several papers [9,5,6,3]. We write these first-order conditions assuming that enough regularity holds.

4.2.1 Hamiltonians

Consider the costate (adjoint) variables λ_j^i, $j = 1, \ldots, m$, $i = 1, 2$ expressed in current value terms and the Kuhn–Tucker multiplier ν associated with the coupled constraint. Define the current-valued Hamiltonians

$$H_j(L_j, K_j^1, K_j^2, I_j^1, I_j^2, \hat{e}_j, \bar{e}; \lambda_j^1, \lambda_j^2, \nu, \alpha_j) =$$

[6] Assuming that the constraint qualification conditions hold.

$$\mathcal{L}_j(L_j, K_j^1, K_j^2, I_j^1, I_j^2, \hat{e}_j, \bar{e}) + \sum_{i=1,2} \lambda_j^{i\,T}(I_j^i(t) - \mu_j^i K_j^i) - \frac{\nu}{\alpha_j}\hat{e}_j.$$

The necessary optimality conditions for a normalized equilibrium are given, for each player j, by

$$0 = \hat{e}_j \frac{\partial}{\partial \hat{e}_j} H_j(L_j, K_j^1, K_j^2, I_j^1, I_j^2, \hat{e}_j, \bar{e}; \lambda_j^1, \lambda_j^2, \nu, \alpha_j) \tag{29}$$

$$0 \leq \hat{e}_j \tag{30}$$

$$0 = I_j^i \frac{\partial}{\partial I_j^i} H_j(L_j, K_j^1, K_j^2, I_j^1, I_j^2, \hat{e}_j, \bar{e}; \lambda_j^1, \lambda_j^2, \nu, \alpha_j); i = 1, 2 \tag{31}$$

$$0 \leq I_j^i \tag{32}$$

$$\dot{\lambda}_j^i = -\frac{\partial}{\partial K_j^i} H_j(L_j, K_j^1, K_j^2, I_j^1, I_j^2, \hat{e}_j, \bar{e}; \lambda_j^1, \lambda_j^2, \nu, \alpha_j) + \rho \lambda_j^i; \quad i = 1, 2. \tag{33}$$

Here the parameter ν is the common Kuhn–Tucker multiplier associated with the coupled global emissions constraint (6) which verifies

$$0 \leq \nu \tag{34}$$

$$0 \leq \bar{E} - \int_0^\infty e^{-\rho t} \bar{e}(t)\, dt \tag{35}$$

$$0 = \nu \left(\bar{E} - \int_0^\infty e^{-\rho t} \bar{e}(t)\, dt \right) \tag{36}$$

$$\bar{e}(t) = \sum_{j=1,\dots,m} \hat{e}_j(t). \tag{37}$$

4.2.2 Determination of Emissions Caps

Country j is informed of the value of the common multiplier ν and its relative weight $\alpha_j > 0$. Then the cap at time t is determined by the condition (29), which can be written as

$$\frac{\partial \mathcal{L}_j}{\partial \hat{e}_j} = \frac{\nu}{\alpha_j} \tag{38}$$

or, also, in view of (21),

$$\frac{1}{c_j}(\pi + \pi'(\bar{e})(\hat{e}_j - e_j)) = \frac{\nu}{\alpha_j}. \tag{39}$$

4.2.3 Determination of Investment Rates

The determination of investment rates is defined by condition (31), which can be written, for country j, when the positivity constraint is not active, as

$$\frac{\partial \mathcal{L}_j}{\partial I_j^i} = -\lambda_j^i, \quad i = 1, 2. \tag{40}$$

It is an easy matter to express $\partial \mathcal{L}_j / \partial I_j^i$ in terms of consumption at time t and to get

$$\frac{1}{c_j} = \lambda_j^i, \quad i = 1, 2, \tag{41}$$

which is the usual equalization of marginal utility of consumption with marginal productivity of capital.

5 Interpretation

It is interesting to compare the results obtained in this formulation with the one obtained in [10] for a static model where a standard Nash equilibrium has been defined for a group of countries that are exposed to different environmental damages due to the total emissions stock. In addition to the introduction of the dynamic effect of investment in clean technologies and development, the model proposed herein provides a way to determine, as part of the negotiation process, the share of the burden that will be incurred by each country. This is clear when one looks at the conditions that determine the permit sellers and buyers.

5.1 Permit Sellers Have High Weights

From (39) we can express the permit supply at time t as

$$\hat{e}_j - e_j = -\frac{\pi - \frac{c_j \nu}{\alpha_j}}{\pi'(\bar{e})}. \tag{42}$$

We can then observe the effect of increasing the relative weight α_j for country j. If α_j is high enough, it will cause this country to be a permit seller on the carbon market at time t. In other terms, the parameter ν is a global marginal cost due to the long-term accumulation of GHGs; the term $c_j \nu / \alpha_j$ is the share of that cost that is allocated to nation j.

5.2 Global Abatement Will Be Effective

It is important to emphasize the fact that the global marginal cost ν and the relative weights $\alpha_j, j = 1, \ldots, m$ are jointly defined in such a way that the equilibrium solution satisfying (29)–(33) will also satisfy in the long run the global cumulative emissions constraint (6).

6 Conclusion

In this chapter we have extended the model of international emissions trading with endogenous allowance choices proposed in [10] in two directions. First we have proposed a dynamic framework retaining the two-level structure, with an international market for GHGs at each time period t; we have also replaced the cost-benefit analysis framework proposed in [10], where each nation has a specific damage function, with a cost-effectiveness framework, where the equilibrium must be reached under a coupled constraint that imposes a global limit on cumulative emissions. The interpretation of the necessary conditions for equilibrium shows the close relationship between the results of [10] and those reported here. However, a major difference lies in the determination of the global marginal cost parameter ν, together with the weighting scheme α_j, $j = 1, \ldots, m$, which leads to an equilibrium that satisfies, in the long run, the global environmental constraint, with a sharing of the burden that could contribute to development aid via the trading of emission rights. We feel that the formalism proposed in this chapter could therefore contribute to a better understanding of the terms of the forthcoming international agreements that seem to be necessary in order to tackle climate change threats.

Indeed, the model proposed here has no representation of international trade, except for the GHG emissions. It could be easily extended to include such a description, e.g., following the description made in the RICE model [14]. Another promising approach consists in implementing the normalized equilibrium model in a multi-country and multi-sector CGEM with full representation of the terms of trade effects. Preliminary results in this direction are reported in [7].

REFERENCES

[1] Viguier L., Vielle M., Haurie A. and Bernard A., A two-level computable equilibrium model to assess the strategic allocation of emission allowances within the European Union, *Computers & Operation Research*, 33, 2, 369–385, 2006.

[2] Carlson D.A., Normalized overtaking Nash equilibrium for a class of distributed parameter dynamic games, in A. Nowak and K. Szajowski (eds.), *Advances in Dynamic Games*, Annals of the International Society of Dynamic Games, 7, 163–182, Birkhäuser, Boston, 2005.

[3] Carlson D.A. and Haurie, A., Infinite horizon dynamic games with coupled state constraints. in J.A. Filar, V. Gaitsgory and K. Mizukami (eds.), *Advances in Dynamic Games and Applications*, Annals of the International Society of Dynamic Games, 5, 196–212, Birkhäuser, Boston, 2000.

[4] Carbone J.C., Rutherford T. and Helm C., Coalition formation and international trade in greenhouse gas emissions, mimeo, 2003.

[5] Haurie A., Environmental coordination in dynamic oligopolistic markets, *Group Decision and Negotiation*, 4, 46–67, 1995.

[6] Haurie A. and Krawczyk J., Optimal charges on river effluent from lumped and distributed sources, *Environmental Modeling & Assessment*, 2, 177–199, 1997.

[7] Haurie A., Moresino F., Vielle M. and Viguier L., *A Coupled Equilibrium Model of International Climate Policy*, University of Geneva, LOGILAB Technical report, August 2005.

[8] Haurie A. and Viguier L., A stochastic dynamic game of carbon emissions trading, *Environmental Modeling & Assessment*, 8, 239–248, 2003.

[9] Haurie A. and Zaccour G., Differential game models of global environmental management, in C. Carraro and J.A. Filar (eds.), *Control and Game-Theoretic Models of the Environment*, Annals of the International Society of Dynamic Games, 3–23, Birkhäuser, Boston, 1995.

[10] Helm C., International emissions trading with endogenous allowance choices, *Journal of Public Economics*, 87, 2737–2747, 2003.

[11] Krawczyk J., Coupled constraint Nash equilibria in environmental games, *Resource & Energy Economics*, 27, 2, 157–181, 2005.

[12] Labriet M. and Loulou R., Coupling climate damages and GHG abatement costs in a linear programming framework, *Environmental Modeling & Assessment*, 8, 261–274, 2003.

[13] Nordhaus W.D., *Managing the Global Commons: The Economics of Climate Change*, MIT Press, Cambridge, MA, 1994.

[14] Nordhaus W.D. and Yang Z., A regional dynamic general-equilibrium model of alternative climate change strategies, *American Economic Review*, 86, 4, 741–765, 1996.

[15] Ramsey F., A mathematic theory of saving, *Economic Journal*, 38, 543–549, 1928.

[16] Rosen J.B., Existence and uniqueness of the equilibrium points for concave n-person games, *Econometrica*, 33, 3, 520–534, 1965.

[17] Scheffran J., Economic growth, emission reduction and the choice of energy technology in a dynamic-game framework, in P. Chamoni et al., *Operations Research Proceedings 2001*, Heidelberg, 329–336, Springer, 2002.

[18] Scheffran J., Interaction in climate games: The case of emissions trading,

in J. Geldermann and M. Treitz (eds.), *Entscheidungstheorie und -praxis in industrieller Produktion und Umweltforschung*, 1–18, Shaker, Aachen, 2004.

A Stochastic Multigeneration Game for Global Climate Change Impact Assessment

Alain Haurie
Logilab-HEC, University of Geneva
CH-1211 Geneva 4, Switzerland
alain.haurie@hec.unige.ch

Abstract

This paper deals with a definition of intergenerational equity in a stochastic game formulation. We propose a piecewise deterministic control model where the control is exerted by a succession of generations, each having a random life duration. Each generation has a concern in both the expected reward received during its own life and the expected reward that will accrue to the next generation when it will take control at the end of the present generation's life. An intergenerational equilibrium is defined. The model is then specialized to the case of exponential random life duration and stationary state equations. A complete characterization of the equilibrium solution is proposed through the use of a family of auxiliary infinite horizon control problems. A numerical approximation method is proposed. The model and the equilibrium concepts are then used in the context of an integrated assessment model of global climate change impacts.

1 Introduction

This paper proposes a game-theoretic approach to deal with the thorny issue of discounting and intergenerational equity in cost-benefit analyses for very long-lived projects. We base our approach on an interpretation of the discounting process as a representation of a random life duration for the economic agent, described as an exponential random variable with mean $1/\rho$ where ρ is the discount rate. In our approach we explore the equilibrium solution to a continuous-time piecewise deterministic game where players represent successive generations, each having a random life duration. The players control a deterministic system and derive a reward at a rate depending on state and control. One assumes that the payoff to one generation is a function of the expected sum of rewards to this generation and the expected sum of rewards to the next generation. This introduces a form of altruism in the behavior of the players. One defines and characterizes the intergenerational equilibrium solution to this game when the system is stationary and the life duration is described by an exponential random variable, hence the link with the discounting schemes. A computational method is proposed for the stationary case.

The approach is then applied to a model of integrated assessment of climate change. The concern about the economic impacts of global climate change gives a new impetus to the search for a rational way to balance current generation and future generation welfare gains when deciding about policies that will have long-lived consequences (see articles [4,9,11,12] and the informative book [34]). The stochastic game model defined above is used to propose an approach that provides a rationale for valuing time in long-term policies without falling into the trap of time inconsistency. The solution concept that is proposed is fully time consistent and even subgame-perfect (more precisely Markov-perfect). This approach could be compared with the one where the players in a stochastic game use a convex combination of discount rates as proposed by Filar and Vrieze [17,18] and recently surveyed by Feinberg and Schwartz [15]. The main difference lies in the consideration of a sequence of generations playing the game, instead of a set of players who are in the same generation and compete in a dynamic game. Also, our approach leads to a time-consistent amd Markov-perfect equilibrium, whereas the games in [18] and [15] lead to equilibria that are not time consistent. From an economic point of view, the approach proposed here introduces an extra valuation for the state variables, i.e., the capital stocks and the environmental state variables, which takes into account the fact that they will impact the welfare of forthcoming generations.

The paper is organized as follows. In Section 2 one briefly recalls the important debate among economists, concerning the proper way to represent a society's concern toward the welfare of future (unborn) generations. In Section 3 a general multigeneration stochastic game model is defined, and the concept of intergenerational equilibrium is introduced. In Section 4 the stationary case is studied in detail and a numerical approximation scheme is proposed. In Section 5 this characterization is applied to a reduced form of an integrated assessment model proposed in [30], and the solution obtained is compared with those coming from a standard cost-benefit analysis. In conclusion, the contribution of this game formulation to the altruism/time discounting debate is discussed.

2 Altruism and the Time Discounting Issue

When dealing with economic choices over time, economic and finance theories introduce discounting functions that are used to represent "pure time preference" by the economic agents. In continuous-time models it is common to use an exponential discounting term $e^{-\rho t}$ to represent the marginal rate of substitution between consumption in year t and consumption now (at $t = 0$).

It is well known that the discounting process can be given an interpretation in terms of a stochastic process as indicated thereafter. Assume that an economic agent will receive a stream of income $c(t) : t \in [0, T]$, where T is a random time horizon characterized by an exponential law of parameter ρ (recall that an exponential random variable has the half line $[0, \infty)$ as support with a distribution function $\mathrm{P}[T \leq t] = 1 - e^{-\rho t}$). The expected value of T is then $1/\rho$. The expected total income is defined as

$$\mathcal{C} = \mathrm{E}\left[\int_0^T c(t)\, dt\right].$$

Now, since the elementary probability of T being in the interval $\theta, \theta + d\theta$ is given by $\rho e^{-\rho\theta}\, d\theta$, we can rewrite $\mathcal{C}$ as

$$\mathcal{C} = \int_0^\infty \left(\int_0^\theta c(t)\, dt\right) \rho e^{-\rho\theta}\, d\theta.$$

Integrate by parts to obtain

$$\mathcal{C} = \int_0^\infty e^{-\rho t} c(t)\, dt,$$

which is the discounted value of the infinite stream of income. The discount rate ρ is therefore associated with uncertainty about the duration of the consumption period (i.e., the random duration of the agent life). The parameter ρ is also called the *killing rate* since $\rho\, dt$ is the elementary probability that a "death occurs" in the elementary time interval $[t, t + dt]$, given that the agent has survived up to time t. A discount rate of 5% corresponds to a random life duration with expected value $1/0.05 = 20$ years. According to this interpretation, discounting occurs because of the finite life of economic agents. It is because we are not sure of still being around in 20 years from now that we discount heavily our consumption in this distant future. The use of a zero discount rate would correspond to the consideration of an infinitely lived species. In such a case the early consumption could be sacrificed in the perspective of higher long-term consumptions.

Discounting plays a crucial role in cost-benefit analysis and therefore in the selection of investment projects (see, e.g., [3]). Economists have recognized early that some long-lived projects should be justified on another basis that would take into account both the finiteness of the economic agents' life and their desire to leave to the next generation a valuable bequest. This need for a rationale to decide on very long-lived investments has become an important issue in the current debate concerning the assessment of global climate change economic impacts. There is a growing consensus about the anthropogenic global climate change (GCC) induced mainly by the emissions of greenhouse gases (GHGs) due to fossil fuel energy uses, industry and agriculture activities. To cope with anthropogenic GCC, economies can rely on abatement, mitigation and adaptation. An interesting aspect of the problem is that abatement decided by the current generation, if any, will be made for the benefit of generations in a rather distant future and for the benefit of populations that are not currently the principal emitters. In such a context, using a relatively high discount rate will tend to minimize the long-term consequences (welfare losses) due to climate change; however, using a zero discount will penalize unduly the current generation who will invest heavily for the benefit of their descendants.

There have been a variety of attempts to define axiomatic foundations of sustainability [11,12], and to introduce time-dependent discount rates (the discount rate tending to 0 when $t \to \infty$) [1,13,40,41]. The consequences of non-exponential discounting were first studied in [38]. The problem with many of these attempts is that the solution concept is not time consistent. It implies that the present generation is able to force the future generations to use a lower discount rate than the present one. If the present generation is not perfectly altruistic it has to recognize that the future generation will also behave with this limited altruism. The solution will then take the form of an intergenerational equilibrium in a dynamic game controlled by a succession of generations. There have been attempts to represent (imperfect) altruism in multigeneration games. In [33] an economic growth model has been proposed where the current generation controls only the initial period decision but has a vested interest in what happens to all the other forthcoming generations. This model has been generalized to a stochastic game framework in [2]. More recently, similar models have been studied in [28]. These ideas have received a new impetus in recent papers dealing with the discounting issue for global warming assessment [5,14]. In these papers, the deterministic economic growth format of [33] is used. More recently, an interesting contribution [37] has proposed a way to construct time-consistent economic plans in a discrete-time model with quasi-geometric discounting and a generation (player) for each time period. Most of the work on non-exponential discounting has been done in discrete-time models. There is a recent contribution, using continuous time and exponentially distributed life duration [21], but these authors do not introduce a stochastic game similar to the one we propose here.

In this chapter we propose a continuous-time model that includes intergenerational altruism and we characterize a Markov-perfect equilibrium solution between players corresponding to the successive generations. The model is based on a stochastic interpretation of the discounting process. It implies both time-consistency and Markov-perfectness of the equilibrium solution that is defined.

3 A Multigeneration Game Model

In this section a general multigeneration game over a controlled dynamical system is introduced and an intergenerational equilibrium is defined. This dynamic game structure will be used in Section 4 to introduce intergenerational concern in the economic discounting schemes.

Consider a dynamical system defined by the state equations

$$\dot{x}(t) = f(t, x(t), u(t)) \tag{1}$$

$$u(t) \in U(t, x(t)) \tag{2}$$

$$x(t^k) = x^k. \tag{3}$$

A reward rate $L(t, x(t), u(t))$ is associated with the state and control variables. The functions $f(\cdot,\cdot,\cdot)$, $L(\cdot,\cdot,\cdot)$ and $U(\cdot,\cdot)$ satisfy the usual smoothness assumptions used for optimal control (see [10] p. 9, for details).

The system is controlled by a succession of players, each one representing a generation denoted $k \in \mathcal{N} = \{1, 2, \ldots, N\}$. The life duration of generation k is a continuous random variable θ_k with support $[0, \infty)$ with expected value $1/\rho_k$, $\rho_k \geq 0$.

The state of the system is represented by $s = (t, x)$, where $t \geq 0$, $\zeta \in \mathcal{N}$ and $x \in X \subset \mathbb{R}^n$. When the game begins, at $t^1 = 0$, generation 1 is in control. After a random lifetime θ_1 generation 1 dies and leaves control to generation 2, etc.

Definition 3.1. A strategy for generation k is a mapping γ_k from $[0, \infty) \times X$ into the class of mappings $w(\cdot) : [0, \infty) \to U$. When generation k takes control of the system at time t^k, with initial state $x^k = x(t^k)$, then its control will be $u(t) = w(t - t^k)$ where $w(\cdot) = \gamma[t^k, x^k]$. This control generates a trajectory $x(\cdot) : [t^k, \infty) \to X$, with $x(t^k) = x^k$. The control generated is admissible if it satisfies $u(t) \in U(x(t))$.

The payoff to generation $k = 1, \ldots, N-1$, when it takes control at time $t^k = \sum_{\ell=1}^{k-1} \theta_\ell$, with state x^k is defined as

$$\begin{aligned} V_k[s^k; \gamma_k, \gamma_{k+1}] \\ = \alpha \mathrm{E}_{\gamma_k} \left[\int_{t^k}^{t^k+\theta_k} L(t, x(t), u(t))\, dt | s^k = (t^k, x^k) \right] \\ + (1-\alpha) \mathrm{E}_{\gamma_{k+1}} \left[\int_{t^k+\theta_k}^{t^k+\theta_k+\theta_{k+1}} L(t, x(t), u(t))\, dt | (t^k + \theta_k, x(t^k + \theta_k)) \right], \end{aligned} \tag{4}$$

where α (resp. $1-\alpha$) $\in [0, 1]$ is a relative weight given to the present (resp. future) generation. This represents a situation where each generation has a vested interest in its own and immediate successor rewards.[1] In the preceding

[1]Indeed one may also consider situations where each generation has an interest in what will happen to grandchildren and great-grandchildren, etc. In a companion paper [25] we consider such an extension. In this chapter we keep the simplest form of an intergenerational dependence as it is often supposed in overlapping generation models in economic literature. See also Remark 3.4 below.

expressions the state trajectory $x(\cdot) : [t^k, t^k + \theta_k] \to \mathbb{R}^n$ is a solution of (1)–(3) with initial state x^k, induced by the control generated by γ_k at time t^k; the rest of the trajectory $x(\cdot) : [t^k + \theta_k), [t^k + \theta_k + \theta_{k+1}) \to \mathbb{R}^n$ is a solution of (1)–(3) with initial state $x(t^k + \theta_k)$, induced by the control generated by γ_{k+1} at time $t^{k+1} = t^k + \theta_k$.

The total duration of this game is random and given by $\sum_{k=1}^{N} \theta_k$. The payoff of the Nth generation is defined as

$$V_N[s^N; \gamma_N] = \mathrm{E}_{\gamma_N} \left[\int_{t^N}^{t^N+\theta_N} L(t, x(t), u(t))\, dt | s^N \right] \tag{5}$$

since there is no descendent.

Definition 3.2. An intergenerational equilibrium is a sequence $\underline{\gamma}^* = (\gamma_k^* : k \in \mathcal{N})$ such that

$$\forall k \in \mathbb{N}, \quad \forall s^k \in S, \quad V_k(s^k; \gamma_k^*, \gamma_{k+1}^*) = \max_{\gamma_k} V_k(s^k; \gamma_k, \gamma_{k+1}^*). \tag{6}$$

According to this definition, each generation reacts optimally to the strategy that will be used by the next generations. Because each generation has a vested interest in what happens to the next generation, this equilibrium concept links all the generations together.

Remark 3.1. In this game structure the current generation determines the initial point of the trajectory controlled by the next one. This is where the coupling between the games played by successive generations occurs.

Remark 3.2. The conditions for existence of an equilibrium for such a game should not be too difficult to meet. In fact, it would suffice to have the possibility of implementing a dynamic programming method where the value functions

$$V_N^*(t, x) = V_N(t, x; \gamma_N^*) = \max_{\gamma_N} V_N(t, x; \gamma_N) \tag{7}$$

and

$$\begin{aligned} V_k^*(t, x) \ &= \ V_k(t, x; \gamma_k^*) = \max_{\gamma_k} \alpha \mathrm{E}_{\gamma_k} \left[\int_t^{t+\theta_k} L(\tau, x(\tau), u(\tau))\, d\tau \right] \\ &\quad + (1 - \alpha)\Phi_{k+1}^*[t + \theta_k, x(t + \theta_k)], \end{aligned} \tag{8}$$

where we have denoted

$$\begin{aligned} &\Phi_{k+1}^*[t + \theta_k, x(t + \theta_k)] \\ &\quad = \mathrm{E}_{\gamma_{k+1}^*} \left[\int_{t+\theta_k}^{t+\theta_k+\theta_{k+1}} L(\tau, x(\tau), u(\tau))\, d\tau | (t + \theta_k, x(t + \theta_k)) \right] \end{aligned} \tag{9}$$

for $k = N - 1, \ldots, 1$, are computed recursively and keep enough regularity for the maximum strategy to exist at each step.

Remark 3.3. It is conceptually possible to extend the equilibrium definition to the case of an infinite number of generations ($\mathcal{N} = \mathbb{N}$). The existence issue will be more difficult to settle, though. However, in the remainder of the chapter we focus on the case of an infinite number of generations, but in a stationary environment.

Remark 3.4. A more general payoff form could be defined as follows:

$$V_k[s^k; \gamma_k, \tilde{\gamma}_{k+}] = \sum_{\ell \geq k} \alpha(k,\ell) \mathrm{E}_{\gamma_\ell} \left[\int_{t^\ell}^{t^\ell + \theta_\ell} L(t, x(t), u(t))\, dt | x^\ell = x(t^\ell) \right]. \quad (10)$$

In the above expressions the state trajectory $x(\cdot)$ is a solution of (1)–(3) with initial state x^k at initial time t^k, when the control is generated by the strategies γ_ℓ. The weights $\alpha(k,\ell)$, $\ell > k$, reflect the concern of generation k for the welfare of generation ℓ. In [25,26] this type of multigeneration game is studied, in particular when $\alpha(k,\ell) = \beta^{\ell-k}$, with $\beta < 1$. In this chapter we keep the simplest form of intergenerational concern, where each generation has a vested interest in the immediately following one.

Definition 3.2 is very general, with a random life duration that can be distributed in many different ways. In the rest of the chapter we shall specialize to the case of exponentially distributed life durations, as they are linked with the discounting process. We shall also restrict our analysis to the case of a stationary system, leaving for further exploration the time-varying case.

4 The Stationary Exponential Case

In this section one assumes an infinite number of generations ($k \in \mathbb{N}$) and a life duration of generation k defined as an exponentially distributed random variable θ_k with expected value $1/\rho$. One also assumes that the controlled dynamical system is stationary,

$$\begin{aligned} \dot{x}(t) &= \tilde{f}(x(t), u(t)) && (11)\\ u(t) &\in \tilde{U}(x(t)) && (12)\\ x(t^k) &= x^k, && (13) \end{aligned}$$

with reward rate $\tilde{L}(x(t), u(t))$.

4.1 Interpretation as a Time-Varying Discount Rate

It is interesting[2] to look at the problem from the point of view of a generation if it could commit the behavior of the next generation and so decide a trajectory

[2] This interpretation has been kindly provided by an anonymous reviewer.

$x(\cdot) : [0,\infty) \to X$ generated by a control $u(\cdot) : [0,\infty) \to U(x(t))$ with the reward flow $\tilde{L}(x(\cdot), u(\cdot))$. Since

$$\mathrm{E}\left[\int_0^{\theta} \tilde{L}(x(t),u(t))\,dt\right] = \int_0^{\infty} e^{-\rho t}\tilde{L}(x(t),u(t))\,dt \tag{14}$$

and similarly

$$\mathrm{E}\left[\int_{\theta}^{\theta+\theta'} \tilde{L}(x(t),u(t))\,dt \,\middle|\, \theta\right] = e^{\rho\theta}\int_{\theta}^{\infty} e^{-\rho t}\tilde{L}(x(t),u(t))\,dt, \tag{15}$$

we obtain readily by applying the law of iterated expectations and changing the order of integration once more

$$\mathrm{E}\left[\int_{\theta}^{\theta+\theta'} \tilde{L}(x(t),u(t))\,dt\right] = \int_0^{\infty} \rho t e^{-\rho t}\tilde{L}(x(t),u(t))\,dt. \tag{16}$$

Therefore,

$$\begin{aligned}\alpha\mathrm{E}\left[\int_0^{\theta} \tilde{L}(x(t),u(t))\,dt\right] &+ (1-\alpha)\mathrm{E}\left[\int_{\theta}^{\theta+\theta'} \tilde{L}(x(t),u(t))\,dt\right]\\ &= \int_0^{\infty} [\alpha+(1-\alpha)\rho t]e^{-\rho t}\tilde{L}(x(t),u(t))\,dt. \end{aligned}\tag{17}$$

So if a generation could commit the control and trajectory for itself and the next generation, it would obtain a utility which is the discounted sum of rewards, with a discount factor at time t given by

$$d_\alpha(t) = [\alpha+(1-\alpha)\rho t]e^{-\rho t}. \tag{18}$$

This corresponds to a time-dependent discount rate

$$r_\alpha(t) = -\frac{d'_\alpha(t)}{d_\alpha(t)} = \frac{\rho[(1-\alpha)\rho t + (2\alpha-1)]}{\alpha+(1-\alpha)\rho t}. \tag{19}$$

We notice that $r(t)$ increases with t and reaches asymptotically the value ρ. At $t=0$ one has $r_\alpha(0) = \rho(2\alpha-1)/\alpha$. So for $\alpha = 0.5$ one has $r_{0.5}(0) = 0$ and for $\alpha < 0.5$ one observes $r_\alpha(0) < 0$.

It is interesting to notice that introducing a concern for the welfare of the next generation means an increasing discount rate. The asymptotic value ρ is easily understood, since, when t increases the probability that it is the second generation which is in charge tends to one, and when the second generation is in charge, the problem boils down to the ρ-discounted optimization.

The system is controlled by a sequence of sovereign decision makers, who use the $r_\alpha(t)$ time-varying discount rate, and have exponentially distributed lifetimes. As each generation cannot commit the control used by the next one, the control policy is the result of an intergenerational equilibrium. *The first generation has a vested interest in the rewards accrued to the second generation, knowing that this second generation will have an interest in the reward accrued to the third, etc.* As indicated before, the equilibrium concept links all generations together because of the overlapping structure of the generations' rewards. In the remainder of this section we characterize the equilibrium solution.

4.2 Stationary Dynamic Programming Equations

Because of time stationarity, the problem will have the same structure for each generation and one can define a generic generational payoff as

$$\begin{aligned}\tilde{V}(x^o;\gamma,\gamma') = \alpha \mathrm{E}_\gamma &\left[\int_0^\theta \tilde{L}(x(t),u(t))\,dt|x^o\right]\\ &+ (1-\alpha)\mathrm{E}_{\gamma'}\left[\int_\theta^{\theta+\theta'} \tilde{L}(x(t),u(t))\,dt|x(\theta)\right], \qquad (20)\end{aligned}$$

where γ, resp. γ', is the current (resp. next) generation strategy, θ (resp. θ') is the life duration of the present (resp. next) generation, x^o is the initial state and $x(\cdot)$ is the trajectory solution to the state equations (11)–(13) induced by these strategies.

Since we aim at implementing a numerical method we have to content ourselves with an approximate equilibrium solution defined as follows.

Definition 4.1. A stationary ε-equilibrium is defined as a strategy γ^* that satisfies

$$\forall x \in X, \quad \tilde{V}(x;\gamma^*,\gamma^*) \geq \max_\gamma \tilde{V}(x;\gamma,\gamma^*) - \varepsilon. \qquad (21)$$

A stationary equilibrium corresponds to the case where $\varepsilon = 0$.

Theorem 4.1. *A stationary intergenerational equilibrium is characterized by the following equations:*

$$\gamma^*(x^o) = \operatorname*{argmax}_{u(\cdot)} \left[\int_0^\infty e^{-\rho t}\left(\alpha\tilde{L}(x(t),u(t)) + \rho(1-\alpha)\tilde{\Phi}^*(x(t))\right)dt|x(0)=x^o\right] \qquad (22)$$

$$\tilde{\Phi}^*(x^o) = \mathrm{E}_{\gamma^*}\left[\int_0^\theta \tilde{L}(x(t),u(t))\,dt|x(0)=x^o\right]. \qquad (23)$$

Proof. Define

$$\tilde{\Phi}^*(x^o) = \mathrm{E}_{\gamma^*}\left[\int_0^\theta \tilde{L}(x(t),u(t))\,dt|x(0)=x^o\right]. \qquad (24)$$

Since the elementary probability of the interval $[\theta, \theta + d\theta)$ is given by $\rho e^{-\rho\theta}\,d\theta$, using Equation (20) one can write

$$\tilde{V}(x^o; \gamma^*, \gamma^*) = \int_0^\infty \left(\int_0^\theta \alpha \tilde{L}(x^*(t), u^*(t))\,dt + (1-\alpha)\tilde{\Phi}^*(x^*(\theta)) \right) \rho e^{-\rho\theta}\,d\theta,$$

where $x^*(\cdot)$ is the trajectory generated at initial state x^o by the control $u^*(\cdot)$ defined by the strategy γ^*. Integrate by parts the first integral to finally obtain

$$\tilde{V}(x^o; \gamma^*, \gamma^*) = \alpha \int_0^\infty e^{-\rho t}\tilde{L}(x^*(t), u^*(t))\,dt + (1-\alpha)\int_0^\infty \Phi^*(x^*(t))\rho e^{-\rho t}\,dt. \tag{25}$$

$$= \int_0^\infty e^{-\rho t}\left(\alpha \tilde{L}(x^*(t), u^*(t)) + \rho(1-\alpha)\tilde{\Phi}^*(x^*(t)) \right) dt. \tag{26}$$

It suffices to write the equilibrium conditions to obtain the conditions (22), (23). □

There is an existence issue, for a stationary equilibrium, that will not be completely addressed in this chapter.[3] In [2] it has already been shown how the existence proof of approximate equilibria, given by Whitt [42] for stochastic sequential games and based on the Kakutani fixed-point theorem, can be adapted to the framework of multigeneration stochastic games. However this proof needs to be adapted to the infinite-dimensional state space context, which requires a heavy mathematical apparatus. We shall proceed differently and use a constructive numerical method to approximate an equilibrium solution. Obtaining a numerical solution satisfying the ε-equilibrium solution conditions will then provide a "local" existence proof.

4.3 The Associated Implicit Infinite Horizon Control Problem

Theorem 4.1 characterizes the equilibrium strategy of a generation as the solution of a family of auxiliary infinite horizon ρ-discounted optimal control problems having an implicitly defined reward function

$$\mathcal{L}^*(x(t), u(t)) = \alpha \tilde{L}(x(t), u(t)) + \rho(1-\alpha)\tilde{\Phi}^*(x(t)). \tag{27}$$

We call it "implicit" because the function $\tilde{\Phi}^*(x(t))$ is itself defined by the very strategy γ^* characterized by the solution of this control problem. There is a fixed-point property which is very typical of equilibrium characterizations. The auxiliary infinite horizon control problem is defined as

$$\max \quad \int_0^\infty e^{-\rho t}\mathcal{L}^*(x(t), u(t))\,dt \tag{28}$$

[3] In [25] an existence proof is given for a particular case where each generation considers the welfare of all coming generations k with a geometrically declining weight β^k.

s.t.

$$\begin{aligned} \dot{x}(t) &= \tilde{f}(x(t), u(t)) & (29)\\ u(t) &\in \tilde{U}(x(t)) & (30)\\ x(0) &= x^o. & (31) \end{aligned}$$

4.4 Turnpikes and the limit cases when $\alpha \to 1$ or $\rho \to 0$

Under sufficient concavity and curvature assumptions (see [10], Chap. 6 for details) the infinite horizon optimal control problem (28)–(31) may have an attractor $\bar{x}$ common to all trajectories, emanating from different initial states x^o. These concavity properties should be verified by the extended reward function $\alpha\tilde{L}(x, u) + (1 - \alpha)\rho\tilde{\Phi}^*(x)$. As $\tilde{\Phi}^*(x)$ is a value function associated with the equilibrium policy as defined in (24), it is not easy to give conditions under which this function is concave. However, as will be observed in Section 5 dealing with our numerical experiment, we can obtain (for some interesting problems) a good approximation of $\tilde{\Phi}^*(x)$ by an affine function. In that case the concavity property of $\tilde{L}(x, u)$ would also impose concavity for the extended function and we can expect the turnpike property to hold. This attractor is a trajectory steady state that solves the following implicit programming problem (as shown in [16]):

$$\max \mathcal{L}^*(x, u) = \alpha\tilde{L}(x, u) + (1 - \alpha)\rho\tilde{\Phi}^*(x) \qquad (32)$$

s.t.

$$\begin{aligned} 0 &= f(x, u) - \rho(x - \bar{x}) & (33)\\ u &\in U(x), & (34) \end{aligned}$$

where $\bar{x}$ is the turnpike itself. When the selfishness parameter α tends to 1, the above turnpike will tend to coincide with the one associated with the usual discounted reward optimal control problem.

$$\max \tilde{L}(x, u) \qquad (35)$$

s.t.

$$\begin{aligned} 0 &= f(x, u) - \rho(x - \bar{x}) & (36)\\ u &\in U(x). & (37) \end{aligned}$$

So, in terms of asymptotic behavior of the optimal trajectories, altruism is modifying the attractor by replacing the optimized reward $L(x, u)$ with a modified reward $\mathcal{L}(x, u) = \alpha L(x, u) + (1 - \alpha)\rho\tilde{\Phi}^*(x)$. We can perform a static comparative analysis to assess the impact of this modification on the asymptotic steady state.

When $\rho \to 0$ one may expect the expression $\rho\,\tilde{\Phi}^*(x)$ to tend toward g^* which is the maximal sustainable reward, solution of

$$g^* = \max \tilde{L}(x, u) \qquad (38)$$

s.t.

$$\begin{aligned} 0 &= f(x,u) && (39)\\ u &\in U(x). && (40) \end{aligned}$$

Therefore, the limit steady-state problem will be

$$\max \mathcal{L}^*(x,u) = \alpha \tilde{L}(x,u) + (1-\alpha)g^* \tag{41}$$

s.t.

$$\begin{aligned} 0 &= f(x,u) && (42)\\ u &\in U(x). && (43) \end{aligned}$$

Clearly, as g^* is a constant, this problem admits the same solution as the maximal sustainable reward defined in (35)–(37). So, when the discount (killing) rate $\rho \to 0$ the optimal trajectory is similar to the one associated with the infinite horizon, undiscounted control problem.

4.5 Strategy Synthesis

We may represent the optimal strategy as a feedback loop by solving the auxiliary control problems by a dynamic programming approach. For that purpose we introduce the notation

$$\tilde{V}^*(x) = \tilde{V}(x;\gamma^*,\gamma^*) \tag{44}$$

to represent the equilibrium value function for the current generation. Now assuming regularity for this value function and using standard dynamic programming arguments (see, e.g., [19]) one can characterize the value function as the solution to the algebraic equation

$$\rho \tilde{V}^*(x) = \max_{u \in \tilde{U}(x)} \left\{ \mathcal{L}^*(x,u) + \frac{\partial}{\partial x}\tilde{V}^*(x)\tilde{f}(x,u) \right\}, \tag{45}$$

which may also be written as

$$\rho \tilde{V}^*(x) = \max_{u \in \tilde{U}(x)} \left\{ \alpha \tilde{L}(x,u) + (1-\alpha)\rho\tilde{\Phi}^*(x) + \frac{\partial}{\partial x}\tilde{V}^*(x)\tilde{f}(x,u) \right\}. \tag{46}$$

4.6 Time Consistency and Markov-Perfectness

The multigeneration equilibrium solution concept is time consistent. If one restarts the whole game at any time t and state $x(t)$ the same strategies will be used and the same trajectory will continue to be optimal for the current generation. When one synthesizes the control used by a generation in a synthesized way (i.e., when one finds an optimal feedback law), the solution is even Markov-perfect, in the sense where the equilibrium strategies remain the same even if for some time a generation has not played correctly.

4.7 Numerical Approximation

To numerically solve the stationary intergenerational equilibrium problem one has to approximate the value function $\tilde{V}^*(x)$ and the associated optimal control strategy γ^*, as well as the second generation reward function $\tilde{\Phi}^*(x)$. These evaluations are coupled since $\tilde{V}^*(x)$ depends on $\tilde{\Phi}^*(x)$ which, in turn, depends on γ^*. A natural way to proceed is to implement a cobweb approach where, for a given candidate for $\tilde{\Phi}^*(x)$, one computes the associated $\tilde{V}^*(x)$ and γ^* by solving a dynamic programming optimization problem and, for a given candidate for the strategy γ^*, one updates the second generation reward function $\tilde{\Phi}^*(x)$. At both stages one can use the Dupuis-Kushner scheme, fully described in the book [27], to approximate the value functions via the solution of the DP equations of an associated approximating Markov decision process (MDP). For that purpose one uses a finite grid $\mathcal{G}$ with nodes $g \in \mathcal{G}$. For each component $i = 1, \dots, n$ of x one defines $x_i^{\min}$, $x_i^{\max}$, with $x_i^{\max} - x_i^{\min} = N\,h$, where h is the mesh size and N is the number of sampled values on each axis. So, to each node $g \in \mathcal{G}$ there corresponds a discretized (or sampled) state value x^g with components x_i^g given by $x_i^g = x_i^{\min} + n_i^g\,h$, $n_i^g \in \{0, 1, \dots, N\}$.

4.7.1 Evaluating the Equilibrium Strategy

For a candidate second generation reward function $\tilde{\Phi}(x)$ one solves, via dynamic programming, the auxiliary control problem. In Equation (46), one approximates the partial derivative $(\partial/\partial x_i)\tilde{V}^*(x)$ by finite differences taken in the direction of the flow, that is:

$$\frac{\partial}{\partial x_i}\tilde{V}^*(x) \to \begin{cases} (\tilde{V}^*(x + e_i h) - \tilde{V}^*(x))/h & \text{if} \quad \tilde{f}_i(x,u) \geq 0 \\ (\tilde{V}^*(x) - \tilde{V}^*(x - e_i h))/h & \text{if} \quad \tilde{f}_i(x,u) < 0 \end{cases} \tag{47}$$

where e_i is the unit vector of the ith axis. Define

$$\begin{aligned} \tilde{f}_i^+(x,u) &= \max\{0, \tilde{f}_i(x,u)\} \\ \tilde{f}_i^-(x,u) &= \max\{0, -\tilde{f}_i(x,u)\}. \end{aligned}$$

Substituting the differences to the partial derivatives in Eq. (46), one obtains

$$\begin{aligned} \rho\tilde{V}^*(x) = \max_{u \in \tilde{U}(x)} \Bigg\{ &\alpha\tilde{L}(x,u) + (1-\alpha)\rho\tilde{\Phi}^*(x) \\ &+ \sum_{i=1}^{n} \Bigg(\frac{(\tilde{V}^*(x + e_i h) - \tilde{V}^*(x))}{h}\, \tilde{f}_i^+(x,u) \\ &+ \frac{(\tilde{V}^*(x) - \tilde{V}^*(x - e_i h))}{h}\, \tilde{f}_i^-(x,u) \Bigg) \Bigg\}. \end{aligned} \tag{48}$$

$$0 = \max_{u\in\tilde{U}(x)} \Bigg\{ (\alpha\tilde{L}(x,u) + (1-\alpha)\rho\tilde{\Phi}^*(x))h \\ - \tilde{V}^*(x)\left(\rho\, h + \sum_{i=1}^{n}(\tilde{f}_i^+(x,u) + \tilde{f}_i^-(x,u))\right) \\ + \sum_{i=1}^{n}(\tilde{f}_i^+(x,u)\tilde{V}^*(x+e_ih) + \tilde{f}_i^-(x,u)\tilde{V}^*(x-e_ih)) \Bigg\}. \tag{49}$$

Define the interpolation interval

$$\Delta_h = \frac{h}{\rho\, h + \sum_{i=1}^{n}(\tilde{f}_i^+(x,u) + \tilde{f}_i^-(x,u))}. \tag{50}$$

One considers an MDP with discrete states x^g, $g \in \mathcal{G}$ and control $u \in \tilde{U}(x^g)$. The transition rewards are given by $(\tilde{L}(x^g,u) + \rho\tilde{\Phi}^*(x^g))\Delta_h$. The transition probabilities $\Pi_i(x^g, x^{g'}, u)$ are defined as follows:

- When $g \in \mathcal{G}\setminus\partial\mathcal{G}$ the transition probabilities x^g to any neighboring sampled value $x^g \pm e_ih$ are given by

$$\pi_i^{\pm}(x^g,u) = \frac{\tilde{f}_i^{\pm}(x^g,u)}{\sum_{i=1}^{n}(\tilde{f}_i^+(x^g,u) + \tilde{f}_i^-(x^g,u))}.$$

- On the boundary $\partial\mathcal{G}$ of the grid, the probabilities are defined according to a reflecting boundary scheme.
- All the other transition probabilities are 0.

A discounting term is defined by

$$\beta(x^g,u) = \sum_{i=1}^{n}(\tilde{f}_i^+(x^g,u) + \tilde{f}_i^-(x^g,u))\frac{\Delta_h}{h}.$$

The DP equations for this approximating MDP are given by

$$v(x^g) = \max_{u\in\tilde{U}(x^g)} \Bigg\{ (\alpha\tilde{L}(x^g,u) + (1-\alpha)\rho\tilde{\Phi}^*(x^g))\Delta_h \\ + \beta(x^g,u)\sum_{g'\in\mathcal{G}}\Pi_i(x^g,x^{g'},u)v(x^{g'}) \Bigg\}, \tag{51}$$

where we use the general notation $\Pi_i(x^g, x^{g'}, u)$ to describe the transition probabilities of the MDP defined above.

4.7.2 Evaluating the Second Generation Reward Function

The solution of (51) defines an optimal policy denoted u^{g*}. A new second generation reward function $\tilde{\Phi}^*(x^g)$ is then obtained by solving

$$\tilde{\Phi}^*(x^g) = \tilde{L}(x^g, u^{g*})\Delta_h + \beta(x^g, u^{g*}) \sum_{g' \in \mathcal{G}} \Pi_i(x^g, x^{g'}, u^{g*})\tilde{\Phi}^*(x^{g'}). \tag{52}$$

To solve both DP equations (51) and (52) one uses a linear programming formulation, as indicated in [35] p. 223.

It is well established that the approximating MDPs will lead to estimates of the value functions that converge weakly toward the continuous-time solutions of the DP equations. Using the classical verification theorems of DP, we conclude that the equilibrium solutions obtained from the approximating MDPs provide ε-equilibrium solutions to the continuous-time game.

5 Application to Integrated Assessment of Global Climate Change

As already indicated in Section 2, the cost-benefit analysis of global climate change mitigation policies poses acutely the question of intergenerational solidarity. In this section the intergenerational equilibrium criterion is applied to a well-known integrated assessment model, and the results are interpreted.

5.1 A Reduced Model Based on DICE94

An integrated model describes the dynamic interplay between the climate system, economy and polity. GHGs are emitted as a by-product of economic activity. A reduction of emissions can be obtained, using our current technology, at a cost that corresponds to a loss of product for consumption or investment purposes. Emissions increase concentrations of GHGs in the earth's atmosphere which defines a forcing effect on the earth's surface temperature. This temperature increase may have a negative economic effect, also expressible as a loss of economic product. These interactions have been nicely summarized by Nordhaus in the models DICE94 [30] and DICE99 [31], which are designed for performing a cost-benefit analysis (CBA) and assessing GCC economic policies. These models are extensions of the classical Ramsey optimal economic growth paradigm [36]. The economy produces a single homogenous good that can be either consumed, with instantaneous utility $\log(c(t))$ for per capita consumption $c(t)$, or invested to obtain more physical capital. The economy produces the good using two factors, labor and capital. There is an exogenous population growth process and an exogenous technical progress process. The production of the economic good generates GHG emissions. An abatement activity can be harnessed with a cost measured in product losses. The emissions generate atmospheric concentrations. Part of these concentrations are captured by the oceans. The temperature increases due to the greenhouse forcing and this generates an economic damage, also expressed as production losses. In references [30] and [31] this integrated assessment model is fully explained. We will summarize this approach in the following, very simplified, continuous-time dynamical system. The variables and parameters are defined in Table 1.

$$\max_{c(\cdot)} \int_0^\infty e^{-\rho t} U(c(t), L(t))\, dt \tag{53}$$
$$U(c(t), L(t)) = L(t) \log(c(t)) \tag{54}$$
$$\dot{L}(t) = g_L(t) L(t) \tag{55}$$
$$\dot{g}_L(t) = -\delta_L g_L(t) \tag{56}$$
$$Q(t) = \Omega(t) A(t) K(t)^\gamma L(t)^{1-\gamma} \tag{57}$$
$$\dot{A}(t) = g_A(t) A(t) \tag{58}$$
$$\dot{g}_A(t) = -\delta_A g_A(t) \tag{59}$$
$$Q(t) = C(t) + I(t) \tag{60}$$
$$c(t) = \frac{C(t)}{L(t)} \tag{61}$$
$$\dot{K}(t) = I(t) - \delta K(t) \tag{62}$$
$$E(t) = (1 - \mu(t))\sigma(t) Q(t) \tag{63}$$
$$\dot{M}(t) = \beta E(t) - \delta_M (M(t) - 590) \tag{64}$$
$$F(t) = 4.1 \frac{\log[M(t)] - \log[590]}{\log[2]} + O(t) \tag{65}$$
$$\dot{T}(t) = \frac{1}{R_1}\{F(t) - \lambda T(t)\} - \frac{R_2}{\tau_{12}}\{T(t) - T^*(t)\} \tag{66}$$
$$\dot{T}^*(t) = \frac{1}{\tau_{12}}\{T(t) - T^*(t)\} \tag{67}$$
$$D(t) = Q(t)\theta_1(T(t) + \theta_2 T(t)^2) \tag{68}$$
$$TC(t) = Q(t) b_1 \mu(t)^{b_2} \tag{69}$$
$$\Omega(t) = \frac{1 - b_1 \mu^{b_2}}{1 + \theta_1 (T + \theta_2 T^2)}. \tag{70}$$

This system is not time homogenous because of population growth and technical progress. However, the model implies asymptotic values for these exogenous variables that are given below:

$$\begin{aligned} \bar{A} &= 0.063 \\ \bar{L} &= 12000 \\ \bar{O} &= 1.15. \end{aligned}$$

Table 1: List of variables in the DICE94 model.

List of endogenous state variables		
$K(t)$	=	capital stock
$M(t)$	=	mass of GHG in the atmosphere
$T(t)$	=	atmospheric temperature relative to base period
$T^*(t)$	=	deep-ocean temperature relative to base period
List of control variables		
$I(t)$	=	gross investment
$\mu(t)$	=	rate of GHG emissions reduction
List of exogenous dynamic variables		
$A(t)$	=	level of technological progress
$L(t)$	=	labor input (= population)
$O(t)$	=	forcing exogenous GHG
List of auxiliary variables		
$C(t)$	=	total consumption
$c(t)$	=	per capita consumption
$D(t)$	=	damage from GH warming
$E(t)$	=	emissions of GHGs
$F(t)$	=	radiative forcing from GHGs
$\Omega(t)$	=	output scaling factor due to emissions control and to damages from climate change
$Q(t)$	=	gross world product

We therefore use these asymptotic values in our computations of an asymptotic attractor or "turnpike" for the optimal trajectory. In Table 2 we show the turnpike values for that control system, when there is no altruism ($\alpha = 0$) and when $\rho = 0\%$ or $\rho = 6\%$, respectively.[4] When $\rho = 0\%$ the turnpike is the asymptotic steady state of an optimal trajectory for a generation having an infinite life duration. When $\rho = 6\%$ the turnpike is the asymptotic steady state of an optimal growth trajectory for a generation having a life duration of 100/6=16.66 years.[5] From these results we see the dramatic influence of discount rates on the asymptotic values for the optimal growth paths. Remember that, in the global climate change context, these asymptotic values become highly relevant. The zero discount rate leads to a much higher capital stock, much lower emissions, much lower GHG concentrations and higher per capita consumption than the 6% rate does. We see very clearly that discounting, even at a moderate

[4]More details on the computation of these turnpike values can be obtained from [23].

[5]The reader may wonder about an infinite time horizon associated with a life expectation of 16.66 years. Indeed one has to recall that an exponential random life with killing rate ρ is such that the expected remaining life, given that one has already lived θ, is still $1/\rho$.

Table 2: Turnpike values when $\rho = 0$ and $\rho = 6\%$ respectively.

State variables			State variables		
$\bar{K}$	=	943	$\bar{K}$	=	505
$\bar{M}$	=	911	$\bar{M}$	=	1170
$\bar{T}$	=	2.64	$\bar{T}$	=	3.69
Control variables			**Control variables**		
$\bar{I}$	=	94.3	$\bar{I}$	=	50.5
$\bar{\mu}$	=	0.68	$\bar{\mu}$	=	0.33
Exogenous variables			**Exogenous variables**		
$\bar{A}$	=	0.063	$\bar{A}$	=	0.063
$\bar{L}$	=	12,000	$\bar{L}$	=	12,000
$\bar{O}$	=	1.15	$\bar{O}$	=	1.15
Auxiliary variables			**Auxiliary variables**		
$\bar{C}$	=	291	$\bar{C}$	=	278
$\bar{E}$	=	41.83	$\bar{E}$	=	75.45
$\bar{F}$	=	3.72	$\bar{F}$	=	5.20
$\bar{Q}$	=	400	$\bar{Q}$	=	342
$\rho = 0\%$			$\rho = 6\%$		

rate, means that in the long term the environment will be quite degraded. Also, discounting discourages capital accumulation.[6] Imposing a 0 discount rate is not feasible because the current generation has a pure time preference linked to its finite life expectation. This is why many economists using these types of models consider discount rates $\varrho(t)$ that are dependent on time and that converge to 0 as $t \to \infty$. With such a choice the economy will have a turnpike (asymptotic steady state) corresponding to the $\rho = 0$ case[7] but will start with investment decisions related to a high value of ρ. This solution is not satisfactory either since it does not satisfy the time consistency rule, because the forthcoming generations are not committed to use the $\rho = 0$ discount rate.

In the next section we look at the effect of introducing altruism in the economic evaluation scheme.

5.2 Intergenerational Equilibria

This model has exactly the format of the control system (1)–(3). We implement the intergenerational equilibrium scheme developed above. Taking the asymptotic value for the exogenous variables, we obtain a time stationary system hav-

[6] Typically for GHG abatement, one could envision investment in high-level production capital like the fusion power processes. These types of investments are discouraged by high discount rates.

[7] We refer to [24] for a discussion of turnpikes in optimal economic growth models with uncertain or variable discount rates.

ing turnpikes. We shall compare the turnpikes when $\rho = 0.06$ and $\alpha = 1$ (total selfishness) or $\alpha = 0.5$, respectively. To compute these turnpikes we implement the approximation scheme with a grid of $10 \times 10 \times 10$ nodes. The mesh value is taken at $h = 50$; the temperature state equation is normalized (multiplied by 200) to harmonize the three dimensions of the grid.

We proceed as indicated in Section 4, using the approximating MDPs to compute the optimal value function and the strategy γ^* for a candidate $\tilde{\Phi}^*(\cdot)$ function and updating $\tilde{\Phi}^*(\cdot)$ when a new γ^* is obtained.[8]

The phase 1 criterion is the sum of values $\sum_{g \in G} v(x^g)$; the phase 2 criterion is $\sum_{g \in G} \tilde{\Phi}^*(x^g)$. The convergence is reached when two successive phase 1 or phase 2 criteria are equal. The following array shows the convergence obtained through the implementation documented in the appendix.

$$\begin{array}{lcl}
Phase1 & = & 2656.218 \\
Phase2 & = & 5312.437 \\
Phase1 & = & 5278.196 \\
Phase2 & = & 5308.574 \\
Phase1 & = & 5276.634 \\
Phase2 & = & 5308.487 \\
Phase1 & = & 5276.613 \\
Phase2 & = & 5308.484 \\
Phase1 & = & 5276.613 \\
Phase2 & = & 5308.486 \\
Phase1 & = & 5276.613 \\
Phase2 & = & 5308.484 \\
\ldots & &
\end{array}$$

When displaying the value function and the next generation reward function, at convergence one observes that these functions are almost affine. We obtain, with an excellent least squares fit $R^2 = 0.999$,

$$\rho\tilde{\Phi}^*(K, M, T) = 311.33 + 0.0265\,K - 0.0091\,M - 0.0019\,T. \tag{71}$$

We are now in a position where we can compare the asymptotic behavior of the growth trajectory with and without altruism. These calculations are made by solving the implicit programming problems defined in Section 4.4. The turnpike, when $\alpha = 1$ and $\rho = 0.06$ (i.e., with total selfishness), is given by

$$\begin{array}{lcl}
\bar{K}_1 & = & 505.37 \\
\bar{M}_1 & = & 1169.84 \\
\bar{T}_1 & = & 3.69
\end{array}$$

[8] In an appendix available on demand from the author, the linear programming models used to solve the approximating MDPs are given in the AMPL format [20].

with steady state controls

$$\begin{aligned}\bar{I}_1 &= 50.537\\ \bar{\mu}_1 &= 0.33\\ \bar{C}_1 &= 278.46.\end{aligned}$$

We notice that the sustainable consumption level is lower than in the case $\rho = 0$, while the capital stock is higher, the GHG concentration and temperature being lower. When the selfishness coefficient is set to 0.5, the turnpike becomes

$$\begin{aligned}\bar{K}_{0.5} &= 638.58\\ \bar{M}_{0.5} &= 1080.43\\ \bar{T}_{0.5} &= 3.35\end{aligned}$$

with steady state controls

$$\begin{aligned}\bar{I}_{0.5} &= 63.86\\ \bar{\mu}_{0.5} &= 0.47\\ \bar{C}_{0.5} &= 285.42.\end{aligned}$$

As expected, the attractor is more capital intensive, and sustains a higher consumption and a lower concentration and temperature. However, these turnpike values are dominated by those corresponding to $\rho = 0$ shown in the previous section, in conformity with what was expected.

As indicated in Equation (27), the approach results in the introduction of a modified utility function, in the problem solved by the current generation, which in the present case will take the form

$$L(t)\log(c(t)) + 311.33 + 0.0265\,K - 0.0091\,M - 0.0019\,T. \tag{72}$$

The intergenerational equilibrium has introduced a utility function that depends directly on the capital stock and the environmental variables, since they are the bequest for the forthcoming generation.

6 Conclusion

The intergenerational equilibrium defined in this model is fully characterized through dynamic programming. It is therefore both time consistent and Markov-perfect. No generation could object to the proposed equilibrium strategy. This property is certainly important when one designs policies that should encompass the life of more than one generation. As shown in the complete analysis of the time stationary case, one may expect the equilibrium trajectory to lie between the two extreme cases of the purely egocentric first generation and the infinitely lived single generation problem, respectively. This property has been illustrated by applying this solution concept to a reduced version of the DICE94 integrated assessment model.

The introduction of this multigeneration equilibrium structure in an economic-environment growth model has permitted the consideration of a form of altruism in the definition of the utility functions of the decision makers. This might bring an interesting contribution to the debate concerning the proper social discount rate to apply when dealing with long-term environmental problems. In a companion paper [25] we extend this stochastic game model to the case where each generation has an interest in the rewards gained by all the forthcoming generations. This variant of the intergenerational stochastic game structure also provides a coherent and time-consistent way to modify the discounting schemes for the consideration of the welfare of future generations.

REFERENCES

[1] Ainslie G. and Haslam N., Hyperbolic discounting, in Loewenstein and Elster (eds.), *Choice Over Time*, Russell Sage Foundation, New York, 1992.

[2] Alj A. and Haurie A., Dynamic equilibria in multigeneration stochastic games, *IEEE Trans. Autom. Control*, AC-28, 2, 1983.

[3] Arrow K. and Kurz M., *Public Investment, the Rate of Return, and Optimal Investment Policies*, John Hopkins Press, Baltimore, MD, 1970.

[4] Arrow K., Bolin B., Costanza R., Dasgupta P., Folke C., Holling C., Jansson B-O., Levin S., Maler K-G., Perrings C. and Pimental D., Economic growth carrying capacity and the environment, *Science*, 268, 520–521, 1995.

[5] Arrow K.J., Discounting, morality, and gaming, in P.R. Portney and J. Weyant (eds.), *Discounting and Intergenerational Effects*, Resources for the Future, Washington DC, 23–29, 1999.

[6] Arrow K.J., Cline W.R., Maeler K.G., Munasinghe M., Squitieri R. and Stiglitz J.E., Intertemporal equity, discounting, and economic efficiency, in J.P. Bruce, H. Lee and E.F. Haites (eds.), *Climate Change 1995: Economic and Social Dimensions*, Contribution of Working Group III to the Second Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, 125–144, 1996.

[7] Baumol W. and Oates W., *Economics, Environmental Policy and the Quality of Life*, Prentice-Hall, Englewood Cliffs, NJ, 1979.

[8] Baumol W. and Oates W., *The Theory of Environmental Policy*, Cambridge University Press, Cambridge, 1988.

[9] Beltratti A., Chichilnisky G. and Heal G., The green golden rule, *Economic Letters*, 49, 175–179, 1995.

[10] Carlson D., Haurie A. and Leizarowitz A., *Infinite Horizon Optimal Control: Deterministic and Stochastic Systems*, Springer-Verlag, Berlin, 1994.

[11] Chichilnisky G., An axiomatic approach to sustainable development, *Soc. Choice Welf.*, 13, 231–257, 1996.

[12] Chichilnisky G., What is sustainable development, *Land Economy*, 73, 467–491, 1997.

[13] Cline W.R., Discounting for the very long term, in P.R. Portney and J. Weyant (eds.), *Discounting and Intergenerational Equity*, Resources for the Future, Washington DC, 131–140, 1999.

[14] Dasgupta P., Mäler K.-G. and Barrett S., Intergenerational equity, social discount rate and global warming, in P.R. Portney and J. Weyant (eds.), *Discounting and Intergenerational Effects*, Resources for the Future, Washington DC, 51–77, 1999.

[15] Feinberg E. and Schwartz A., Markov decision processes with weighted discounted criteria, *Mathematics of Operations Research*, 19, 152–168, 1995.

[16] Feinstein C.D and Luenberger D.G., Analysis of the asymptotic behavior of optimal control trajectories, *SIAM Journal on Control and Optimization*, 19, 561–585, 1981.

[17] Filar J. and Vrieze K., Weighted reward criteria in competitive markov decision processes, *Zeitschrift für O.R.*, 36, 343–358, 1992.

[18] Filar J. and Vrieze K., *Competitive Markov Decision Processes*, Springer-Verlag, Berlin, 1996.

[19] Fleming W.H. and Rishel R.W., *Deterministic and Stochastic Optimal Control*, Springer-Verlag, Berlin, 1975.

[20] Fourer R., Gay D.M. and Kernighan B.W., *AMPL*, The Scientific Press, Redwood City, CA, 1993.

[21] Harris C. and Laibson D., Instantaneous gratification, mimeo, Harvard University, 2001.

[22] Harvey C.M., The reasonableness of non-constant discounting, *Journal of Public Economics*, 53, 31–51, 1994.

[23] Haurie A., Integrated assessment modeling for global climate change: an infinite horizon viewpoint, *Environmental Modeling and Assessment*, 2003.

[24] Haurie A., Turnpikes in multi-discount rate environments and GCC policy evaluation, in G. Zaccour (ed.), *Optimal Control and Differential Games; Essays in Honor of Steffen Jørgensen*, Kluwer, Dordrecht 2002.

[25] Haurie A., A multigenerational game model to analyze sustainable development, 2005, to appear in *Annals of OR*, 137, 369–386, 2005.

[26] Haurie A. and Moresino F., Computing equilibria in stochastic games of intergenerational equity, 2005, to appear in *International Game Theory Review*.

[27] Kushner H.J. and Dupuis P.G., *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer-Verlag, Berlin–New York, 1992.

[28] Laibson D., Golden eggs and hyperbolic discounting, *Quart. J. Econ.*, 62, 443–478, 1997.

[29] Nordhaus W.D., An optimal transition path for controlling greenhouse gases, *Science*, 258, 1315–1319, 1992.

[30] Nordhaus W.D., *Managing the Global Commons: The Economics of Climate Change*, MIT Press, Cambridge, MA, 1994.

[31] Nordhaus W.D., A market based discount rate, in P.R. Portney and J. Weyant (eds.), *Discounting and Intergenerational Effects*, Resources for the Future, Washington DC, 145–162, 1999.

[32] Nordhaus W.D. and Boyer J., *Warming the World: Economic Models of Global Warming*, MIT Press, Cambridge, MA, 2000.

[33] Phelps E.S. and Polak R., On second-best national saving and game-equilibrium growth, *Review of Economic Studies*, 35, 2, 185–199, 1968.

[34] Portney P.R. and Weyant J. (eds.), *Discounting and Intergenerational Effects*, Resources for the Future, Washington DC, 1999.

[35] Puterman M. (ed.), *Markov Decision Processes*, Wiley Interscience, New York, 1994.

[36] Ramsey F., A mathematic theory of saving, *Economic Journal*, 38, 543–549, 1928.

[37] Sorger G., Consistent planning under quasi-geometric discounting, mimeo, 2002.

[38] Strotz R.H., Myopia and inconsistency in dynamic utility maximization, *Review of Economic Studies*, 23, 165–180, 1956.

[39] Tóth F., Intergenerational equity and discounting, *Integrated Assessment*, 1, 127–136, 2000.

[40] Weitzman M.L., Why the far-distant future should be discounted at its lowest possible rate, *Journal of Environmental Economics and Management*, 201–208, 1998.

[41] Weitzman M.L., Just keep discounting, but ... , in P.R. Portney and J. Weyant (eds.), *Discounting and Intergenerational Effects*, Resources for the Future, Washington DC, 23–29, 1999.

[42] Whitt W., Representation and approximation of noncooperative sequential games, *SIAM J. Contr. Optimiz.*, 18, 1, 33–48, 1980.

An Impulsive Differential Game Arising in Finance with Interesting Singularities*

Pierre Bernhard
I3S, University of Nice-Sophia Antipolis and CNRS
France
Pierre.Bernhard@polytech.unice.fr

Naïma El Farouq
I3S, University of Nice-Sophia Antipolis and CNRS
France
elfarouq@i3s.unice.fr

Stéphane Thiery
I3S, University of Nice-Sophia Antipolis and CNRS
France
thiery@i3s.unice.fr

Abstract

We investigate a differential game motivated by a problem in mathematical finance. This game displays two interesting features. On the one hand, one of the players, **P**ursuer say, may, and will, use infinitely large controls, i.e., impulses, producing "jumps" in the state variables. Standard optimal trajectories are made of such a jump followed by a "coasting period" where **P** exerts no control. This leads to barriers of a somewhat new type. But because the cost of jumps is only proportional to their amplitude, some singular optimal trajectories arise where **P** uses an intermediary control, nonzero but finite. (In classical impulse control, there is a minimum positive cost to any use of the control, forbidding such a mixed situation.)

On the other hand, the complete solution of the game exhibits a type of singularity, the existence of which had long been conjectured (noticeably by Arik Melikyan in discussions with the first author) but, as far as we know, never shown in actual examples: a two-dimensional focal manifold traversed by noncollinear optimal fields depending on the control used by **E**vader. It is on this manifold that intermediary controls for **P** arise.

Finally, we show that the Isaacs equation of a discrete-time version of the problem provides a discretization scheme that converges to the value function of the differential game. This is done through the investigation of a (degenerate) quasi-variational inequality and its viscosity solution, with

*We acknowledge an important numerical work performed by intern students at the University of Nice-Sophia Antipolis: Nicéphore Allaglo, Carole Bouvelot, Charlotte Pouderoux, and Laetitia Richter.

the help of an equivalent, but nonimpulsive, differential game—a method of interest per se that we credit to Joshua—to which we apply essentially the classical method of Capuzzo Dolcetta extended to differential games by Pourtallier and Tidball, with some technical adaptations.

1 The Differential Game Considered

We consider a differential game arising in finance, specifically in the theory of option pricing with an "interval model." (We refer to [4,16] for the context in finance.) This is a game in two dimensions plus time with an integral payoff, or three dimensions plus time with a terminal payoff, and two scalar controls (pursuer **P** and evader **E**), with the peculiarity that the pursuer may, and will, use arbitrarily large control values, up to the point of producing "impulses." Thus, this player may cause discontinuities in some state variables, incurring a related cost.

1.1 Dynamics

The (3-D) dynamics are as follows. We call (x, y, z) the state variables, and u and v the controls of pursuer and evader respectively. The continuous (nonimpulsive) part of the dynamics is given in terms of $\varepsilon = \mathrm{sign}(u)$ and two numbers C_{+1} and C_{-1}, also written C^+ and C^- respectively, with $C^+ > 0$, $C^- < 0$, as follows:

$$\dot{x} = vx\,, \tag{1}$$
$$\dot{y} = vy + u\,, \tag{2}$$
$$\dot{z} = vy - C_\varepsilon u\,, \tag{3}$$

with the control constraints on v specified by two positive numbers α and β as

$$-\alpha \le v \le \beta\,. \tag{4}$$

Since u is not bounded, we allow the pursuer to cause discontinuities in the state variables at isolated time instants t_k according to the rule

$$y(t_k^+) = y(t_k^-) + u_k\,, \tag{5}$$
$$z(t_k^+) = z(t_k^-) - C_{\varepsilon_k} u_k\,. \tag{6}$$

Of course, we have set $y(t_k^-) = \lim_{t\uparrow t_k} y(t)$ and $y(t_k^+) = \lim_{t\downarrow t_k} y(t)$ and likewise for z. The jump amplitude in y is $u_k \in \mathbb{R}$, and $\varepsilon_k = \mathrm{sign}(u_k)$.

To avoid an unessential discussion later on, we shall further assume that

$$\alpha \le \beta\,, \qquad \text{and} \qquad 0 < (1 + C^+)(1 + C^-) \le 1\,. \tag{7}$$

1.2 Payoff

The game is played over a fixed time interval $[0,T]$, and is a capture-evasion *game of kind*, with capture defined in terms of a given positive number Z as

$$z(T) \geq \max\{0\,,\, x(T)-Z\} =: M(x(T))\,. \tag{8}$$

Again this rather strange setup is motivated by its finance application in [4].

We may notice that, since z does not appear in the right-hand side of its dynamics, it integrates so that (8) is equivalent to

$$z(0) \geq \int_0^T (-vy + C_\varepsilon u)\,\mathrm{d}t + \sum_k C_{\varepsilon_k} u_k + M(x(T))\,.$$

As a consequence, we may consider the *game of degree* in dimension 2 plus time with state variables (x,y), the same dynamics (1) (2) and (5)(6), and payoff $\min_u \max_v G$ with

$$G = \int_0^T (-vy + C_\varepsilon u)\,\mathrm{d}t + \sum_k C_{\varepsilon_k} u_k + M(x(T))\,. \tag{9}$$

Let $W(t,x,y)$ be its value function, an initial state is capturable iff $z \geq W(0,x,y)$, so that the graph of the value function W is the barrier of the *game of kind.*

1.3 Strategies

In this game, the pursuer chooses the function $u(t)$, the jump instants t_k, and the jump amplitudes u_k. It does so knowing past values of the state. It is a classical fact that it will only use an (instantaneous) *state feedback* which we write symbolically $u = \varphi(t, X(t))$, where X stands for the whole state. Admissible strategies are those such that the dynamical equations have for any initial state a unique solution with $y(\cdot)$ uniformly bounded over admissible $v(\cdot)$'s.

We are looking for capturable states of the *game of kind.* It is known that this is equivalent to looking for the upper value of the game of degree, and that then, whether the evader plays open loop or closed loop is irrelevant. Thus we may always assume that v is chosen open loop, as a measurable time function from $[0,T]$ into $[-\alpha,\beta]$. (This remark will play an important role in the investigation of the convergence.)

2 A Geometric Analysis: The Isaacs–Breakwell Theory

2.1 Jumps as Ordinary Trajectories

In [4], we introduced a quasi-variational inequality (QVI) naturally related to the game of degree with impulse controls. However, due to its very degenerate nature, it is not accounted for by the literature on viscosity solutions of first order QVI such as [3,2]. We prefer to use the 3-D plus time representation (1), (2), (3), (5), (6), and the formulation as a *game of kind*, and apply to it the geometrical tools of the semipermeability.

In that representation, jumps are just trajectories orthogonal to the t axis. As a matter of fact, Equations (5) and (6) show that these trajectories are also orthogonal to the x axis and have a slope either $-C^+$ or $-C^-$ in the (y, z) plane. We stress the following fact.

Proposition 2.1. *Given a smooth two-dimensional manifold $\mathcal{M}$ transverse to the jump trajectories, the hypersurface made of jump trajectories of the same slope through each point of $\mathcal{M}$ is a "safe hypersurface" for* **P***, (i.e.,* **E** *cannot force the state to cross it against* **P***'s will).*

Proof. Indeed by choosing a jump, **P** causes the state to traverse these trajectories in no time, so that **E**'s control v has no time to act. (**P** has chosen to be in the dynamics (5), (6) where v does not enter.) □

We shall in effect construct manifolds $y = \check{y}(t, x)$, $z = \check{z}(t, x)$ for some functions $\check{y}$ and $\check{z}$, construct barriers made of jump trajectories reaching that manifold, and show that upon reaching it, **P** still has a means of preventing a crossing of the composite surface.

2.2 The Natural Barrier

We proceed with the classical construction of the natural barrier through the boundary of the capture set, which here is $t = T$, $z = M(x)$, y arbitrary. This has been published in [4]. We summarize it here.

The natural barrier is made up of two sheets, one towards $x \le Z$ and one towards $x \ge Z$. They are given below, together with a corresponding *inward* semipermeable normal as the vector $(n, p, q, 1)$ (corresponding to the state variables (t, x, y, z)), leading to Isaacs' "main equation"

$$0 = \max_{u} \inf_{v \in [-\alpha, \beta]} \left[n + v\Big(px + (q+1)y\Big) + u(q - C_\varepsilon) \right] ,$$

and the adjoint equations

$$\dot{p} = -vp\,, \tag{10}$$
$$\dot{q} = -v(q+1). \tag{11}$$

The analysis depends on the fact that the maximum in u of $(q - C_\varepsilon)u$ is reached at $u = 0$ provided that $C^- \le q \le C^+$. (Remember that $\varepsilon = \mathrm{sign}(u)$.) When q leaves that range, there is no maximum anymore. (Or u should be infinite: we shall have a jump.)

Sheet $\boldsymbol{\alpha}$ towards $\boldsymbol{x \le Z}$. We set the parameters $x(T) = s \le Z$, $y(T) = r$. It yields $v^* = -\alpha$ and

$$\begin{array}{rclcrcl} & \text{sheet } (\alpha) & & \qquad & & \text{semipermeable normal } \nu_\alpha & \\ t &=& t & & n(t) &=& \alpha r\,, \\ x(t) &=& s\mathrm{e}^{\alpha(T-t)}\,, & & p(t) &=& 0\,, \\ y(t) &=& r\mathrm{e}^{\alpha(T-t)}\,, & & q(t) &=& \mathrm{e}^{-\alpha(T-t)} - 1\,, \\ z(t) &=& r(\mathrm{e}^{\alpha(T-t)} - 1)\,, & & 1 &=& 1\,. \end{array}$$

This is a valid solution as long as $q \ge C^-$, i.e., for $t \ge t_\alpha$ with

$$\mathrm{e}^{-\alpha(T-t_\alpha)} = 1 + C^-\,, \quad \text{i.e.,} \quad T - t_\alpha = \frac{1}{\alpha}\ln\left(\frac{1}{1+C^-}\right). \tag{12}$$

Sheet $\boldsymbol{\beta}$ towards $\boldsymbol{x \ge Z}$. On this sheet, $x(T) = s \ge Z$, $y(T) = r$. We find that $v^* = \beta$, and

$$\begin{array}{rclcrcl} & \text{sheet } (\beta) & & \qquad & & \text{semipermeable normal } \nu_\beta & \\ t &=& t & & n(t) &=& \beta(s-r)\,, \\ x(t) &=& s\mathrm{e}^{-\beta(T-t)}\,, & & p(t) &=& -\mathrm{e}^{\beta(T-t)}\,, \\ y(t) &=& r\mathrm{e}^{-\beta(T-t)}\,, & & q(t) &=& \mathrm{e}^{\beta(T-t)} - 1\,, \\ z(t) &=& r(\mathrm{e}^{-\beta(T-t)} - 1) + s - Z\,, & & 1 &=& 1\,. \end{array}$$

This is a valid solution as long as $q \le C^+$, i.e., for $t \ge t_\beta$ with

$$\mathrm{e}^{\beta(T-t_\beta)} = 1 + C^+\,, \quad \text{i.e.,} \quad T - t_\beta = \frac{1}{\beta}\ln(1+C^+)\,. \tag{13}$$

From the hypothesis (7), we have $t_\alpha < t_\beta$.

Moreover, from final states on the boundary $z = x - Z \ge 0$ of the admissible set, a 2-D singular sheet can be constructed with $r = s$, v arbitrary, leading to

$$x = y = z - Z = s\exp\left(-\int_t^T v(\tau)\,\mathrm{d}\tau\right),\ -p = q + 1 = \exp\left(\int_t^T v(\tau)\,\mathrm{d}\tau\right).$$

Intersection and Composite Barrier. The two main sheets (α) and (β) intersect along a two-dimensional edge $\mathcal{D}$ that spans the domain $t \ge t_\beta$, $Z\mathrm{e}^{-\beta(T-t)} \le x \le Z\mathrm{e}^{\alpha(T-t)}$, and that can be parametrized by (t,x) as $y = \check{y}(t,x)$, $z = \check{z}(t,x)$ given by

$$\check{y}(t,x) = \frac{(x\mathrm{e}^{\beta(T-t)} - Z)}{\mathrm{e}^{\beta(T-t)} - \mathrm{e}^{-\alpha(T-t)}}, \quad \check{z}(t,x) = (1 - \mathrm{e}^{-\alpha(T-t)})\check{y}(t,x). \tag{14}$$

Notice that for $x = Z\exp(-\beta(T-t))$, we have $\check{y} = \check{z} = 0$, which corresponds to the sheet (α) with $r = 0$. For smaller x's, only the sheet (α) plays a role.

We find it convenient to extend the definition of $\check{y}$ and $\check{z}$ by 0 for both. For $x = Z\exp(\alpha(T-t))$, we have $\check{y} = x$, $\check{z} = x - Z$, which corresponds to the sheet (β) with $r = s = Z\exp((\alpha+\beta)(T-t_\beta))$. For larger x's, only the sheet (β) plays a role. Again, we extend the definitions of $\check{y} = x$ and $\check{z} = x - Z$ to larger x's.

We easily check that $\mathcal{D}$ is an **E**-dispersal line. States "above" (with larger z's) are indeed capturable, and the edge does not "leak" since **P**'s control on both sheets is the same: $u = 0$. Therefore, this same control prevents crossing of both barrier sheets.

The singular sheet $x = y = z + Z$ is imbedded in the sheet (β). But it can be used against $v = -\alpha$ until time t_α. In the region $x \le Z\exp(\alpha(T-t))$ it plays no role. However, it will be seen to play a prominent role in the region $x \ge Z\exp(\alpha(T-t))$ for $t \le t_\beta$, when the sheet (β) does not exist. There it behaves as a manifold drawn on an extension of the sheet (α) for $s \ge Z$.

2.3 Junction of a Jump Manifold and the Natural Barrier

For $t \le t_\beta$, the sheet (β) of the natural barrier does not exist, since it would entail a $q \ge C^+$, leading to $u = +\infty$. We therefore expect a positive jump manifold, i.e., trajectories in the (y,z) plane with slope $-C^+$. They must join on a two-dimensional manifold $\mathcal{E}$ drawn on the sheet (α), and such that, whatever v, **P** can maintain the state on or above both that sheet and the jump manifold. The manifold $\mathcal{E}$ will indeed be an "equivocal" one (in Isaacs' parlance), constructed according to the technique of a "safe contact" on a barrier, as originally discovered by Breakwell and Merz [9,12].

We first determine a control $u(v)$ that maintains the state on the barrier sheet (α). Let ν_α be the normal to that sheet; we have

$$\langle \nu_\alpha, \dot{X}\rangle = \mathrm{e}^{-\alpha(T-t)}(v+\alpha)y - u(1 + C_\varepsilon - \mathrm{e}^{-\alpha(T-t)})$$

so that we keep the state on the sheet (α) by choosing

$$u = \frac{\mathrm{e}^{-\alpha(T-t)}(v+\alpha)y}{1 + C^+ - \mathrm{e}^{-\alpha(T-t)}}\,.$$

With that control, keeping in mind that the normal to the jump manifold, say ν_j, has to be of the form $\nu_j = (n_j, p_j, C^+, 1)$, we get on $\mathcal{E}$:

$$\langle \nu_j, \dot{X}\rangle = n_j + v(p_j x + (1 + C^+)y)\,.$$

Furthermore, we want $\mathcal{E}$ to join on the boundary of $\mathcal{D}$ at $t = t_\beta$. Therefore, ν_j there should be normal to that boundary. This gives $p_j = -(1 + C^+)$, i.e., the same as in ν_β as it should be, and hence at $t = t_\beta$,

$$\langle \nu_j, \dot{X}\rangle = n - v(1 + C^+)(x - y)\,.$$

The domain considered thus far, the boundary of $\mathcal{D}$, is such that $x \ge y$. As a consequence, the minimizing v is $v = \beta$. Furthermore, if we construct the manifold $\mathcal{E}$ using $v = \beta$ in the above construction, we can check[1] that n_j remains positive, hence $(p_j x + (1+C^+)y)$ is negative. Hence $v = \beta$ is indeed minimizing, or, equivalently, we check that the strategy $u(v)$ above guarantees that the state lies on the sheet (α) *and* on the desired side of the jump manifold.

We therefore obtain the following.

Theorem 2.1. *The equations of the equivocal manifold $\mathcal{E}$ are*

$$\begin{array}{ll}
\dot{x} = \beta x\,, & x(t_\beta) = \frac{s}{1+C^+}\,,\\
\dot{y} = \frac{\beta(1+C^+)+\alpha \mathrm{e}^{-\alpha(T-t)}}{1+C^+-\mathrm{e}^{-\alpha(T-t)}} y\,, & y(t_\beta) = \frac{s-Z}{1+C^+-\mathrm{e}^{-\alpha(T-t_\beta)}}\,,\\
\dot{z} = \frac{\beta(1+C^+)(1-\mathrm{e}^{-\alpha(T-t)})-C^+\alpha \mathrm{e}^{-\alpha(T-t)}}{1+C^+-\mathrm{e}^{-\alpha(T-t)}} y\,, & z(t_\beta) = (1-\mathrm{e}^{-\alpha(T-t_\beta)})y(t_\beta).
\end{array} \tag{15}$$

We can integrate these backwards as long as the sheet (α) exists, i.e., down to $t = t_\alpha$. However, due to our restricted set of initial conditions, this will only take care of the domain $s \in [Z, Z\exp((\alpha+\beta)(T-t_\beta))]$, i.e., $Z\exp(-\beta(T-t)) \leq x \leq Z\exp(\alpha(T-t_\beta) - \beta(t_\beta - t))$. We need to find the extension of the manifold $\mathcal{E}$ to all values of (t, x) for $t \in [t_\alpha, t_\beta]$.

In the region $x \leq Z\exp(-\beta(T-t))$, the above equations are to be taken with terminal conditions $y = z = 0$, and thus yield $y = z = 0$ down to $t = t_\alpha$.

In the region $x \geq Z\exp(\alpha(T-t))$, we do not have the sheet (α) to perform the above construction, but we do it with the singular sheet $y = x$, $z = x - Z$. A completely similar analysis yields a u proportional to $y - x$, i.e., zero on the singular sheet, which turns out itself to be the manifold $\mathcal{E}$.

This joins smoothly with $\mathcal{D}$ in the region $t \geq t_\beta$. We shall use it as terminal conditions for the equations of $\mathcal{E}$ along the boundary $x = Z\exp(\alpha(T-t))$, $t_\alpha \leq t \leq t_\beta$. That way, we have defined the manifold $\mathcal{E}$ in all the required domain. Again, for $t \in [t_\alpha, t_\beta]$, let $y = \check{y}(t, x)$, $z = \check{z}(t, x)$ describe this manifold. The functions $\check{y}$ and $\check{z}$ thus defined extend continuously those for $t \geq t_\beta$ defined on $\mathcal{D}$.

It turns out that the equations for y integrate analytically. See Appendix A.

2.4 The Focal Manifold

2.4.1 Principle

For $t \leq t_\alpha$, neither of the two sheets of the natural barrier exist. We must therefore replace them both by jump manifolds that will join on a new manifold, which is thus a focal surface, (but with adjoining trajectories that are jump trajectories). Let us call it $\mathcal{F}$.

To explain how to construct $\mathcal{F}$, we need to introduce some notation. We shall have two jump manifolds, one with negative jump and one with positive jump.

[1] We did it numerically. There should exist an analytical proof.

Let ν^- and ν^+ be the corresponding normals. They are of the form

$$\nu^- = \begin{pmatrix} n^- \\ p^- \\ C^- \\ 1 \end{pmatrix}, \qquad \nu^+ = \begin{pmatrix} n^+ \\ p^+ \\ C^+ \\ 1 \end{pmatrix}.$$

Let also $\dot X = f(X,u,v)$ denote the dynamics. Upon reaching $\mathcal{F}$, player **P** will have to choose a control $u(v)$ that will maintain the state on $\mathcal{F}$ or above the composite barrier thus constructed. Assume that for the extreme values of v, i.e., $-\alpha$ and β, the state can just be maintained on $\mathcal{F}$. Let u_α and u_β be the corresponding controls. Now, we must have the following equalities:

$$\begin{array}{ll} 0 = \langle \nu^-, f(X,u_\alpha,-\alpha)\rangle\,, & 0 = \langle \nu^+, f(X,u_\alpha,-\alpha)\rangle\,, \\ 0 = \langle \nu^-, f(X,u_\beta,\beta)\rangle\,, & 0 = \langle \nu^+, f(X,u_\beta,\beta)\rangle\,. \end{array}$$

We have six unknowns, $n^-, p^-, n^+, p^+, u_\alpha, u_\beta$. We need two more equations to determine them.

2.4.2 Trajectories $v = \beta$

We choose to describe $\mathcal{F}$ as the set of trajectories obtained for $v = \beta$. Later we shall discuss this arbitrary choice. In this description, let $X^\beta(s,t)$ be our state,[2] depending on the parameter s characterizing the trajectory (say reaching the boundary of $\mathcal{E}$ at t_α at the point $u = s\exp(\beta(T - t_\alpha))$ and on t. Thus

$$\frac{\partial X^\beta}{\partial t} = X^\beta_t = \dot X^\beta = f(X, u_\beta, \beta)\,.$$

We need further express that all trajectories lie in the same manifold $\mathcal{F}$. Hence, let $X^\beta_s := \partial X^\beta / \partial s$; we must further have

$$0 = \langle \nu^-, X^\beta_s\rangle\,, \qquad 0 = \langle \nu^+, X^\beta_s\rangle\,.$$

We now have six equations in six unknowns at each X. We want to use them to recover u_β and put it in the equations of the dynamics. Surprisingly, this is rather easy to do. The first four equations yield

$$u_\alpha = \frac{\alpha+\beta}{C^+ - C^-}[p^+ x + (1 + C^+)y]\,, \quad u_\beta = \frac{\alpha+\beta}{C^+ - C^-}[p^- x + (1 + C^-)y]\,.$$

The equation $0 = \langle \nu^-, X^\beta_s\rangle$ yields

$$p^- x^\beta_s = -(C^- y^\beta_s + z^\beta_s)\,.$$

[2] Obviously, β here is a superindex, not a power!

Now, we still have $\dot{x^\beta} = \beta x$, i.e., $x^\beta = s\exp(-\beta(T-t))$. Thus $x^\beta(s,t) = s x^\beta_s$. Hence, the above equation reads

$$p^- x = -s(C^- y^\beta_s + z^\beta_s)\,.$$

Put this back in u_β; this finally yields a pair of coupled partial differential equations (PDEs). We use the notation

$$\gamma = \frac{\alpha+\beta}{C^+ - C^-}\,, \tag{16}$$

to get the following fact (we have dropped the superindices β).

Theorem 2.2. *The focal manifold satisfies the following system of partial differential equations:*

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = s\gamma \begin{pmatrix} -C^- & -1 \\ C^+C^- & C^+ \end{pmatrix} \begin{pmatrix} y_s \\ z_s \end{pmatrix} + \begin{pmatrix} \beta + \gamma(1+C^-) \\ \beta - C^+\gamma(1+C^-) \end{pmatrix} y\,. \tag{17}$$

Domain and Boundary Conditions. We need to know $\mathcal{F}$ for all $t \le t_\alpha$ and all $x \ge 0$. However, for $x \le Z\exp(-\beta(T-t))$, we have previously argued that we expect the optimal (y,z) to be $(0,0)$. Also, for $x \ge Z\exp(\alpha(T-t))$, we expect the optimal (y,z) to be $(x, x-Z)$. Notice first that each of these two pairs, with $x = s\exp(-\beta(T-t))$, satisfies the PDE (17). It remains to fill the domain $\Omega := \{t \le t_\alpha,\ Z\exp(-\beta(T-t)) \le x \le Z\exp(\alpha(T-t))\}$, using the above known values at the boundaries in x, and the previously computed values on $\mathcal{E}$ at $t = t_\alpha$ for $Z\exp(-\beta(T-t_\alpha)) \le x \le Z\exp(\alpha(T-t))$.

This may entail discontinuities of the gradients of y and z along the "lateral" boundaries of Ω. Appendix B provides a proof that these two lines are precisely the possible support of such discontinuities. It also provides a further mathematical and numerical investigation of this PDE.

We therefore have a manifold $\mathcal{F}$ defined for all $t \le t_\beta$, all positive x's. We still call $\check{y}(t,x)$, $\check{z}(t,x)$ this manifold, and observe that the functions $\check{y}$ and $\check{z}$ are continuous.

2.4.3 Trajectories $v = -\alpha$

As stressed above, the choice to analyze $\mathcal{F}$ through the trajectories generated by $v = \beta$ was arbitrary. The same analysis could have been made using the trajectories $v = -\alpha$. Let them be parametrized by $u = r\exp(\alpha(T-t))$, and $X^\alpha(r,t)$ be the resulting manifold. One obtains the PDE

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = r\gamma \begin{pmatrix} -C^+ & -1 \\ C^-C^+ & C^- \end{pmatrix} \begin{pmatrix} y_r \\ z_r \end{pmatrix} + \begin{pmatrix} -\alpha + \gamma(1+C^+) \\ -\alpha - C^-\gamma(1+C^+) \end{pmatrix} y\,. \tag{18}$$

(Notice that one gets these equations upon interchanging in (17) $-\alpha$ with β on the one hand, and C^- with C^+ on the other hand.)

Proposition 2.2. *The PDEs (17) and (18) (with the same boundary conditions) describe the same manifold in the* (t, x, y, z) *space.*

Proof. Let us pick

$$s = r \exp((\alpha + \beta)(T - t)) \tag{19}$$

so that the coordinates x coincide. Let $Y = (y\ z)^t$. We want to show (with transparent notation) that $Y^\alpha(r, t) = Y^\beta(r \exp((\alpha + \beta)(T - t)), t)$. Therefore,

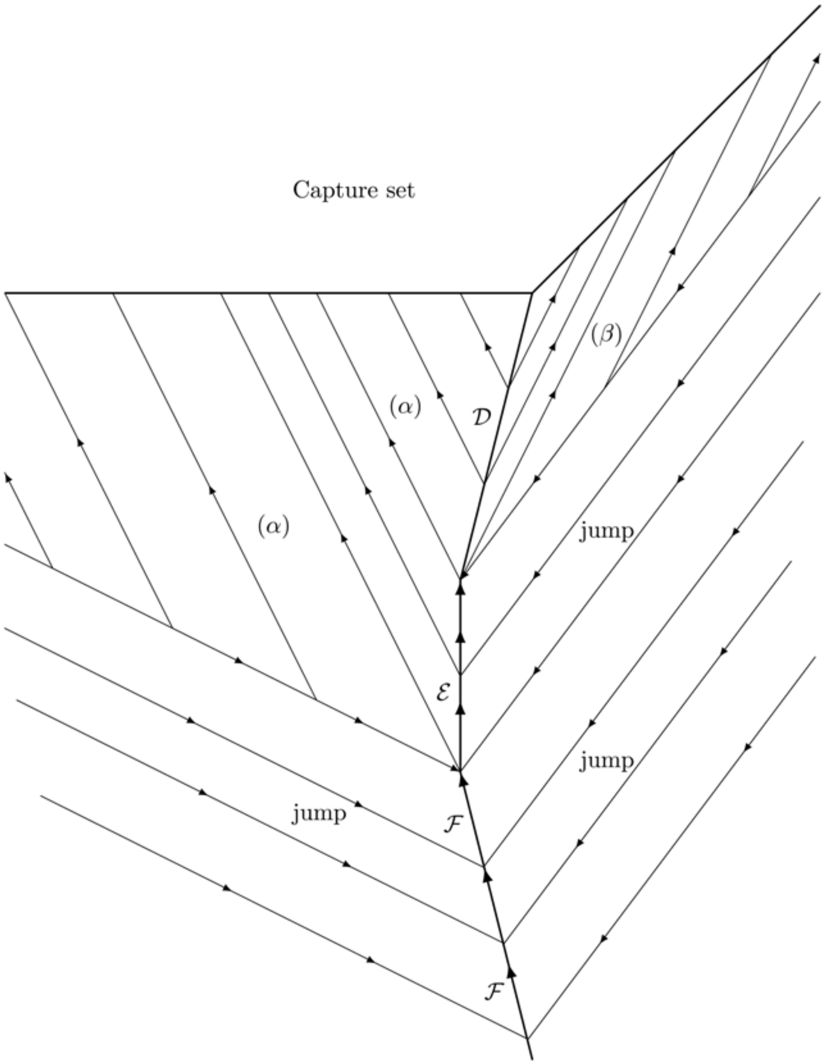


Figure 1: A 2-D sketch of the 4-D geometry of the barrier.

we should have

$$
\begin{aligned}
Y_r^\alpha &= \mathrm{e}^{(\alpha+\beta)(T-t)} Y_s^\beta \,, && (20)\\
Y_t^\alpha &= -(\alpha+\beta) r \mathrm{e}^{(\alpha+\beta)(T-t)} Y_s^\beta + Y_t^\beta \,. && (21)
\end{aligned}
$$

Write (17) and (18) respectively as

$$
Y_t^\beta = s A^\beta Y_s^\beta + B^\beta Y^\beta \,, \quad Y_t^\alpha = r A^\alpha Y_r^\alpha + B^\alpha Y^\alpha \,.
$$

We also need the notation $D := \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$. We therefore have

$$
\begin{aligned}
A^\beta &= \gamma \begin{pmatrix} -1 \\ C^+ \end{pmatrix} \begin{pmatrix} C^- & 1 \end{pmatrix}, \;\; B^\beta = (\beta I - A^\beta) D \,,\\
A^\alpha &= \gamma \begin{pmatrix} -1 \\ C^- \end{pmatrix} \begin{pmatrix} C^+ & 1 \end{pmatrix}, \;\; B^\alpha = (-\alpha I - A^\alpha) D \,.
\end{aligned}
$$

Substituting both (17) and (20) into (21) and also using (19), we get

$$
Y_t^\alpha = r[-(\alpha+\beta) I + A^\beta] Y_r^\alpha + B^\beta Y^\alpha \,. \tag{22}
$$

The proposition then results from the easy fact that (remembering (16))

$$
A^\alpha = A^\beta - (\alpha+\beta) I \,, \quad \text{and therefore} \quad B^\alpha = B^\beta \,,
$$

so that (22) coincides with (18). □

2.5 Synthesis

The boundary of the set of capturable states is given by $z = W(t,x,y)$ defined, in the domain $t \in [0,T]$, $x \geq 0$, $y \in [0,x]$, by

$$
W(t,x,y) = \check{z}(t,x) + D_\eta(y - \check{y}(t,x)) \,,
$$

where

- The functions $\check{y}$ and $\check{z}$ are given by the requirement that they be continuous (which specifies the boundary values of the differential equations) and
  - (i) $\check{y} = \check{z} = 0$ if $x \leq Z\exp(-\beta(T-t))$,
  - (ii) $\check{y} = x$, $\check{z} = x - Z$ if $x \geq Z\exp(\alpha(T-t))$,
  - (iii) if $x \in [Z\exp(-\beta(T-t)), Z\exp(\alpha(T-t))]$,
    - · if $t \geq t_\beta$, equations (14)
    - · if $t \in [t_\alpha, t_\beta]$, differential equations (15) with terminal conditions as in (15) $t = t_\beta$, and and by continuity with region (ii) above on the boundary in x,
    - · if $t \leq t_\alpha$, equations (17) with terminal conditions by continuity with the above at $t = t_\alpha$, and by continuity with (i) and (ii) on the boundaries in x,

- $\eta = \text{sign}(y - \check{y})$, and $D_{+1} = D^+$ and $D_{-1} = D^-$ are given by

$$D^+ = \begin{cases} -C^- & \text{if } t \le t_\alpha\,, \\ 1 - \mathrm{e}^{-\alpha(T-t)} & \text{if } t \ge t_\alpha\,, \end{cases}$$

$$D^- = \begin{cases} -C^+ & \text{if } t \le t_\beta\,, \\ 1 - \mathrm{e}^{\beta(T-t)} & \text{if } t \ge t_\beta\,. \end{cases}$$

This function W is therefore also the upper value function of the *game of degree* in (x, y) with payoff given by (9). Figure 1 shows a sketch of this compound manifold.

3 Discretization

3.1 The Multistage Game

In [4], we investigated a discrete-time version of the same problem. In discrete time, there are no such things as impulse controls (or there are *only* such things!), so that this is now a classical multistage game. Let $h = T/N$, with N an integer, be our time step. We shall often use a dyadic division, i.e., $N = 2^d$, with d an integer. Write $x(kh^-) = x_k$, and likewise for y, z and $W(kh, x, y) = W_k(x, y)$.

The following system is the natural discretization of our game (and is of interest per se in the finance application):

$$x_{k+1} = (1 + v_k)x_k\,, \tag{23}$$

$$y_{k+1} = (1 + v_k)(y_k + u_k)\,, \tag{24}$$

$$z_{k+1} = z_k + v_k(y_k + u_k) - C_\varepsilon u_k\,, \tag{25}$$

$$\alpha_h = 1 - \exp(-\alpha h)\,, \quad \beta_h = \exp(\beta h) - 1\,, \quad v_k \in [-\alpha_h,\, \beta_h]\,. \tag{26}$$

It is also convenient to separate the effect of the two controls via the two-step description:

$$\begin{array}{ll} x_k^+ = x_k\,, & x_{k+1} = (1 + v_k)x_k^+\,, \\ y_k^+ = y_k + u_k\,, & y_{k+1} = (1 + v_k)y_k^+\,, \\ z_k^+ = z_k - C_\varepsilon u_k\,, & z_{k+1} = z_k^+ + v_k y_k^+\,. \end{array}$$

The 3-D plus time *game of kind* is the same pursuit-evasion game as in the continuous theory, and capturable states are here again defined by $z_k \ge W_k(x_k, y_k)$ where the sequence of functions $\{W_k\}_{k\in\mathbb{N}}$ is the uppervalue function of the 2-D plus time game of degree (23), (24) with payoff

$$G = M(x_N) + \sum_{k=0}^{N-1} \Big(-v_k(y_k + u_k) + C_{\varepsilon_k} u_k \Big)\,. \tag{27}$$

Straightforward application of Isaacs' equation (see [4]) yields

Proposition 3.1. *The value function of the above discrete-time game is the only solution of the recursion*

$$W_k(x,y) = \min_u \max_{v\in[-\alpha_h,\beta_h]} [W_{k+1}((1+v)x,(1+v)(y+u)) - v(y+u) + C_\varepsilon u] \tag{28}$$

with

$$\forall x,y\,, \quad W_N(x,y) = M(x)\,. \tag{29}$$

Equation (28) is equivalent to the two-step procedure

$$\begin{aligned} W_k^+(x,y^+) &= \max_{v\in[-\alpha_h,\beta_h]} [W_{k+1}((1+v)x,(1+v)y^+) - vy^+], \\ W_k(x,y) &= \min_u [W_k^+(x,y+u) + C_\varepsilon u]. \end{aligned}$$

The two-step formulation separates the maximization and minimization operations. It proves useful in the numerical implementation.

We also recall the following theorem from [4].

Theorem 3.1. *The functions* $(x,y) \mapsto W_k(x,y)$ *are all convex.*

Proof. Notice that $(x,y) \mapsto M(x)$ is convex. Assume that W_{k+1} is convex. Then $(x,y) \mapsto W_{k+1}((1+v)x,(1+v)y) - vy$ is convex, so that W_k^+ is the maximum of a family of convex functions, and hence is convex. Now, changing u in $-u'$, W_k appears as the inf-convolution of W_k^+ and the convex extended function $\Gamma(x,y)$ equal to $+\infty$ if $x \neq 0$ and to $C_\varepsilon(-y)$ (with $\varepsilon = \mathrm{sign}(-y)$) if $x = 0$. Hence it is convex. □

This, in turn, helps us in devising an efficient numerical procedure to compute that value. Because the functions $v \mapsto W_k((1+v)x,(1+v)y)$ are convex, the maximum in v is reached at either $v = -\alpha$ or $v = \beta$. As for the inf-convolution, it is easy to see that, for each fixed (k,x), one should look for

$$\begin{aligned} y_k^-(x) &= \max\{y \mid -C^+ \in \partial_y W_k(x,y)\}\,, \\ y_k^+(x) &= \min\{y \mid -C^- \in \partial_y W_k(x,y)\}\,. \end{aligned} \tag{30}$$

Then, for $y \in [y^-,y^+]$, W_k and W_k^+ coincide. For $y \le y^-$, W_k must be extended continuously with a slope in y equal to $-C^+$. For $y \ge y^+$, W_k must be extended continuously with a slope equal to $-C^-$:

$$W_k(x,y) = \begin{cases} W_k^+(x,y_k^-(x)) - C^+(y - y_k^-(x)) & \text{if } y \le y_k^-(x)\,, \\ W_k^+(x,y) & \text{if } y_k^-(x) \le y \le y_k^+(x), \\ W_k^+(x,y_k^+(x)) - C^-(y - y_k^+(x)) & \text{if } y \ge y_k^+(x)\,. \end{cases} \tag{31}$$

Implementing that procedure is much faster than computing a min via a standard search procedure.

Understanding the shape of the functions W_k will be useful in the sequel. We emphasize it in the following remark.

Remark 3.1.
- For $y < y^-$, for $0 < h < y^- - v$, $W_k(x,y) = W_k(x,y+h) + C^+h \le W_k^+(x,y)$, and $W_k(x,y) \le W_k(x,y-h) - C^-h$,
- for $y \in [y^-, y^+]$, for all $h > 0$, $W_k(x,y) = W_k^+(x,y) \le W_k(x,y-h) - C^-h$, and $W_k(x,y) \le W_k(x,y+h) + C^+h$,
- for $y > y^+$, for $0 < h < v - y^+$, $W_k(x,y) = W_k(x,y-h) - C^-h \le W_k^+(x,y)$, and $W_k(x,y) \le W_k(x,y+h) + C^+h$.

3.2 Convergence

3.2.1 Main Theorem

We introduce the function $W^h(t,x,y)$ defined as the linear interpolation in time of the functions $W_k(x,y)$ and $W_{k+1}(x,y)$ where $kh \le t < (k+1)h$, and where the functions W_k are given by Equations (28) and (29) for a time step h (in (26)). The objective of this section is to prove the following theorem.

Theorem 3.2. *Let $h = 2^{-d}T$. As d goes to infinity, the sequence of functions $\{W^h\}$ converges uniformly on every compact (and monotonously decreasing) to the value function W of the continuous-time, impulse control game of degree.*

To prove this theorem, we need to introduce another way of looking at the impulse control problem, via yet another game. Thus we name our games. Let $\mathcal{G}$ be the original, continuous-time game, with controls u either finite or impulsive. Its (upper) value function is W. We shall also use the game $\mathcal{G}'$ which is the same as $\mathcal{G}$, but where $\mathbf{P}$ may only use impulses. Let $\mathcal{G}^h$ be the discretized game of this section, and its upper value the sequence $\{W_k^h\}_k$ (the W_k's above). Let also $\mathcal{G}^{h,\ell}$ be the discrete-time game with time step h where, in addition, the variable u has been discretized with a step ℓ, i.e., $u_k \in \ell\mathbb{Z}$. Its value function is a sequence $\{W_k^{h,\ell}\}_{k\in\mathbb{N}}$, which we interpolate in a function $W^{h,\ell}(t,x,y)$ as we did for W^h.

3.2.2 Joshua's Transformation

Finally, we introduce a game $\mathcal{J}$ according to an idea initially due to Joshua [11]. The players are still $\mathbf{P}$ and $\mathbf{E}$ as previously, but $\mathbf{P}$ has a control j which can take only the values -1, 0 or +1. We shall for convenience let $\bar{\jmath} = 1 - |j|$. The game happens in an artificial time that we call τ. We denote with a prime the derivatives with respect to τ. The natural time is now a state variable, and the final τ is defined as the first instant $\tau = \mathcal{T}$ such that $t(\mathcal{T}) = T$. The dynamics of the game are

$$
\begin{aligned}
t' &= \bar{\jmath}\,, \\
x' &= \bar{\jmath} v x\,, \\
y' &= \bar{\jmath} v y + j\,,
\end{aligned}
$$

and the payoff is

$$J = M(x(\mathcal{T})) + \int_0^{\mathcal{T}} (-\bar{\jmath} v y + C_j j)\,\mathrm{d}\tau\,,$$

(with C_0 arbitrary, 0 for instance).

Observe that this is now a standard differential game, which no longer has impulse controls. Its Isaacs equation can be written in the following way:

$$
\begin{aligned}
0 = \min\Bigg\{ & \frac{\partial W}{\partial t} + \max_{v\in[-\alpha,\beta]} v \left[\frac{\partial W}{\partial x} x + \left(\frac{\partial W}{\partial y} - 1 \right) y \right], \\
& \frac{\partial W}{\partial y} + C^+, - \left(\frac{\partial W}{\partial y} + C^- \right) \Bigg\}.
\end{aligned}
$$

This is a less degenerate form of the quasi-variational inequality of [4]:

$$
\begin{aligned}
0 = \min\Bigg\{ & \frac{\partial W}{\partial t} + \max_{v\in[-\alpha,\beta]} v \left[\frac{\partial W}{\partial x} x + \left(\frac{\partial W}{\partial y} - 1 \right) y \right], \\
& \min_u [W(t,x,y+u) - W(t,x,y) + C_\varepsilon u] \Bigg\}
\end{aligned}
$$

(where we required $\partial W/\partial y \in [-C^+, -C^-]$ everywhere).

We claim the important following fact.

Proposition 3.2. *The game $\mathcal{J}$ has the same value as the game $\mathcal{G}$.*

Proof. The game $\mathcal{J}$ is in fact completely equivalent to the game $\mathcal{G}'$. When **P** chooses a control $j = 0$, the game proceeds exactly as the game $\mathcal{G}'$ between two impulses. When **P** chooses $j = +1$ or -1, the time stops (hence the reference to Joshua), and y evolves in no real time of a quantity equal to j times the duration, in artificial time, of that control, at a cost C_ε times the variation of y.

The rest of the proof depends on the following easy lemma.

Lemma 3.1. *For any* **P***'s control strategy φ in the game $\mathcal{G}$, and any positive δ, there exists an admissible (causal) strategy in the game $\mathcal{G}'$ that yields against any admissible $v(\cdot)$ a payoff within δ of the payoff obtained with φ in the game $\mathcal{G}$.*

The proof of the lemma is given in Appendix C. We immediately have the following:

Corollary 3.1. *The games $\mathcal{G}$ and $\mathcal{G}'$ have the same value.*

And this, together with the fact that $\mathcal{J}$ and $\mathcal{G}'$ have the same value, proves the proposition. □

To complete the proof of the theorem, we need two more lemmas.

Lemma 3.2. *For every positive h, ℓ and every $(t = kh, x, y)$, $N \geq k \in \mathbb{N}$, one has*

$$W(t, x, y) \leq W_k^h(x, y) \leq W_k^{h,\ell}(x, y)\,. \tag{32}$$

Proof. We notice that due to our choice of α_h and β_h in (26), the quantity

$$\exp\left(\int_{t-h}^{t} v(\tau)\,\mathrm{d}\tau\right)$$

exactly spans the interval $[-\alpha_h,\ \beta_h]$. As a consequence, due to the linearity of the dynamics, the game $\mathcal{G}^h$ is an exact time sampling of the game $\mathcal{G}'$ where $\mathbf{P}$ is further constrained to placing its impulses at time instants $t_k = kh$, $k \in \mathbb{N}$. Since constraints have been placed on the admissible strategies of the minimizer, but not on the controls of the maximizer, we have the first inequality in (32). (Here and in the next lemma, the fact that $v(\cdot)$ can be taken open loop in defining the upper value plays a crucial role.)

In the game $\mathcal{G}^{h,\ell}$, further constraints are placed on the admissible strategies of $\mathbf{P}$. Hence the second inequality follows. □

Lemma 3.3. *The functions W^h and $W^{h,h}$ with $h = 2^{-d}T$ decrease as $d \to \infty$ and converge, uniformly on any compact, to functions $\widehat{W}$ and $\widetilde{W}$ respectively.*

Proof. We have noticed that the various games $\mathcal{G}^h$ are variants of the game $\mathcal{G}'$. They differ by the frequency at which player $\mathbf{P}$ is allowed to play. The game with $h = 2^{-d}T$ can be considered itself as a variant of the game with $h = 2^{-(d+1)}T$ but where $\mathbf{P}$ is constrained to play $u = 0$ at every odd-numbered stage. Since $\mathbf{P}$ is minimizing, this constraint increases the value of the game. Hence $W^h(t, x, y)$ is decreasing for every fixed (t, x, y). Being bounded from below by zero, it converges to some $\widehat{W}(t, x, y)$. Now, the W^h are convex, thus continuous, in (x, y), and continuous in time by construction. We therefore have a monotonous convergence of continuous functions, hence it is uniform on every compact.

Concerning the functions $W_k^{h,h}$, they correspond to games where a further constraint has been imposed on u. And again, for $h = 2^{-d}T$, the admissible u's for $d + 1$ are a superset of those admissible for d. Hence the value function decreases also. The rest follows. □

The main theorem is now a consequence of a last lemma.

Lemma 3.4. *Let W be the value function of the game $\mathcal{G}$ and $\widetilde{W} = \lim_{h\to 0} W^{h,h}$. Then*

$$\widetilde{W} = W\,. \tag{33}$$

Proof. The detailed proof is given in Appendix C. It uses the method of [10] for the game $\mathcal{J}$, whose value is W according to Corollary 3.1, and uses the fact that it follows from our analysis of (28), and specifically from Remark 3.1, that the sequence $\{W_k^{h,h}\}_k$ can be identified with the value function of the discretized version of the game $\mathcal{J}$. Hence (33) follows. □

Proof of the Main Theorem. of the main theorem: It follows from Lemma 3.3 that W^h converges to some $\widehat{W}$ as $h \to 0$ in a dyadic way. It follows from (32) that $W \le \widehat{W} \le \widetilde{W}$, and from (33) that $\widehat{W} = W$. □

4 Numerical results

We have implemented the recursion (28). We have used the two-step formulation and the procedures of Section 3.1 for the maximization and minimization. We are, of course, obliged to discretize x and y. To evaluate W and W^+ between discretization points, we have used a piecewise affine interpolation on triangles, and to evaluate them beyond the domain of discretization (the evaluation at $((1+\beta)x, (1+\beta)y)$ may require that), a linear extrapolation. Notice that this affine interpolation is essentially equivalent to the space discretization procedure analyzed in [13,14]. Hence, we may expect it to converge to the desired function as the discretization step goes to zero.

We have found that in some very narrow ranges of discretization steps, depending on the parameters, one may get wide numerical instabilities. Yet, being carefull to validate the results as “reasonable,” we have a very efficient program. With a 600×600 grid in the (x, y) domain, it runs in about .22 second per time step on a 1.7 GHz PC.

The numerical results corroborate our continuous-time theory. The results we discuss here correspond to the following set of parameters: a time step of $h = 0.02$, $\alpha = .10$, $\beta = .15$, $c_0 = .02$, $c_1 = .05$, and a discretization step of .005 in x and y.

For large t's (the first time steps) the program finds y^- and y^+ at both ends of the domain of y. Then for $T-t$ larger than .52, it finds y^- within the range of discretization and y^+ at the boundary. The theoretical value is $T-t_\beta = .46$. For $T-t$ larger than .70, it finds y^- and y^+ either equal or within one discretization step, the latter being a normal discretization effect. The theoretical value is $T - t_\alpha = .71$. Thus t_α has been recovered with a good accuracy (within one time step) while t_β is recovered with an error of three time steps. When y^- and y^+ differ within one discretization step, we have taken $(1/2)(y^- + y^+)$ as the approximation of $\check{y}$, and the smallest W as the approximation of $\check{z}$.

We have also implemented a numerical integration of the differential equations for $\mathcal{E}$ (or used the closed form found later. It makes no observable difference) and of the PDE for $\mathcal{F}$. The latter can exhibit numerical instabilities with bad choices of method. We got good results with a second-order centered finite difference scheme in "space" and a Runge–Kutta method of order two in time. Our computer code (in MATLAB) is still far from being optimized in terms of computing time. Thus this aspect will not be discussed here.

We have made the comparisons with a short maturity of $T = 5$ to save computation time in the computation of the focal surface. Both methods gave the same graphs for $\check{y}$ within one or two discretization steps ($.005Z$), except close to the boundary of the discretized domain, and almost the same graphs for $\check{z}$ to within two discretization steps, the discrete time $\check{z}$ being slightly larger, as expected. Both graphs are plotted in Figure 2.

5 Variants and Related Works

5.1 Another Terminal Target

Another game, maybe more significant for the finance application, but less rich in terms of game theory, is obtained by replacing $M(x)$ by $N(x,y) := M(x) + C_\varepsilon(-y)$, with $\varepsilon = \operatorname{sign}(-y)$. (See [4] for a motivation.) Then the sheet (α) of the barrier does not exist any longer, and thus neither does the dispersal manifold. The first singularity met (rearward) is an equivocal junction on the sheet (β), and before (in forward time) a focal surface. The theory is essentially the same.

5.2 The Viability Approach

In a series of papers [1,15] and in private communications, Aubin, Saint-Pierre, Pujal, and collaborators have considered, with the same motivation, essentially the same continuous-time problem, slightly more general in some aspects (they allow for constraints that were not considered in our work). They put a bound on the magnitude of our u to avoid impulses. But this is mainly for theoretical reasons, to get existence results for the viscosity solution of Isaacs' equation, an issue we did not tackle. They use a capture basin type of approach (similar to our *game of kind* approach) and discretize the corresponding PDE, leading to the same recursion (28) as ours, or a slightly different one depending on whether they use an explicit or implicit scheme. There, if taken large enough, the bounds on u are inactive.

A noteworthy feature is Saint-Pierre's "decoupling algorithm," which, for the above variant (Section 5.1), let him compute the locus $\check{y}$ and $\check{z}$, as the locus of

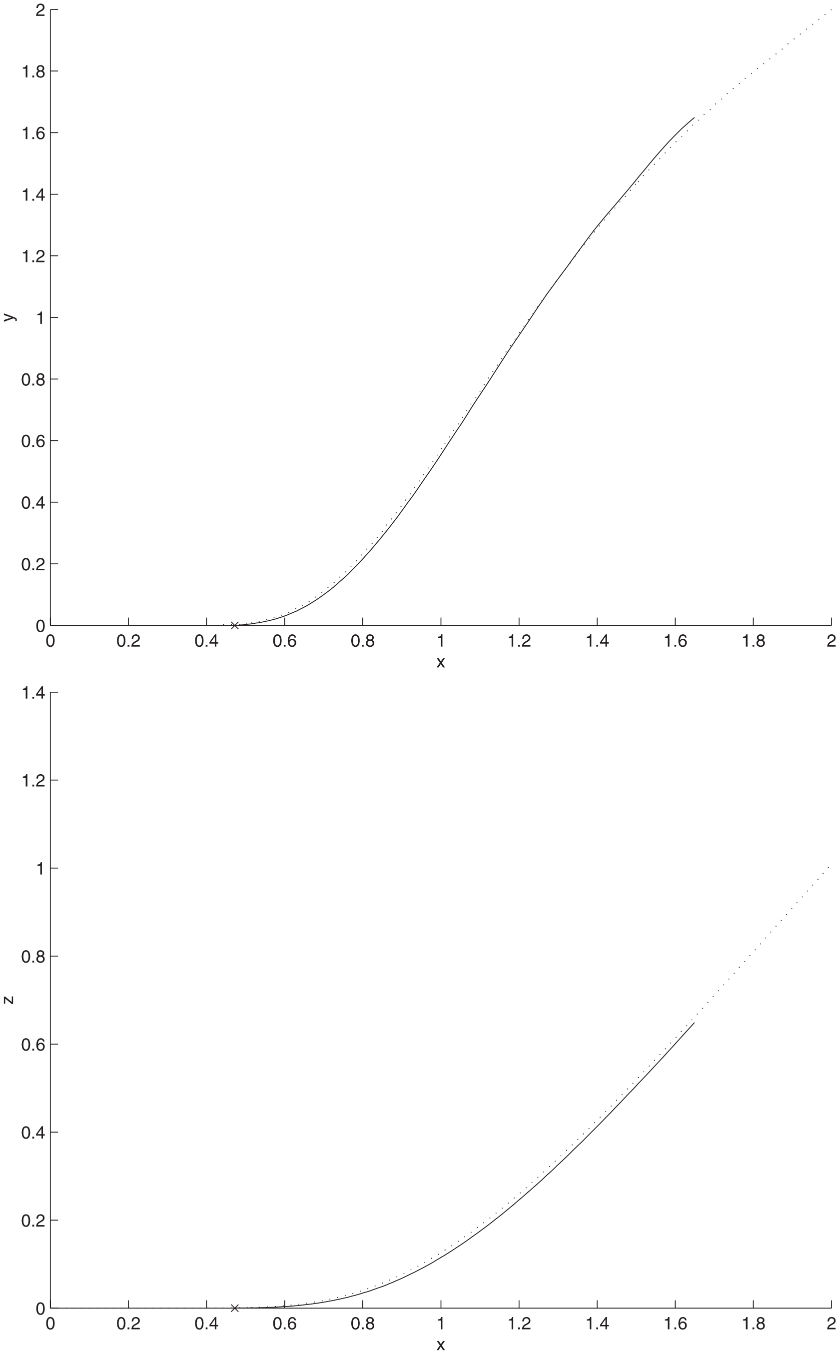


Figure 2: Cut of the focal manifold $\mathcal{F}$: $\check{y}(t, x)$ and $\check{z}(t, x)$ for $T - t = 5$. Dotted line: discrete time. Solid line: continuous time.

the minimum in y of the solution, with a computing effort comparable to two dynamic programming algorithms in dimension one instead of one in dimension two. Coupled with our theory as synthesized above, this is the fastest known way to compute the value of that game.

6 Conclusion

We have provided two closely related ways of investigating impulse controls in a differential game, both linked to the fact that “jump trajectories” can be regarded as ordinary trajectories. In this game, the optimal strategy of the pursuer contains both an impulse at initial time, and finite controls later on as the state traverses the singular manifolds. Admittedly, here the optimal strategy has the weakness that it needs to sense instantaneously the opponent’s control, i.e., here the time derivative of the first state variable. Breakwell has discussed this feature and approximate implementation in other papers [7,8]. Here, our discrete-time theory points to a practical solution of that problem.

This approach is feasible only because the cost of jumps was supposed to be proportional to the amplitude of the jump. It would be interesting to consider a cost affine in the amplitude, with a positive infimum. This would probably entail an investigation of the QVI according to the theory of [3].

More significantly perhaps, this analysis proves correct an old conjecture by Arik Melikyan that in higher dimensions, focal surfaces would be traversed by noncollinear optimal fields of trajectories. We have shown in detail that this is indeed the case here.

There remains to derive from the above analysis a general construction of higher-dimensional focal surfaces, which was missing in our constructive theory of singularities of co-dimension one in the Isaacs equation of (deterministic) two-person zero-sum differential games [5,6].

We have also proved and checked numerically that the continuous-time solution can be approached by the natural discrete-time game associated to our differential game. Yet, while that approach lets one numerically compute the value function, it does not give the more explicit form of Section 2.5, nor our detailed description of the optimal continuous-time strategies.

Appendix A: Equations of the Manifold $\mathcal{E}$

We recall the equations of the manifold $\mathcal{E}$:

$$\begin{aligned}
\dot{x} &= \beta x \,, \\
\dot{y} &= \frac{\beta(1+C^+) + \alpha \mathrm{e}^{-\alpha(T-t)}}{1+C^+ - \mathrm{e}^{-\alpha(T-t)}} y \,, \\
\dot{z} &= \frac{\beta(1+C^+)(1-\mathrm{e}^{-\alpha(T-t)}) - C^+ \alpha \mathrm{e}^{-\alpha(T-t)}}{1+C^+ - \mathrm{e}^{-\alpha(T-t)}} y \,,
\end{aligned}$$

to be integrated backwards from the terminal states $(t_0, x(t_0), y(t_0), z(t_0))$ either on the boundary of the manifold $\mathcal{D}$ at $t = t_\beta$ or on the boundary parametrized by $x = Z\mathrm{e}^{(\alpha(T-t))}$, $t_\alpha \le t \le t_\beta$. These equations admit a closed form solution as follows:

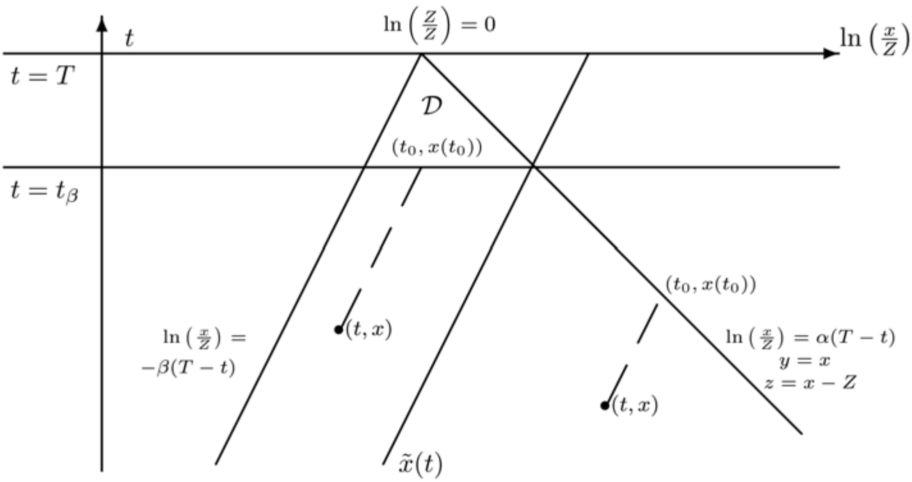
$$\begin{aligned}
x(t) &= x(t_0)e^{\beta(t-t_0)} \,, \\
y(t) &= y(t_0)e^{\beta(t-t_0)} \left(\frac{1 + C^+ - e^{\alpha(t_0-T)}}{1 + C^+ - e^{\alpha(t-T)}}\right)^{\frac{\alpha+\beta}{\alpha}} \,, \\
z(t) &= (1 - e^{-\alpha(T-t)})y(t) + z(t_0) - (1 - e^{-\alpha(T-t_0)})y(t_0),
\end{aligned}$$

as can be checked by direct differentiation. The expressions for y and z can be rewritten in terms of x and t, upon substituting for t_0, $y(t_0)$, and $z(t_0)$, to yield $\check{y}(t, x)$ and $\check{z}(t, x)$.

Let $\tilde{x}(t) = Ze^{(\alpha+\beta)(T-t_\beta)}e^{-\beta(T-t)}$:

$$\begin{cases}
\text{if } x \le \tilde{x}(t): & \begin{cases} t_0 = t_\beta \,, \\ y(t_0) = \frac{xe^{\beta(T-t)} - Z}{1 + C^+ - e^{-\alpha(T-t_\beta)}} \,, \\ z(t_0) = (1 - e^{-\alpha(T-t_\beta)})y(t_0) \,, \end{cases} \\[2ex]
\text{if } x \ge \tilde{x}(t): & \begin{cases} t_0 = \frac{1}{\alpha+\beta}(\alpha T + \beta t)\ln\left(\frac{x}{Z}\right) \,, \\ y(t_0) = xe^{\beta(T-t_0)} \,, \\ z(t_0) = y(t_0) - Z \,. \end{cases}
\end{cases}$$

We can remark that in the region $x \le \tilde{x}(t)$ we still have $w(t) = (1 - e^{-\alpha(T-t)})y(t)$ as on the manifold $\mathcal{D}$.



Appendix B: The PDE for the Focal Manifold $\mathcal{F}$

B.1 Analysis

As the trajectories $v = \beta$ and $v = -\alpha$ describe the same focal manifold $\mathcal{F}$ in (s, x, y, z) space, we only solve the PDE system (17), which we rewrite as

$$Y_t^\beta = sA^\beta Y_s^\beta + B^\beta Y^\beta.$$

Let us pick

$$\sigma = \ln\left(\frac{s}{Z}\right),$$

which transforms the PDE system in a linear PDE system of first order with constant coefficients in (t, σ):

$$Y_t^\beta = A^\beta Y_\sigma^\beta + B^\beta Y^\beta. \tag{34}$$

Moreover, the domain of interest Ω simplifies into the new domain in (t, σ): $\Omega_\sigma := \{t \le t_\alpha,\ 0 \le \sigma \le (\alpha + \beta)(T - t)\}$.

We notice that the known solutions $\breve{y}$ and $\breve{z}$ outside Ω, namely $(0, 0)$ to the "left" of Ω and $(x, x - Z)$ to the right, satisfy the PDE for $\mathcal{F}$ (34). Moreover, we have the following fact.

Proposition B.1. *If (34) admits a continuous solution on* $[0, T] \times [-\infty, \infty]$ *(in the domain* (t, σ)*) with simple discontiuities in* $(\nabla y, \nabla z)$*, these discontinuities are born by lines of slope 0 or* $-(\alpha + \beta)$ *in the plane* t, σ*.*

Hence, if such discontinuities follow from the discontinuity in $\sigma = 0$ at terminal time, they will precisely be born by the boundaries of Ω.

Proof. Let Δy_t, Δy_σ, Δz_t, and Δz_σ be the discontinuities. Let (p, q) be the direction of a smooth curve bearing the discontinuity in the (t, σ) domain. The continuity of both y and z implies that

$$p\begin{pmatrix}\Delta y_t\\ \Delta z_t\end{pmatrix} + q\begin{pmatrix}\Delta y_\sigma\\ \Delta z_\sigma\end{pmatrix} = 0\,.$$

Moreover, because at the discontinuity both sides satisfy the PDE (34), it follows that

$$\begin{pmatrix}\Delta y_t\\ \Delta z_t\end{pmatrix} = A^\beta \begin{pmatrix}\Delta y_\sigma\\ \Delta z_\sigma\end{pmatrix}\,.$$

Hence, combining these two equations, we obtain

$$(pA^\beta + qI)\begin{pmatrix}\Delta y_\sigma\\ \Delta z_\sigma\end{pmatrix} = 0.$$

Since, by hypothesis, the vector is nonzero, p cannot be 0, and $-q/p$ is an eigenvalue of A^β. These are 0 and $\alpha + \beta$. $\square$

B.2 Numerical Integration

We decided to use the fact that $\check{y}$ and $\check{z}$ are known outside of the domain of interest Ω in the numerical procedure. We compared this approach with a global integration relying on the preceding analysis. But the latter gave, not surprisingly, less precise results close to the boundary of Ω.

Hence, the boundary conditions in (t, σ) are also affine but the domain is not rectangular, the range in σ is a function of t. Let $\sigma_\ell = \sigma_0 + \ell\delta_\sigma$ where $\ell = 0, \ldots, N-1$ are the values of the discretization of the variable σ with a step of δ_σ on the domain $(t, \sigma) = [0, t_\alpha] \times [0, (\alpha+\beta)T + \delta_\sigma]$ including Ω_σ. We shall explain this choice hereafter.

At any time $t \leq t_\alpha$, we consider the vector of fixed dimension $2N \times 1$:

$$Y^\beta(t) = \begin{pmatrix} Y^\beta(t, \sigma_0) \\ \vdots \\ Y^\beta(t, \sigma_\ell) \\ \vdots \\ Y^\beta(t, \sigma_{N-1}) \end{pmatrix} \text{ with } Y^\beta(t, \sigma_\ell) = \begin{pmatrix} y(t, \sigma_\ell) \\ z(t, \sigma_\ell) \end{pmatrix} .$$

We denote by $Y^\beta_\sigma(t)$ the vector of the derivatives in σ of the vector $Y^\beta(t)$. We will approach it by finite differences in σ. Thus the PDE system leads to an ODE system of $2N$ equations of the form

$$Y^\beta_t(t) = M(t)Y(t). \tag{35}$$

The interest of the new domain is that we will work with a matrix $M(t)$ of constant dimension.

In the domain $\sigma \geq (\alpha+\beta)(T-t)$, we replace the PDE system (34) by the system satisfied by $x = y$, $z = x - Z$ with $\dot{x} = \beta x$, i.e.:

$$\left\{ \begin{array}{lcl} \dot{y} &=& \beta y \ , \\ \dot{z} &=& \beta y \ . \end{array} \right. \tag{36}$$

This leads to a matrix $M(t)$ whose lines corresponding to $\sigma \leq (\alpha+\beta)(T-t)$ implement Equation (34) while those corresponding to $\sigma \geq (\alpha+\beta)(T-t)$ implement Equation (36). Hence, M is time varying.

To solve the ODE system (35), we have tried different numerical methods of lower order (1 or 2). Some methods exhibit numerical instabilities, but we got good results with a second-order centered finite difference scheme in "space" and a Runge–Kutta method of order two in time.

Appendix C: Proofs of the Lemmas

C.1 Proof of Lemma 3.1

Lemma 3.1. *For any* **P***'s control srategy φ in the game $\mathcal{G}$, and any positive δ, there exists an admissible (causal) strategy in the game $\mathcal{G}'$ that yields against any admissible $v(\cdot)$ a payoff within δ of the payoff obtained with φ in the game $\mathcal{G}$.*

Proof. We shall only prove that y can be approximated uniformly arbitrarily well. The proof for its integral follows. In fact, integrating (2) yields

$$y(t) = y(0)\exp\left(\int_0^t v(\tau)\,\mathrm{d}\tau\right) + \bar{y}(t),$$

where only $\bar{y}$ depends on u. Thus it suffices to approximate $\bar{y}$. Let δ be a given positive number; we shall show how to approximate $\bar{y}(t)$ within δ uniformly in t and $v(\cdot)$.

Pick a strategy φ. For a disturbance $v(\cdot)$ given, it generates a time function (or distribution) $u(\cdot)$ that may contain impulses. One has

$$\bar{y}(t) = \int_0^t \exp\left(\int_s^t v(\tau)\,\mathrm{d}\tau\right) u(s)\,\mathrm{d}s\,.$$

We decompose $u(\cdot)$ as $u(t) = u^+(t) - u^-(t)$, its positive and negative parts (including the positive and negative impulses). In an obvious way, this induces a decomposition $\bar{y} = \bar{y}^+ - \bar{y}^-$.

Proposition C.1. *Under our hypotheses, we may assume that both $\bar{y}^+$ and $\bar{y}^-$ are uniformly bounded over all admissible $v(\cdot)$'s for any initial state.*

Proof of the Proposition. In investigating the value of the game $\mathcal{G}$, we may restrict our attention to strategies φ that do better than a given strategy φ_0. Choose, for instance, φ_0 as the strategy made of an initial jump to $y = 0$ at time $t = 0$ (i.e., $t_0 = 0$ and $u_0 = -y(0)$), and $u = 0$ from then on. It yields $z(T) = z(0) - C_\varepsilon(-y(0))$. Thus we restrict our attention to strategies φ that yield a larger $z(T)$ for all admissible $v(\cdot)$'s. Now, since $y(t)$ is by hypothesis uniformly bounded, so is $\int_0^T v(t)y(t)\,\mathrm{d}t$. According to (3), $z(T) = z(0) + \int_0^T (vy - C_\varepsilon u)\,\mathrm{d}t$. Therefore, $\int_0^T C_\varepsilon u(t)\,\mathrm{d}t$ is also uniformly bounded. But if we let $C = \max\{C^+, -C^-\}$ (C is positive), we have $C_\varepsilon u \geq C|u|$. Hence the integral of $|u|$ is uniformly bounded, and a fortiori those of u^+ and u^-, and finally also $\bar{y}^+$ and $\bar{y}^-$. □

We shall do the approximation for each of these two parts separately. Hence, from now on, we may assume that $u(t) \geq 0$, or more precisely that $\int |u(s)|\,\mathrm{d}s = \int u(s)\,\mathrm{d}s$.

Let therefore $y_{\max}$ be an (uniform) upper bound of $\bar{y}(t)$, pick ϵ (< 1 and) such that $\epsilon y_{\max} \leq \delta/2$, and let h be a positive number such that, for every admissible $v(\cdot)$, and $\forall t \in [h, T]$,

$$\left|\exp\left(\int_{t-h}^{t} v(\tau)\,\mathrm{d}\tau\right) - 1\right| \leq \frac{\epsilon}{2} \leq \frac{1}{2}$$

(this is possible uniformly in $v(\cdot)$ because $|v(t)|$ is bounded), and thus a fortiori, $\forall s \in [t-h, t]$,

$$\left|\exp\left(\int_{s}^{t} v(\tau)\,\mathrm{d}\tau\right) - 1\right| \leq \frac{\epsilon}{2} \leq \epsilon \exp\left(\int_{s}^{t} v(\tau)\,\mathrm{d}\tau\right). \tag{37}$$

We advocate the impulses-only strategy using impulses of amplitude u_k at the instants $t_k = kh$, $k \in \mathbb{N}$ as follows:

$$u_k = \int_{t_k-h}^{t_k} u(t)\,\mathrm{d}t\,.$$

This yields for $\bar{y}$ a time function that we denote $\tilde{y}$:

$$\tilde{y}(t) = \sum_{k|t_k<t} u_k \exp\left(\int_{t_k}^{t} v(\tau)\,\mathrm{d}\tau\right).$$

The difference $\Delta(t) = |\bar{y}(t) - \tilde{y}(t)|$ can be written as

$$\Delta(t) = \left|\sum_{k|t_k<t} \int_{t_k-h}^{t_k} \left[\exp\left(\int_{s}^{t_k} v(\tau)\,\mathrm{d}\tau\right) - 1\right] u(s)\,\mathrm{d}s \exp\left(\int_{t_k}^{t} v(\tau)\,\mathrm{d}\tau\right)\right|,$$

hence

$$\Delta(t) \leq \sum_{k|t_k<t} \int_{t_k-h}^{t_k} \left|\exp\left(\int_{s}^{t_k} v(\tau)\,\mathrm{d}\tau\right) - 1\right| u(s)\,\mathrm{d}s \exp\left(\int_{t_k}^{t} v(\tau)\,\mathrm{d}\tau\right).$$

According to (37),

$$\Delta(t) \leq \epsilon \sum_{k|t_k<t} \int_{t_k-h}^{t_k} \exp\left(\int_{s}^{t_k} v(\tau)\,\mathrm{d}\tau\right) u(s)\,\mathrm{d}s \exp\left(\int_{t_k}^{t} v(\tau)\,\mathrm{d}\tau\right) = \epsilon\bar{y}(t)\,.$$

Hence, for each of the positive and negative parts of $\bar{y}$ we have

$$\tilde{y}(t) \in [(1-\epsilon)\bar{y}(t)\,,\ (1+\epsilon)\bar{y}(t)] \subset \left[\bar{y}(t) - \frac{\delta}{2}\,,\ \bar{y}(t) + \frac{\delta}{2}\right].$$

C.2 Proof of Lemma 3.4

We consider the following discrete scheme associated to Joshua's transform:

$$\begin{cases} t((k+1)h) &= t(kh) + h\bar{\jmath}, \\ x((k+1)h) &= x(kh) + h\bar{\jmath}vx(kh), \\ y((k+1)h) &= y(kh) + h(\bar{\jmath}vy(kh) + j), \end{cases}$$

with the payoff defined with $t(\mathcal{T}) = T$ and

$$J = M(u(\mathcal{T})) + \sum_{k=0}^{N-1} h(-\bar{\jmath}vy(kh) + C_j j)\,,$$

and the controls $j \in \{-1, 0, 1\}$ and $v \in [-\alpha^h, \beta^h]$, where (see (26)):

$$\alpha^h = \frac{\alpha_h}{h} = \frac{1}{h}(1 - \mathrm{e}^{-\alpha h})\,, \qquad \beta^h = \frac{\beta_h}{h} = \frac{1}{h}(\mathrm{e}^{\beta h} - 1)\,.$$

We notice that $\alpha^h \to \alpha$ and $\beta^h \to \beta$ as $h \to 0$.

The Isaacs equation of the above multistage game concerns a function V^h and reads:

$$\begin{aligned} &\forall (t,x,y) \in [0,T] \times \mathbb{R}^+ \times \mathbb{R}\,, \\ &0 = \min_{j\in\{-1,0,1\}} \max_{v\in[-\alpha^h,\beta^h]} \Big[V^h(t + h\bar{\jmath}, x + h\bar{\jmath}vx, y + h(\bar{\jmath}vy + j)) \\ &\qquad\qquad\qquad\qquad - V^h(t,x,y) + h(-\bar{\jmath}vy + C_j j)\Big], \\ &\forall t \geq T\,, \quad V^h(t,x,y) = M(x). \end{aligned} \tag{38}$$

Now, we want to prove that V^h converges towards V, where V is the viscosity solution of the following Isaacs equation, associated to the continuous Joshua form:

$$0 = \min_j \max_{v\in[-\alpha,\beta]} \left[\frac{\partial V}{\partial t}\bar{\jmath} + \frac{\partial V}{\partial x}\bar{\jmath}vx + \frac{\partial V}{\partial y}(\bar{\jmath}vy + j) + (-\bar{\jmath}vy + C_j j) \right]$$

with the same boundary condition.

We recall the definition of a viscosity solution of the last Isaacs equation. A bounded uniformly continuous function V is called a viscosity solution of the Isaacs equation above if for each $\phi \in C^1(\mathbb{R}^3)$, the following hold:

(1) if $V - \phi$ attains a strict local maximum at $a_0 = (t_0, x_0, y_0)$, then

$$\min_j \max_v \left[\frac{\partial \phi}{\partial t}(a_0)\bar{\jmath} + \frac{\partial \phi}{\partial x}(a_0)\bar{\jmath}vx_0 + \frac{\partial \phi}{\partial y}(a_0)(\bar{\jmath}vy_0 + j) - \bar{\jmath}vy_0 + C_j j \right] \geq 0\,,$$

(2) if $V-\phi$ attains a strict local minimum at $a_1=(t_1,x_1,y_1)$, then

$$\min_j \max_v \left[\frac{\partial\phi}{\partial t}(a_1)\bar{\jmath}+\frac{\partial\phi}{\partial x}(a_1)\bar{\jmath}vx_1+\frac{\partial\phi}{\partial y}(a_1)(\bar{\jmath}vy_1+j)-\bar{\jmath}vy_1+C_j j\right]\le 0\,.$$

Proof. Notice first that, expanding the $\min_j$ according to the three possible values of j, and replacing $hv\in[-\alpha^h,\beta^h]$ by
$v\in[-\alpha_h,\beta_h]$, (38) also reads

$$\min\Big\{\max_{v\in[-\alpha_h,\beta_h]}\left[V^h(t+h,(1+v)x,(1+v)y)-V^h(t,x,y)-vy\right]\,,\\ V^h(t,x,y-h)-C^-h\,,\quad V^h(t,x,y+h)+C^+h\Big\}\,,$$

so that, using Remark 3.1 we may conclude that V^h coincides with $W^{h,h}$. Thus, we know from Lemma 3.3 that there exists a V (called W in the body of the paper) such that

$$V^h\to V\qquad \text{uniformly on any compact of } \mathbb{R}^3\text{when}\quad h\to 0\,. \tag{39}$$

Let $\phi\in C^1(\mathbb{R}^3)$ and a_0 be a strict local maximum for $V-\phi$. Then there exists a closed ball B centered at a_0 such that

$$(V-\phi)(a_0)>(V-\phi)(a),\qquad \forall a\in B. \tag{40}$$

Let now a_0^h be a maximum point for $V^h-\phi$ over B.

Lemma C.1.

$$a_0^h\to a_0,\qquad \text{when}\quad h\to 0. \tag{41}$$

Proof. a_0^h remains in the compact B. Let $\bar{a}$ be a cluster point of the sequence $\{a_0^h\}$ and $\{a_0^{h_i}\}$ be a subsequence converging to $\bar{a}$. By definition we have that $(V^{h_i}-\phi)(a_0^{h_i})\ge(V^{h_i}-\phi)(a)$, for all $a\in B$, and then, by continuity of V^{h_i} and ϕ and using (39), we get $(V-\phi)(\bar{a})\ge(V-\phi)(a)$, $\forall a\in B$. By unicity of the maximum, we have that $\bar{a}=a_0$. The cluster point $\bar{a}$ is then unique, so the whole sequence a_0^h converges towards a_0. $\square$

Now since $h\to 0$, we have that $(t_0^h+h\bar{\jmath},x_0^h+h\bar{\jmath}vx_0^h,y_0^h+h(\bar{\jmath}vy_0^h+j))$ remains in B. Since a_0^h is a maximum point for $V^h-\phi$ over B, we have

$$V^h(t_0^h,x_0^h,y_0^h)-\phi(t_0^h,x_0^h,y_0^h)\\ \ge V^h\Big(t_0^h+h\bar{\jmath},x_0^h+h\bar{\jmath}vx_0^h,y_0^h+h(\bar{\jmath}vy_0^h+j)\Big)\\ -\phi\Big(t_0^h+h\bar{\jmath},x_0^h+h\bar{\jmath}vx_0^h,y_0^h+h(\bar{\jmath}vy_0^h+j)\Big),$$

Using the last inequality together with Equation (38) and also using the monotonicity of the "minmax" operator, we get the following, where v is always understood to lie in $[-\alpha^h, \beta^h]$:

$$\begin{aligned} 0 &= \min_j \max_v \Big[V^h(t_0^h + h\bar{\jmath}, x_0^h + h\bar{\jmath} v x_0^h, y_0^h + h(\bar{\jmath} v y_0^h + j)) \\ &\qquad\qquad - V^h(t_0^h, x_0^h, y_0^h) + h(-\bar{\jmath} v y_0^h + C_j j) \Big] \\ &\le \min_j \max_v \Big[\phi(t_0^h + h\bar{\jmath}, x_0^h + h\bar{\jmath} v x_0^h, y_0^h + h(\bar{\jmath} v y_0^h + j)) \\ &\qquad\qquad - \phi(t_0^h, x_0^h, y_0^h) + h(-\bar{\jmath} v y_0^h + C_j j) \Big]. \end{aligned}$$

Since $\phi \in C^1(\mathbb{R}^n)$, from the last inequality, we get

$$0 \le \min_j \max_v h \left[\frac{\partial \phi}{\partial t}(b^h)\bar{\jmath} + \frac{\partial \phi}{\partial x}(b^h)\bar{\jmath} v x_0^h + \frac{\partial \phi}{\partial y}(b^h)(\bar{\jmath} v y_0^h + j) - \bar{\jmath} v y_0^h + C_j j \right],$$

where b^h is in the segment $[(t_0^h, x_0^h, y_0^h), (t_0^h + h\bar{\jmath}, x_0^h + h\bar{\jmath} v x_0^h, y_0^h + (\bar{\jmath} v y_0^h + j))]$. Since $h > 0$, we may divide through by h; then it follows that

$$0 \le \min_j \max_v \left[\frac{\partial \phi}{\partial t}(b^h)\bar{\jmath} + \frac{\partial \phi}{\partial x}(b^h)\bar{\jmath} v x_0^h + \frac{\partial \phi}{\partial y}(b^h)(\bar{\jmath} v y_0^h + j) - \bar{\jmath} v y_0^h + C_j j \right].$$

Since a_0^h converges towards a_0 and since h converges towards zero, it follows that b^h also converges towards a_0. Moreover, the bracket is continuous in $(v, (t, x, y))$, therefore in (v, h) and therefore uniformly continuous in (v, h) in a (closed) neighborhood of $[-\alpha, \beta] \times \{0\}$. Thus we may pass to the limit for each value of j to conclude that

$$\min_j \max_{v \in [-\alpha, \beta]} \left[\frac{\partial \phi}{\partial t}(a_0)\bar{\jmath} + \frac{\partial \phi}{\partial x}(a_0)\bar{\jmath} v x_0 + \frac{\partial \phi}{\partial y}(a_0)(\bar{\jmath} v y_0 + j) - \bar{\jmath} v y_0 + C_j j \right] \ge 0.$$

The proof is the same, mutatis mutandis, for point (2) of the definition of the viscosity solution of the Isaacs equation considered. □

REFERENCES

[1] Aubin J.-P., Pujal D. and Saint-Pierre P., *Dynamic Management of Portfolios with Transaction Costs under Tychastic Uncertainty*, preprint, 2001.

[2] Barles G., Deterministic impulse control problems, *SIAM Journal on Control and Optimization*, 23, 419–432, 1985.

[3] Bensoussan A. and Lions J-L., *Contrôle impulsionnel et inéquations quasivariationnelles*, Gauthier-Villars, Paris, 1982.

[4] Bernhard P., A robust control approach to option pricing including transaction costs, in A. Nowak (ed.), *Annals of the ISDG*, 7, Birkhauser, 2002.

[5] Bernhard P., Singular surfaces in differential games: an introduction, in P. Hagedorn, H.W. Knobloch, and G.J. Olsder (eds.), *Differential Games and Applications*, Proceedings of a workshop held in Twente University of Technology, 1977, Springer Verlag Lecture Notes in Control and Information Sciences, 3, 1–33, 1977.

[6] Bernhard P., *Contribution à l'étude des jeux différentiels à deux joueurs et somme nulle*, Ph.D. Dissertation, Université Paris 6, 1989.

[7] Breakwell J.V., Time optimal pursuit inside a circle, in T. Başar and P. Bernhard (eds.), *Differential Games and Applications*, Springer Lecture Notes in Control and Information Sciences, 119, Springer, Berlin, 1989.

[8] Breakwell J.V. and Bernhard P., A simple game with a singular focal line, *JOTA*, 64, 1990.

[9] Breakwell J.V. and Merz A.W., Toward a Complete Solution of the Hommicidal Chauffeur Game, *First International Conference on the Theory and Applications of Differential Games*, Amherst, 1969.

[10] Capuzzo Dolcetta I., On a discrete approximation of the Hamilton–Jacobi equation of dynamic programming, *Applied Mathematics and Optimization*, 10, 367–377, 1983.

[11] Joshua: The Book of Joshua, 10:12–13, in *The Bible* (circa 1000 to 500 AD).

[12] Merz A.W., *The Homicidal Chauffeur: A Differential Game*, Ph.D. Dissertation, SUDAAR Report, Stanford University, 1971.

[13] Pourtallier O. and Tidball M., Approximation of the value function for a class of differential games with target, *Sixth International Symposium on Dynamic Games and Applications*, St. Jovite, Canada, July 1994.

[14] Pourtallier O., Tidball M. and Altman E., Approximations in dynamic zero-sum games, II, *SIAM J. Control and Optimization*, 35, 2101–2117, 1997.

[15] Pujal D., *Évaluation et gestion dynamiques de portefeuilles*, Ph.D. Dissertation, Université Paris Dauphine, 2000.

[16] Roorda B., Engwerda J. and Schumacher H., *Performance of Hedging Strategies in Interval Models*, to appear.

Incentives for Retailer Promotion in a Marketing Channel*

Steffen Jørgensen
Department of Organization and Management
University of Southern Denmark, Odense, Denmark
stj@sam.sdu.dk

Sihem Taboubi
GERAD and Department of Marketing, HEC Montréal
Montréal (Québec) Canada H3T 2A7
sihem.taboubi@hec.ca

Georges Zaccour
GERAD and Department of Marketing, HEC Montréal
Montréal (Québec) Canada H3T 2A7
georges.zaccour@gerad.ca

Abstract

This chapter analyzes a differential game model of a two-member marketing channel. A manufacturer invests in national advertising with the purpose of improving (or sustaining) the image of one of her brands, and her retailer makes local promotions for the brand. The game is played à la Stackelberg with the manufacturer as leader. We characterize and compare equilibria for two scenarios. In the first one, the manufacturer designs an incentive strategy to affect the retailer's promotion strategy with the objective of maximizing the total channel profit. In the second, the manufacturer's objective is the maximization of her own payoff.

Key words. Incentive strategies, channel coordination, advertising and promotion, Stackelberg equilibrium.

1 Introduction

This chapter considers a two-member channel of distribution and the setup concerns the sales of a particular brand of a manufacturer to the retailer. In the

*Research supported by NSERC, Canada. We wish to thank an anonymous reviewer for constructive comments on an earlier draft.

absence of cooperation, channel members determine their marketing decision strategies independently and noncooperatively. It is well known, in the literature and in practice, that uncoordinated decision making can create "channel inefficiencies," in the sense that channel members' noncooperative payoffs are less than those they could obtain in a coordinated setup.

Assume that the manufacturer can, and will, take the role of a channel leader, and that the retailer agrees to take the role of a follower. By a channel leader we mean an agent who announces her strategy in advance and commits to playing that strategy. The manufacturer wishes to induce the retailer to implement a certain outcome and, given the leader's announcement, the retailer can do no better than to react rationally to the leader's strategy.

The timing of events is the following. At the start of the game the manufacturer announces and commits to an incentive strategy. Note that commitment to a strategy does not mean commitment to a predetermined course of action. Rather, a strategy specifies what type of action to take at a decision point, contingent on available information. Typically, an incentive strategy of the manufacturer is conditioned upon the retailer's past or present actions (and even upon the past actions of the manufacturer).

The paper provides an answer to the question of how to achieve channel coordination, in two specific senses. How can the manufacturer, through her choice of marketing strategy, induce the retailer to implement an outcome that (i) is favored by both channel members or (ii) is favored by the manufacturer only? Each scenario involves a coordination problem and an assumption on what the manufacturer has in mind when attempting to coordinate the channel.

(1) The manufacturer wishes to induce the retailer, through the announcement of an incentive, to act in accordance with a *cooperative* outcome. By a cooperative outcome we mean the one which maximizes the sum of the two channel members' individual payoffs. In the sequel we refer to this as the *joint maximization outcome*. Note that the manufacturer acts in the best interest of the *channel* when seeking to maximize the overall profit. The resulting outcome is Pareto optimal (group rational). When the manufacturer acts in the best interest of the channel, it is plausible that the retailer will accept the role of a follower.

(2) The manufacturer wishes to induce the retailer, through the announcement of an incentive, to act in such a way that the *manufacturer's individual* payoff is maximized. The manufacturer acts in her *own* best interest and it is not evident that the retailer will accept the role of a follower. However, since the manufacturer announces and commits to her strategy, the retailer has no better option than to accept the role of a follower (and implements her best reply to the manufacturer's strategy).

The incentive problem under consideration is *asymmetric* (one-sided). One channel member (the manufacturer) assumes the role of a channel leader who designs an incentive for the retailer. The latter acts as a follower. Clearly, the roles of channel members can be reversed, making the retailer the channel leader. This is observed in practice when nation-wide or international retail chains deal with individual manufacturers. Also note that the assignment of roles is *exogenous*, that is, the manufacturer is the channel leader by assumption.

Studies of marketing channels have identified a number of mechanisms that can lead to coordination, under alternative hypotheses about the manufacturer's objective. Coordinating mechanisms are, for instance, quantity discounts, advertising allowances, and retailer promotion cost subsidies. Most of these works have assumed a *static* environment and the prime example of coordination is to remedy pricing inefficiency (the double marginalization problem). Here, incentives can be designed such that a retailer will choose a consumer price that maximizes total channel profits. Jeuland and Shugan [3] studied profit sharing as a coordinating device. Through a quantity discount scheme the manufacturer induces a relationship between total channel profits and channel members' individual profits such that if any one of these is maximized, they are all maximized. Moorthy [9], on the other hand, advocates an asymmetric relationship where the manufacturer assumes responsibility for the implementation of a coordinated outcome. Bergen and John [1] analyzed a situation where the manufacturer decides a transfer price and a subsidy for the retailer's local advertising expenditures.

Marketing decisions often have carry-over effects, and therefore they do not only affect the current payoff but also future ones. Capturing the full impact of decisions thus requires a dynamic model. There are still not many attempts to examine the channel coordination issue in an *intertemporal* setting. Jørgensen and Zaccour [4,5], Jørgensen, Sigué, and Zaccour [6,7], and Jørgensen, Taboubi, and Zaccour [8] use differential games to demonstrate that channel coordination can be achieved, primarily through advertising allowances. However, these studies did not address the question of whether coordinated outcomes can be sustained (enforced) over time. This issue is important and is actually the central point in this chapter. We shall show that sustainability can indeed be achieved; the key to this result is to design the incentive in such a way that a coordinated outcome becomes a (Nash) equilibrium.

In the scenario where the manufacturer wishes to implement the joint maximization outcome, the follower must be induced to select her part of the outcome at any instant of time. This means (among other things) that as of any instant of time, the retailer will obtain a payoff-to-go in the coordinated solution which exceeds that of the uncoordinated (noncooperative) solution. If such a "dynamic individual rationality" constraint is not satisfied at some instant of time, the retailer has an incentive to switch to noncooperative play. In the two scenarios to be considered here, the individual rationality problem is resolved from the viewpoint that the manufacturer announces and commits to her strategy in advance, and the retailer has no better choice than to implement her best reply. Under the assumption that the retailer cannot withdraw from the channel relationship, the manufacturer does not need to care about the satisfaction of dynamic individual rationality.

The paper proceeds as follows. Section 2 introduces a differential game model of the distribution channel. Section 3 addresses the incentive problem and identifies in Section 3.1 the joint maximization solution and the incentive needed for sustaining this outcome. Thus, Section 3.1 analyzes the problem posed above in Scenario 1, and establishes an equilibrium of the incentive game in which the manufacturer wishes to have the joint maximization outcome implemented. Section 3.2 establishes an equilibrium of the incentive game in which the manufacturer wishes to have her individual maximization outcome implemented, cf. Scenario 2. In Section 4 we compare the results obtained for the two scenarios. Section 5 states our conclusions.

2 A Differential Game

Consider the sales of a particular brand of a manufacturer to a single retailer. The manufacturer advertises the brand in national media, whereas the retailer promotes the brand locally. Suppose that the channel members have infinite planning horizons and let t denote time, $t \in [0, \infty)$. Let $A(t)$ represent the manufacturer's national advertising rate which influences the brand image, represented by the state variable $G(t)$. (In a capital accumulation sense, the brand image can be seen as a stock of advertising goodwill.) The evolution of the stock G is assumed to follow the rather classical formulation à la Nerlove–Arrow, i.e.,

$$\dot{G}(t) = A(t) - \delta G(t), \quad G(0) = G_0 > 0, \tag{1}$$

where $\delta > 0$ is a decay rate. Since a feasible control $A(t)$ must be nonnegative, it follows from (1) that the state constraint $G(t) \geq 0$ is satisfied for all t. This means that the brand image is never negative (which would mean negative goodwill (badwill)).

The retailer controls her local promotional activities, represented by the effort rate $P(t)$. Following Jørgensen, Sigué, and Zaccour [6,7], we assume that the retail sales revenue rate is given by

$$S(t) = \theta + \beta P(t)\sqrt{G(t)},$$

where $\theta > 0$ is a constant representing the baseline sales revenue for the brand, in the absence of promotional effort. The hypothesis here is that the retailer needs to promote the brand locally if she wishes to increase her sales revenue above its baseline level. The parameter $\beta > 0$ represents the effect of promotion on sales revenue. Note that retailer's promotional efforts are reinforced by the level of goodwill, but at a decreasing marginal rate.

We assume that both channel members face convex cost functions which we take as quadratic for tractability,

$$C_M(A(t)) = \frac{u_M A(t)^2}{2}, \quad C_R(P(t)) = \frac{u_R P(t)^2}{2}.$$

Obviously, the controls must take nonnegative values, $A(t) \geq 0, P(t) \geq 0$. Following Chintagunta and Jain [2] we suppose that a fixed percentage $0 < \pi < 1$ of retail sales revenue accrues to the manufacturer. Thus, a fixed profit sharing mechanism has been agreed upon before playing the game.

Letting ρ denote a constant and positive discount rate, the objective functionals of the manufacturer and the retailer, respectively, are

$$J_M = \int_0^\infty e^{-\rho t} \left[\pi S(t) - \frac{u_M A(t)^2}{2} \right] dt$$
$$J_R = \int_0^\infty e^{-\rho t} \left[(1-\pi) S(t) - \frac{u_R P(t)^2}{2} \right] dt.$$

3 Incentive Strategies

In this section we solve the two incentive problems outlined in Section 1. Section 3.1 deals with the case in which the manufacturer wishes to design a promotion incentive for the retailer such that the joint maximization outcome is implemented. In Section 3.2 we solve the analogous problem with the maximization of the manufacturer's individual profit as the objective. Section 4 compares the outcomes of the two incentive problems.

3.1 Joint Maximization

To identify the joint profit maximization (cooperative) outcome, we solve an optimal control problem with objective

$$J^C = J_M + J_R = \int_0^\infty e^{-\rho t} \left[S(t) - \frac{u_M A(t)^2}{2} - \frac{u_R P(t)^2}{2} \right] dt,$$

subject to the state dynamics in (1). Note that in the above objective, the revenue sharing parameter π vanishes and hence the optimal strategies and cooperative outcome will be obviously independent of π. The superscript C refers to "cooperative." The solution of the joint maximization problem is summarized in the following lemma.

Lemma 3.1. *The joint profit maximization problem has the following solution:*

$$A^C = \frac{\beta^2}{2 u_M u_R (\rho + \delta)} \tag{2}$$

$$P^C(G) = \frac{\beta\sqrt{G}}{u_R} \tag{3}$$

$$V^C(G) = \frac{\beta^2 G}{2u_R(\rho+\delta)} + \frac{\theta}{\rho} + \frac{\beta^4}{8u_M u_R^2 \rho(\rho+\delta)^2}. \tag{4}$$

Proof. We use a sufficiency theorem for the optimal control problem. One needs to find a bounded and continuously differentiable function $V^C(G)$ which for all $G \geq 0$ satisfies the Hamilton–Jacobi–Bellman (HJB) equation

$$\rho V^C = \max_{A^C, P^C \geq 0} \left[S - \frac{u_R}{2} P^2 - \frac{u_M}{2} A^2 + \frac{dV^C}{dG}(A - \delta G) \right].$$

Cooperative strategies, whenever they take positive values, are given by

$$A^C = \frac{1}{u_M}\frac{dV^C}{dG}, \quad P^C = \frac{\beta\sqrt{G}}{u_R}.$$

Substituting A^C and P^C into the HJB equation yields

$$\rho V^C = \theta + \frac{1}{2u_M}\left(\frac{dV^C}{dG}\right)^2 + \left[\frac{\beta^2}{2u_R} - \delta\frac{dV^C}{dG}\right] G. \tag{5}$$

It is readily verified that the ordinary differential equation in (5) admits the solution

$$V^C(G) = \frac{\beta^2 G}{2u_R(\rho+\delta)} + \frac{\theta}{\rho} + \frac{\beta^4}{8u_M u_R^2 \rho(\rho+\delta)^2}.$$

Finally, substituting $dV^C/dG = \beta^2/[2(\rho+\delta)u_R]$ into the expressions for A^C and P^C gives the results stated in (2) and (3). □

Lemma 3.1 shows that the manufacturer's national advertising rate is constant and positive. (The advertising rate being constant is a non-essential artifact of the model structure.) The retailer's local promotional rate is represented as a feedback (or Markovian) strategy such that promotion efforts are positively affected by the brand image G. Thus, the retailer promotes more intensively a brand with a strong image than one with a weak image. Promotional efforts are, however, increasing at a decreasing rate in G, which means that there is a saturation effect[1]. The value function $V^C(G)$ represents the optimal cooperative payoff-to-go when one starts out at time t in state G. In the absence of

[1] It could be interesting to test the result empirically, i.e., to check if retailers do indeed promote more the products enjoying a high goodwill than otherwise. A possible explanation of this would be that the higher the goodwill, the higher the sales of the product and also the sales of other products via the increase of the number of customers of the store attracted by the promoted product.

any promotion one would have $V^C(G) = \theta/\rho$, that is, the present value of a perpetual constant stream of baseline sales revenue.

Notice that neither A^C, P^C, nor V^C depends on time. The reason is that we have confined our interest to stationary strategies, that is, strategies which depend on state G only (i.e., not on time). Restricting one's interest to stationary strategies is a standard working hypothesis in autonomous dynamic optimization problems with infinite time horizon. (A problem is autonomous if the instantaneous payoffs and the right-hand side of the state dynamics do not depend explicitly on time.)

Remark 3.1. By (4) the value function is linearly increasing in G. To verify that the value remains bounded for $t \longrightarrow \infty$, insert A^C into the state dynamics (1). Solving the differential equation provides the optimal state trajectory

$$G^C(t) = G_0 \exp\{-\delta t\} + \frac{A^C}{\delta}\left[1 - \exp\{-\delta t\}\right],$$

which shows that $G^C(t)$, and hence $V^C(G(t))$, converges for $t \longrightarrow \infty$. Convergence of $G^C(t)$ is intuitive, in view of (1). Since advertising effort is constant, the decay term $-\delta G$ guarantees that $G^C(t)$ does not diverge.

In order to implement the joint maximization outcome identified in Lemma 3.1, the manufacturer designs an incentive strategy $D^C(P)$. The assumption is that at time $t = 0$ the manufacturer announces her advertising strategy A^C and the incentive $D^C(P)$. The manufacturer commits to use these strategies throughout the game.

The incentive is supposed to be a promotion allowance where the manufacturer pays $D^C(P)$ *per unit of* P. (Alternatively, the manufacturer might pay the retailer an allowance per unit of sales, or a fraction of the retailer's promotion cost.) The promotion incentive strategy then is

$$D^C(P)(t) = w^C P(t), \tag{6}$$

where w^C ("the incentive coefficient") is a positive constant, to be determined by the manufacturer at the start of the game. Note that the manufacturer offers promotional support only when the retailer makes some promotional effort of her own. Also note that the incentive strategy is fully determined by the choice of w^C. Since A^C and w^C both are constants, fixed at the start of the game, the manufacturer's strategy commitment amounts to keeping the values A^C and w^C fixed throughout the course of the game.

Remark 3.2. The choice of the functional form of the incentive is assumed here. In general, given the retailer's best reply, the manufacturer must solve a (nonstandard) optimal control problem that involves the (unknown) strategy $D^C(P)$. This can be any function, and it is not clear how to obtain an *optimal* incentive strategy. The simplest (and the standard) way to resolve this difficulty is to restrict the family of functions from which the manufacturer can choose, for example, to linear strategies (as was done in (6)).

The following proposition characterizes the solution of the incentive problem when the manufacturer wishes to implement the cooperative solution.

Proposition 3.1. *The incentive coefficient* $w^C = \pi u_R/2$ *leads to an implementation of the channel joint maximization solution. The incentive strategy depends on the state and is given by* $D^C(G) = \pi\beta\sqrt{G}/2$.

Proof. Given the manufacturer's commitment, the retailer wishes to design an equilibrium strategy $P(G)$. For this purpose she solves an optimal control problem, having the HJB equation

$$\rho V_R(G) = \max_{P\geq 0}\left\{(1-\pi)\left[\theta+\beta P\sqrt{G}\right]-\frac{u_R P^2}{2}+w^C P^2+V_R'(G)\left[A^C-\delta G\right]\right\}.$$

Maximizing the right-hand side of the above HJB equation with respect to P leads to

$$P(G) = \frac{\beta(1-\pi)\sqrt{G}}{u_R - 2w^C}. \tag{7}$$

To implement the desired promotion rate $P^C(G)$, given by (3), the manufacturer needs to select the incentive coefficient w^C in (6) such that $P^C(G)$ equals $P(G)$ in (7). Thus w^C must satisfy

$$\frac{\beta\sqrt{G}}{u_R} = \frac{\beta(1-\pi)\sqrt{G}}{u_R - 2w^C},$$

which leads to

$$w^C = \frac{\pi u_R}{2} > 0. \tag{8}$$

Finally, substituting w^C into (6) yields the incentive strategies

$$D^C(G) = w^C P^C(G) = \frac{\pi\beta\sqrt{G}}{2}. \tag{9}$$

□

The retailer promotes the brand (i.e., $P(G) > 0$) if and only if $u_R - 2w^C > 0$ which is obviously satisfied by the manufacturer's optimal choice of the coefficient w^C. Note that the inequality $u_R - 2w^C > 0$ is equivalent to $u_R P^2/2 > w^C P^2$. Hence, the manufacturer's total support ($w^C P^2$) is less than the full cost of promotion. From (9) we see that the promotion allowance D^C is always positive. Hence, in the joint maximization scenario the manufacturer always supports the retailer's promotion efforts. To interpret the incentive strategy, note that

$$D^C(G) = \frac{\pi\beta\sqrt{G}}{2} = \frac{\pi}{2}\frac{\partial S}{\partial P},$$

which tells that the per unit of promotion support provided by the manufacturer to the retailer is equal to half of the manufacturers's share in revenue times the marginal sales revenue with respect to promotion. This is a kind of "fair" sharing rule of marginal revenue due to promotion. To establish a relationship between total promotional cost and total amount of support given by the manufacturer to the retailer, recall that

$$C_R\left(P(G)\right) = \frac{u_R P(G)^2}{2} = \frac{\beta^2 G}{2u_R}$$
$$TS^C(G) = D^C(G)P^C(G) = \pi\frac{\beta^2 G}{2u_R},$$

where $TS^C(G)$ denotes the total support provided by the manufacturer in this scenario. It is easy to see that

$$TS^C(G) = \pi C_R\left(P(G)\right),$$

showing that the manufacturer applies the same agreed-upon splitting rules of the revenues to the promotional cost (recall that π is the share of the manufacturer in the total revenue). Hence, the higher her share in revenues, the higher her support. Further, the relationship between the total support and the goodwill is linear increasing. It is of interest to note that the elasticity of the total support with respect to the goodwill is precisely equal to one. Indeed, this elasticity, denoted ε, is given by

$$\varepsilon = \frac{\partial TS^C(G)}{\partial G}\frac{G}{TS^C(G)} = \frac{\pi\beta^2}{2u_R}\frac{2u_R G}{\pi\beta^2 G} = 1.$$

3.2 Manufacturer Profit Maximization

We now consider the scenario where the manufacturer wishes the retailer to implement the promotion rate that follows from the maximization of the manufacturer's profit. The manufacturer announces and commits to a national advertising rate A^I and an incentive strategy $D^I(P) = w^I P$. The superscript "I" refers to "individual." The manufacturer then needs to determine an advertising rate A^I and an incentive coefficient w^I such that her individual payoff J_M is maximized, subject to (1). This problem admits the following solution.

Proposition 3.2. *The equilibrium strategies and manufacturer's value functions are given by*

$$P^I(G) = \begin{Bmatrix} \frac{\beta(1+\pi)\sqrt{G}}{2u_R} \\ \frac{\beta(1-\pi)\sqrt{G}}{u_R} \end{Bmatrix} \text{ for } \begin{Bmatrix} \pi > \frac{1}{3} \\ \pi \leq \frac{1}{3} \end{Bmatrix} \tag{10}$$

$$A^I = \begin{Bmatrix} \frac{\beta^2(1+\pi)^2}{8u_M u_R(\rho+\delta)} \\ \frac{\beta^2\pi(1-\pi)}{u_M u_R(\rho+\delta)} \end{Bmatrix} \text{ for } \begin{Bmatrix} w^I > 0 \\ w^I = 0 \end{Bmatrix} \tag{11}$$

$$w^I = \begin{Bmatrix} \frac{(3\pi-1)u_R}{2(1+\pi)} \\ 0 \end{Bmatrix} \text{ for } \begin{Bmatrix} \pi > \frac{1}{3} \\ \pi \leq \frac{1}{3} \end{Bmatrix} \tag{12}$$

$$V_M^I(G) = \frac{\beta^2 (1+\pi)^2 G}{8(\rho+\delta) u_R} + \frac{\pi\theta}{\rho} + \frac{\beta^4 (1+\pi)^4}{128 u_M u_R^2 \rho (\rho+\delta)^2} \text{ for } \pi > 1/3 \tag{13}$$

$$V_M^I(G) = \frac{\beta^2 \pi (1-\pi) G}{(\rho+\delta) u_R} + \frac{\pi\theta}{\rho} + \frac{\beta^4 \pi^2 (1-\pi)^2}{2 u_M u_R^2 \rho (\rho+\delta)^2} \text{ for } \pi \leq 1/3.$$

Proof. The derivation follows the same approach as the one used in the proof of Proposition 3.1. The retailer's best response is derived from her HJB equation

$$\rho V_R^I = \max_{P \geq 0} \left[(1-\pi)S - \frac{u_R P^2}{2} + w^I P^2 + \frac{dV_R^I}{dG}(A - \delta G) \right].$$

Indeed, performing the maximization on the right-hand side yields

$$(1-\pi)\beta\sqrt{G} + (2w^I - u_R)P = 0 \text{ if } P > 0$$

from which the best reply retailer's strategy follows:

$$P(G) = \begin{Bmatrix} \frac{\beta(1-\pi)\sqrt{G}}{u_R - 2w^I} \\ \frac{\beta(1-\pi)\sqrt{G}}{u_R} \end{Bmatrix} \text{ for } \begin{Bmatrix} w^I > 0 \\ w^I = 0 \end{Bmatrix}. \tag{14}$$

The manufacturer's HJB equation is

$$\rho V_M^I = \max_{A^I \geq 0, w^I \geq 0} \left\{ \pi\theta + \frac{\pi\beta^2 (1-\pi) G}{u_R - 2w^I} - \frac{u_M A^2}{2} - \frac{w^I \beta^2 (1-\pi)^2 G}{(u_R - 2w^I)^2} + \frac{dV_M^I}{dG}(A - \delta G) \right\}.$$

Maximization of the term in braces yields

$$A^I = \frac{1}{u_M} \frac{dV_M^I}{dG}, \tag{15}$$

$$w^I = \frac{(3\pi - 1)u_R}{2(1+\pi)}, \tag{16}$$

which shows that A^I is positive whenever the shadow price dV_M^I/dG is positive. Furthermore, w^I is positive whenever $\pi > 1/3$. If $\pi \leq 1/3$ we have $w^I = 0$. From (14) and (16) we get the result for w^I. Inserting A^C and w^I on the right-hand side of the manufacturer's HJB equation provides

$$\rho V_M^I = \pi\theta - \delta G \frac{dV_M^I}{dG} + \frac{1}{2u_M}\left(\frac{dV_M^I}{dG}\right)^2 + \frac{\beta^2(1+\pi)^2 G}{8u_R} \text{ for } \pi > 1/3$$
$$\rho V_M^I = \pi\theta - \delta G \frac{dV_M^I}{dG} + \frac{1}{2u_M}\left(\frac{dV_M^I}{dG}\right)^2 + \frac{\beta^2(1-\pi)\pi G}{u_R} \text{ for } \pi \leq 1/3.$$

It is straightforward to verify that this differential equation admits the solution $V_M^I(G)$ stated in (13). A^C follows from (15) and (13). □

The proposition shows that the retailer's promotion is always positive and depends on the goodwill. The equilibrium advertising rate is constant and positive. It shows also that the optimal value of the incentive coefficient w^I is positive only if the manufacturer's share in sales revenue is greater than one-third.

In the event of positive support, the total cost and support are given by

$$C_R\left(P(G)\right) = \frac{u_R P^I(G)^2}{2} = \frac{\beta^2\left(1+\pi\right)^2 G}{8u_R}$$
$$TS^I(G) = D^C(G)P^C(G) = \frac{\beta^2\left(1+\pi\right)(3\pi-1)G}{8u_R},$$

where $TS^I(G)$ denotes the total support provided by the manufacturer in this scenario. It is easy to see that

$$\frac{TS^I(G)}{C_R\left(P(G)\right)} = \frac{(3\pi-1)}{(1+\pi)},$$

showing that the proportion of the promotional cost paid for by the manufacturer is an increasing function in her share in the revenues π. Although in this scenario the total support does not obey a simple rule as in the previous case, the previous elasticity result still holds. Indeed, this elasticity, denoted ξ, is given by

$$\xi = \frac{\partial TS^I(G)}{\partial G}\frac{G}{TS^C(G)} = \frac{\beta^2\left(1+\pi\right)(3\pi-1)}{8u_R}\frac{8u_R G}{\beta^2\left(1+\pi\right)(3\pi-1)G} = 1.$$

4 Comparing Strategies and Outcomes

In this section we compare the promotion and advertising strategies as well as incentive coefficients identified in Sections 3.1 (joint maximization) and 3.2 (individual maximization). Furthermore, we compare manufacturer's profits in the two scenarios. We have the following results.

Proposition 4.1. *Promotion and advertising strategies obtained in the two scenarios are related as follows:*
(a) $A^C > A^I, \forall w^I$ *and* A^I *is largest when* w^I *is positive.*
(b) $P^C(G) > P^I(G)$ *for all feasible* G *and irrespective of* w^I.
(c) $w^C > w^I$.

Proof. The first part in (a) is straightforward from (11) and (2) and the second part from (11). Comparing (10) to (3) leads to (b). Item (c) is straightforward from (12) and (8). □

The proposition shows that the manufacturer advertises more in the joint maximization solution than under individual maximization. Interestingly, this result does not depend on whether or not the manufacturer supports the retailer's promotion in the individual maximization case. Further, when the manufacturer implements the individual maximization outcome she advertises at the highest rate when she supports the retailer's promotion. Similarly, the retailer promotes more in the joint optimization case than under individual maximization; this result does not depend on whether or not the manufacturer supports the retailer's promotion in the individual maximization case.

The result in (c) shows that in the joint maximization case the manufacturer offers a higher promotion support rate to the retailer than in the individual maximization case.

Assuming that higher advertising and promotion levels mean better information to the consumer, the latter would prefer a manufacturer who optimizes total channel profit to one focusing on her own interest. However, what is important for the manufacturer is her own profit; hence, choosing between the two options would simply result from a comparison of the profits under the two regimes. Recall that in the joint maximization solution, the sharing parameter π vanishes, and hence comparing the two profits requires one first to make an assumption on how the joint profits would be shared between the two players. There are many possible ways in which to divide joint profits. One option which is rather intuitively appealing is to apply the same sharing rule to the joint profit as the one used for the sales revenue. Then the manufacturer receives the profit share $\pi V^C(G_0)$ (where $V^C(G_0)$ follows from (4)).

First we suppose that $\pi > 1/3$, i.e., the manufacturer will offer positive support in the individual maximization case. Using the value functions $V^C(G)$ and $V_M^I(G)$ from (4) and (13), respectively, yields

$$\pi V^C(G_0) - V_M^I(G_0) > 0 \Longleftrightarrow G_0 < \frac{\beta^2(\pi^3 + 5\pi^2 + 11\pi - 1)}{16 u_M u_R \rho(\rho+\delta)(1-\pi)} \triangleq f_1(\pi). \quad (17)$$

Noting that $\pi > 1/3$ implies that $\pi^3 + 5\pi^2 + 11\pi - 1 > 0$ which shows that $f_1(\pi)$ is positive. From (17) we see that it is in the manufacturer's best interest (in terms of maximal individual payoff) to implement the joint maximization outcome if the initial brand image is sufficiently weak ($G_0 < f_1(\pi)$). The intuition is that the manufacturer wishes to improve a (weak) brand image which most efficiently is done by implementing the joint maximization outcome, in which advertising and promotion rates are higher than in the individual maximization case.

On the other hand, if the brand image already is "strong" ($G_0 \geq f_1(\pi)$), the manufacturer has no incentive to implement the joint maximization outcome and hence acts in her own best interest (maximizing her individual profit).

Next suppose that $\pi \leq 1/3$. Then the manufacturer offers no support in the individual maximization case. Use value functions $V^C(G)$ and $V_M^I(G)$ from (4) and (13), respectively, to obtain

$$\pi V^C\left(G_0\right) - V_M^I\left(G_0\right) = \frac{\pi\beta^2(2\pi-1)G_0}{2u_R\left(\rho+\delta\right)} + \frac{\beta^4\pi(1-4\pi(1-\pi)^2)}{8u_Mu_R\rho\left(\rho+\delta\right)^2}. \qquad (18)$$

From this we conclude that if $\pi \leq 1/3$, then

$$\pi V^C\left(G_0\right) - V_M^I\left(G_0\right) > 0 \Longleftrightarrow G_0 < \frac{\beta^2[1-4\pi(1-\pi)^2]}{4u_Mu_R\rho\left(\rho+\delta\right)\left(1-2\pi\right)} \triangleq f_2(\pi).$$

The result in (18) is qualitatively similar to the result in (17). Thus, the manufacturer implements the joint maximization outcome if the initial brand image is sufficiently weak ($G_0 < f_2(\pi)$).

5 Concluding remarks

This chapter has considered a two-member marketing channel in which the manufacturer assumes the role of a channel leader and offers the retailer an incentive in the form of a promotion allowance. The allowance is paid per unit of the retailer's promotion effort. Depending on the share of retail sales revenue that the manufacturer receives, and the initial brand image level, the manufacturer can implement joint as well as individual maximization outcomes. Such outcomes are enforceable since they are Nash equilibria.

The reader should be aware that our results depend critically on one specific assumption, viz., that the manufacturer's share of the retail sales revenue π is constant. The hypothesis here is that the transfer price is negotiated between channel members at the start of the game, and the manufacturer commits to maintain this price level throughout the game. (Similarly, but less restrictive in the context at hand, the consumer price has been assumed constant.) It may happen, however, that the manufacturer does not want to commit to a fixed transfer price; rather she wishes to use the transfer price as a strategic variable, the choice of which will influence the share π. This puts the manufacturer in an even more dominating position, since then she can, by choosing an appropriate transfer price, dictate the outcome (joint or individual maximization) that will be implemented.

The derivation of the analytical results has been simplified by the assumed functional forms of the dynamics, revenues, and costs. The methodology proposed would remain unaltered if one adopts other forms, but of course the results will be affected. As a final remark we wish to note that an important issue in incentive problems, which has not been addressed in this chapter, is the assignment of roles. In practice it may very well happen that the roles are reversed, making the retailer the channel leader who wishes to design an incentive to make the manufacturer implement a particular outcome. More interesting than just a change of roles, it may be worthwhile to investigate the impacts of making the assignment of roles *endogenous*. Then the channel members will decide (on the basis of their individual payoffs) whether or not they will accept a particular assignment of roles. Another research avenue here would be to abandon the concept of a channel leader and consider a "symmetric" channel in which no member has a first-mover advantage (cf. Jeuland and Shugan [3]). A first attempt in this direction has been made in the dynamic game analyzed in Jørgensen and Zaccour [5].

REFERENCES

[1] Bergen M. and John G., Understanding cooperative advertising participation rates in conventional channels, *Journal of Marketing Research,* 34, 3, 357–369, 1997.

[2] Chintagunta P.K. and Jain D., A dynamic model of channel member strategies for marketing expenditures, *Marketing Science,* 11, 168–188, 1992.

[3] Jeuland A.P. and Shugan S.M., Managing channel profits, *Marketing Science,* 2, 239–272, 1983.

[4] Jørgensen S. and Zaccour G., Equilibrium pricing and advertising strategies in a marketing channel, *Journal of Optimization Theory and Applications,* 102, 1, 111–125, 1999.

[5] Jørgensen S. and Zaccour G., Channel coordination over time: Incentive equilibria and credibility, *Journal of Economic Dynamics and Control,* 27, 5, 801–822, 2002.

[6] Jørgensen S., Sigué S.P. and Zaccour G., Dynamic cooperative advertising in a channel, *Journal of Retailing,* 76, 1, 71–92, 2000.

[7] Jørgensen S., Sigué S.P. and Zaccour G., Stackelberg leadership in a marketing channel, *International Game Theory Review,* 3, 1, 13–26, 2001.

[8] Jørgensen S., Taboubi S. and Zaccour G., Cooperative advertising in a marketing channel, *Journal of Optimization Theory and Applications*, 110, 1, 145–158, 2001.

[9] Moorthy K.S., Managing channel profits: comment, *Marketing Science,* 6, 375–379, 1988.

Farsighted Behavior Leads to Efficiency in Duopoly Markets*

Akihiro Suzuki
Department of Public Policy and Social Studies
Faculty of Literature and Social Sciences
Yamagata University, Japan
asuzuki@human.kj.yamagata-u.ac.jp

Shigeo Muto
Department of Social Engineering
Graduate School of Decision Science and Technology
Tokyo Institute of Technology, Japan
muto@soc.titech.ac.jp

Abstract

This chapter studies the farsighted behavior of firms in Cournot and Bertrand duopoly markets. The solution concepts used here are the von Neumann–Morgenstern stable set with indirect domination and the consistent set. Principal findings are: (1) in both Cournot and Bertrand duopoly markets, the largest consistent set consists of all possible outcomes, and (2) the stable set with indirect domination yields firms' joint profit maximization in a Cournot duopoly, and monopoly pricing by two firms in a Bertrand duopoly. In both duopoly markets, the stable set with indirect domination leads to efficient outcomes for the firms.

1 Introduction

This chapter studies what results from firms' farsighted behavior in duopoly markets.

It is well known that, in one-shot duopoly markets, Nash equilibrium outcomes are not efficient in general (from the standpoint of the firms). If duopoly competition is repeated without end, then the Folk theorem holds; there exist too many equilibrium outcomes, though efficient ones are included in them.

*The authors wish to thank an anonymous referee for valuable comments. The second author wishes to acknowledge the support from the Japan Society for the Promotion of Science through Grants-in-Aid for Scientific Research (#16310107, 2004).

A different analysis of a firm's behavior in a duopoly is given in Greenberg [3]. The "individual contingent threats situation" assumes that each player, facing a pair of strategies, declares against its rival, "If you stick to the current strategy, I will shift to a new strategy." Each firm can make such contingent threats realizing the rival can make contingent threats in turn. Firms can revise their strategies; no one is committed to anything. The stability notion that Greenberg used was the stable set originally defined by von Neumann and Morgenstern [10].[1] Muto and Okada [8,9] applied Greenberg's idea to Bertrand and Cournot duopoly markets. Though several interesting outcomes, such as "equal pricing" of two firms in a Bertrand duopoly, were obtained, puzzling results, e.g., strategy pairs in which neither firm gains positive profit, also appeared as stable outcomes.

Later Chwe [2] criticized the myopic behavior of players assumed in von Neumann and Morgenstern [10] and also in Greenberg [3]. He introduced players' farsightedness into the definition of the stable set.[2] That is, he assumed that a deviating player took into consideration not only a direct response by the rival but also a sequence of responses between two players that might ensue. He proposed two solutions: one was the stable set with a sequence of deviations, and the other was its modification, which he called a consistent set.[3] He claimed that his solution, the consistent set, was superior to the stable set by showing examples in which a consistent set (precisely, the largest consistent set) existed but no stable set existed.

The aim of this chapter is to study in detail the two solutions, the stable set with a sequence of deviations and the largest consistent set, in Cournot and Bertrand duopoly markets.

Principal findings are the following. (1) In both duopoly markets, the stable set with a sequence of deviations produces only efficient outcomes (from the viewpoint of the firms): firms set the (equal) price that maximizes their profits in the Bertrand duopoly and produce the amounts that maximize their joint profit in the Cournot duopoly. (2) The largest consistent set is very large, and contains most of the possible pairs of strategies in both the Bertrand and Cournot duopolies.

Thus the von Neumann–Morgenstern stability together with the firms' farsighted behavior produces only efficient outcomes even though the firms act independently.

The rest of the chapter is organized as follows. Section 2 describes the duopoly markets under discussion. Section 3 gives definitions of the stable set with a sequence of deviations and of the consistent set, together with definitions of

[1] Though the stable set has been considered a solution for cooperative games, von Neumann and Morgenstern [10] defined this concept in a more general setting.

[2] Harsanyi [4] gave a similar criticism of the myopic behavior; but his notion of farsightedness was slightly different.

[3] These solutions were defined not only in 2-person but also in n-person games.

other basic terms. Section 4 presents the main theorems, which describe stable sets and the largest consistent sets in Cournot and Bertrand duopoly markets. Sections 5 and 6 give proofs of the theorems. Section 7 ends the paper with remarks on relations between our analysis and equilibrium analyses by Bhaskar [1] and by Maskin and Tirole [5–7] in alternating move games.

2 Duopoly Markets and Nash Equilibria

We consider two types of duopoly markets: the Cournot quantity-setting duopoly and the Bertrand price-setting duopoly. To simplify the discussion, we will consider a simple duopoly model in which firms' cost functions and a market demand function are all linear. Similar results, however, hold in more general duopoly models.

There are two firms, 1, 2, each producing a homogeneous good with the same marginal cost $c > 0$. No fixed cost is assumed.

(1) Cournot duopoly: The firms' strategic variables are their production levels. Let x_1 and x_2 be the production levels of firms 1 and 2, respectively. The market price $p(x_1, x_2)$ for x_1 and x_2 is given by

$$p(x_1, x_2) = \max(a - (x_1 + x_2), 0),$$

where $a > c$. We restrict the domain of production of both firms to $0 \leq x_i \leq a - c$, $i = 1, 2$. This is reasonable since a firm would not overproduce to make negative profits. When x_1 and x_2 are produced, firm i's profit is given by

$$\pi_i(x_1, x_2) = (p(x_1, x_2) - c)x_i.$$

Thus the Cournot duopoly is formulated as the following strategic form game:

$$G^C = (N,\ \{X_i\}_{i=1,2},\ \{\pi_i\}_{i=1,2}),$$

where the set of players is $N = \{1, 2\}$, the players' strategy sets are $X_1 = X_2 = [0, a - c]$, a closed interval between 0 and $a - c$, and the players' payoff functions are the profit functions π_i, $i = 1, 2$. The product of strategy sets $X_1 \times X_2$ is denoted by X. The total profit of two firms is maximized when $x_1 + x_2 = (a - c)/2$.

(2) Bertrand duopoly: The firms' strategic variables are their price levels. Let

$$D(p) = \max(a - p, 0)$$

be the market demand at price p. Then the total profit at p is

$$\Pi(p) = (p - c)D(p).$$

We restrict the domain of price level p of both firms to $c \le p \le a$. This assumption is also reasonable since a firm would avoid negative profits. The total profit $\Pi(p)$ is maximized at $p = (a+c)/2$. The price is called the "monopoly price."

Let p_1 and p_2 be the prices of firms 1 and 2, respectively. We assume that if firms' prices are equal they equally share the total profit; and otherwise all sales go to a lower pricing firm. Thus firm i's profit is given by

$$\rho_i(p_i, p_j) = \begin{cases} \Pi(p_i) & \text{if } p_i < p_j \\ \Pi(p_i)/2 & \text{if } p_i = p_j \\ 0 & \text{if } p_i > p_j \end{cases} \quad i, j = 1, 2, \; i \neq j.$$

Thus the Bertrand duopoly is formulated as the strategic form game

$$G^B = (N, \; \{Y_i\}_{i=1,2}, \; \{\rho_i\}_{i=1,2}),$$

where $N = \{1, 2\}$, $Y_1 = Y_2 = [c, a]$, and ρ_i $(i = 1, 2)$ is i's profit function. The product of strategy sets $Y_1 \times Y_2$ is denoted by Y.

It is well known that the Nash equilibrium is uniquely given in either market: $x_1 = x_2 = (a-c)/3$ in the Cournot duopoly, and $p_1 = p_2 = c$ in the Bertrand duopoly.

3 Stable Sets and Consistent Sets

Let $G = (N, \; \{S_i\}_{i=1,2}, \; \{u_i\}_{i=1,2})$, $N = \{1, 2\}$, be a two-person strategic form game where S_i and u_i are i's strategy set and payoff function, respectively.

For any two strategy pairs $s = (s_1, s_2)$ and $t = (t_1, t_2) \in S = S_1 \times S_2$, we say *$s$ is induced from t via player 1*, denoted $t \rightarrow_1 s$, if $s_2 = t_2$; i.e., s is reached from t by player 1's unilateral move. Similarly, *s is induced from t via player 2*, denoted $t \rightarrow_2 s$, if $s_1 = t_1$. We say *s indirectly dominates t*, denoted s indom t, if there exist a sequence of strategy pairs $t = s^0, s^1, \ldots, s^{m-1}, s^m = s$ and a sequence of players $i^1, i^2, \ldots, i^m$ such that for all $j = 1, \ldots, m$, $s^{j-1} \rightarrow_{i^j} s^j$ and $u_{i^j}(s) > u_{i^j}(s^{j-1})$.[4] Hence s indirectly dominates t when there are a sequence of strategy pairs starting from t and ending at s and a corresponding sequence of deviating players such that in each deviation a deviating player is better off at the end point s. We sometimes say that *s indirectly dominates t starting with i^1*, denoted s indom_{i^1} t, to specify the player who first deviates from t. If $m = 1$, we say *s directly dominates t,* denoted s dom t. The direct domination is considered in von Neumann and Morgenstern [10] and Greenberg [3]. The indirect domination is borrowed from Chwe [2], which is slightly different from the definition of indirect domination by Harsanyi [4].

[4]We may assume $i^{j-1} \neq i^j$ for all $j = 1, \ldots, m$ since consecutive moves by the same player can be combined into one move.

By just replacing direct domination in the stable set of von Neumann and Morgenstern [10] by indirect domination, we obtain the stable set with indirect domination. A set $K \subseteq S$ is called a *stable set with respect to (w.r.t.) indirect domination (indom)* if (1) for any $s, t \in K$, neither s indom t nor t indom s, and (2) for any $t \in S - K$, there exists $s \in K$ such that s indom t. Properties (1) and (2) are called *internal and external stability*, respectively.

Chwe [2] defined a notion of consistent set by slightly altering the definition of internal and external stability. *A consistent set* is a set $L \subseteq S$ satisfying the following two properties. (1) For any $s \in L$, $u \in S$, and any i ($i = 1$ or 2) such that $s \rightarrow_i u$, there exists $t \in L$, $t = u$ or t indom u, such that $u_i(s) > u_i(t)$ holds. (2) For any $t \in S - L$, there exist $u \in S$ and i ($i = 1$ or 2) such that for all $s \in L$, $s = u$ or s indom u, $u_i(s) > u_i(t)$ holds. We call properties (1) and (2) *internal and external consistency,* respectively. The *largest consistent set* is a consistent set that contains all others.

It should be noted that in the definition of the consistent set if we replace "there exists" in (1) by "for all" and "for all" in (2) by "there exists" then we obtain the definition of the stable set w.r.t. indirect domination. Thus it can be shown that any stable set w.r.t. indirect domination is contained in the largest consistent set. See Chwe [2], Proposition 3. Moreover, the largest consistent set is always unique, but stable sets w.r.t. indirect domination are generally not unique.

We will use symbols K and L to denote the stable set w.r.t. indirect domination and the largest consistent set, respectively.

4 Main theorems

The following two theorems, Theorems 4.1 and 4.2, are for the Cournot duopoly.

Theorem 4.1. *Take any strategy pair* $(x_1, x_2) \in X$ *such that* $x_1 + x_2 = (a - c)/2$*. Then the singleton set* $\{(x_1, x_2)\}$ *is a stable set. Furthermore, every stable set is of the form* $\{(x_1, x_2)\}$, $x_1 + x_2 = (a - c)/2$, $x_1, x_2 \geq 0$.

As mentioned before, any strategy pair (x_1, x_2) with $x_1 + x_2 = (a - c)/2$ and $x_1, x_2 \geq 0$ maximizes two firms' joint profit. Therefore we would claim that the stable set together with the firms' farsighted behavior produces joint profit maximization even if the firms act independently.

The next theorem shows that the largest consistent set is too large, i.e., it consists of all strategy pairs that give nonnegative profits to both firms.

Theorem 4.2. *Let* $L = \{(x_1, x_2) \in X \ : \ x_1 + x_2 \leq a - c\}$*. Then* L *is the largest consistent set.*

As for the Bertrand duopoly, the following two theorems hold. Theorem 4.3 claims that the monopoly price pair is itself a stable set and that no other stable set exists. Therefore the stable set together with the firms' farsighted behavior also attains efficiency (from the standpoint of the firms) in the Bertrand duopoly.

Theorem 4.3. *Let $p^* = (p_1^*, p_2^*)$ be the pair of monopoly prices, i.e., $p_1^* = p_2^* = (a+c)/2$. Then the singleton set $\{p^*\}$ is a unique stable set.*

The next theorem shows that the largest consistent set is the whole set of strategy pairs. Thus the largest consistent set can give a sharp prediction of stable outcomes in neither duopoly market.

Theorem 4.4. *$Y = Y_1 \times Y_2$ is the largest consistent set.*

5 Proofs of Theorems 4.1 and 4.2—Cournot Duopoly

Before proving the theorems, we present a simple lemma on the firms' profits. Its proof is omitted since it is straightforward.

Lemma 5.1. *Let*

$$\begin{aligned}
A &= \{(x_1, x_2) \in X \mid x_1 + x_2 < a - c, x_1 > 0, x_2 > 0\}, \\
B &= \{(x_1, x_2) \in X \mid x_1 = 0, 0 < x_2 < a - c\}, \\
C &= \{(x_1, x_2) \in X \mid 0 < x_1 < a - c, x_2 = 0\}, \\
D &= \{(x_1, x_2) \in X \mid x_1 = x_2 = 0\}, \\
E &= \{(x_1, x_2) \in X \mid x_1 + x_2 = a - c\} \quad \textit{and} \\
F &= \{(x_1, x_2) \in X \mid x_1 + x_2 > a - c\}.
\end{aligned}$$

See Figure 1. Then (1) $\pi_1(x_1, x_2) > 0$, $\pi_2(x_1, x_2) > 0$ if and only if $(x_1, x_2) \in A$, (2) $\pi_1(x_1, x_2) = 0$, $\pi_2(x_1, x_2) > 0$ if and only if $(x_1, x_2) \in B$, (3) $\pi_1(x_1, x_2) > 0$, $\pi_2(x_1, x_2) = 0$ if and only if $(x_1, x_2) \in C$, (4) $\pi_1(x_1, x_2) = 0$, $\pi_2(x_1, x_2) = 0$ if and only if $(x_1, x_2) \in D \cup E$, (5) $\pi_1(x_1, x_2) < 0$, $\pi_2(x_1, x_2) < 0$ if and only if $(x_1, x_2) \in F$.

Proof of Theorem 4.1. Take a strategy pair $x = (x_1, x_2) \in X$ such that $x_1 + x_2 = (a-c)/2$. The internal stability of $\{x\}$ is clear; so we will show only the external stability.

Take any strategy combination $y = (y_1, y_2) \in X$, $y \neq x = (x_1, x_2)$. Then at least one firm's profit in y is lower than its profit in x. Without loss of generality, we suppose $\pi_1(x) > \pi_1(y)$. First consider the case of $y_2 = x_2$. In this case x directly dominates y since $\pi_1(x) > \pi_1(y)$. Suppose $y_2 \neq x_2$. Pick a sequence of strategy pairs $y^0 = y = (y_1, y_2)$, $y^1 = (a-c, y_2)$, $y^2 = (a-c, x_2)$, $y^3 = x = (x_1, x_2)$ and a sequence of firms $i^1 = 1$, $i^2 = 2$, $i^3 = 1$. See Figure 2. Then for all $j = 1, 2, 3$, $y^{j-1} \to_{i^j} y^j$ and $\pi_{i^j}(y^3) > \pi_{i^j}(y^{j-1}), j = 1, 2, 3$. In

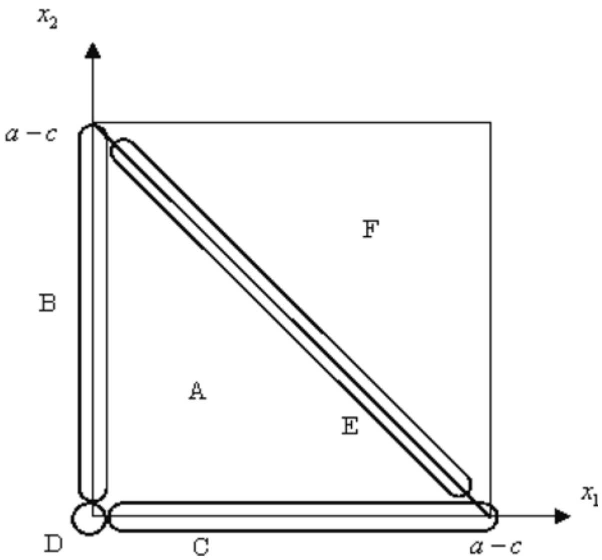


Figure 1: Regions A, B, C, D, E, F in Lemma 5.1.

fact, since $\pi_1(x) > \pi_1(y)$, $\pi_1(y^3) > \pi_1(y^0)$ is trivial; $\pi_1(y^3) > \pi_1(y^2)$ holds since $\pi_1(y^3) \geq 0$ (= holds when $x_1 = 0$ and thus $x_2 = (a-c)/2$), $\pi_1(y^2) \leq 0$ (= holds when $y_2^2 = 0$), and $x_2 = y_2^2$; and $\pi_2(y^3) > \pi_2(y^1)$ holds since $\pi_2(y^3) \geq 0$ (= holds when $x_2 = 0$), $\pi_2(y^1) \leq 0$ (= holds when $y_2^1 = 0$), and $y_2^1 = y_2 \neq x_2$. Therefore x indirectly dominates y, and thus the external stability holds.

We next show that any stable set must be of the form $\{(x_1, x_2)\}$, $x_1 + x_2 = (a-c)/2$, $x_1, x_2 \geq 0$. Take any strategy pair $y = (y_1, y_2) \in X$ that is not of this form. Thus $y_1 + y_2 \neq (a-c)/2$. We will show that y is not contained in any stable set. Then any stable set must contain a pair (x_1, x_2), $x_1 + x_2 = (a-c)/2$, $x_1, x_2 \geq 0$. From the proof of the external stability above, each such pair dominates all other strategy pairs, and thus the latter half of the theorem follows.

Suppose y is in a stable set. Call this stable set K. Since two firms' joint profit is maximized on the line segment $x_1 + x_2 = (a-c)/2$, $x_1, x_2 \geq 0$, we can take a strategy pair $z = (z_1, z_2)$ such that $z_1 + z_2 = (a-c)/2$, $z_1, z_2 \geq 0$ and $\pi_1(z) > \pi_1(y)$ and $\pi_2(z) > \pi_2(y)$. Hence if $z_1 = y_1$ or $z_2 = y_2$, then z directly dominates y. Thus we can assume $z_1 \neq y_1$ and $z_2 \neq y_2$. As for the strategy pair $z = (z_1, z_2)$, Lemma 5.1 shows the following: $\pi_1(z) \geq 0$, $\pi_2(z) \geq 0$, $\pi_1(z) = 0$ if and only if $z = (0, (a-c)/2)$, and $\pi_2(z) = 0$ if and only if $z = ((a-c)/2, 0)$.

We now show that z indirectly dominates y. Suppose first $y_2 = 0$. Lemma 5.1 shows the following: $\pi_1(y) \geq 0$, $\pi_2(y) = 0$, and $\pi_1(y) = 0$ if and only if $y_1 = 0, a-c$. Since $\pi_1(z) > \pi_1(y)$ and $\pi_2(z) > \pi_2(y)$, we have $\pi_1(z) > 0$ and $\pi_2(z) > 0$. Take a sequence of strategy pairs $y^0 = y = (y_1, y_2)$, $y^1 = (a-c, y_2)$, $y^2 =$

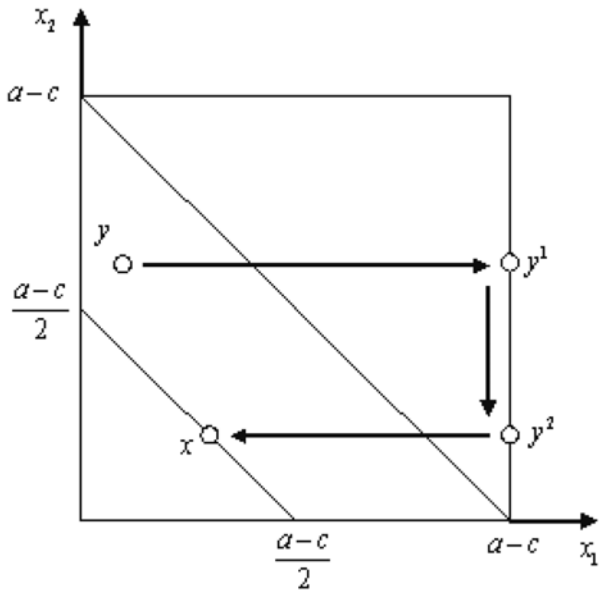


Figure 2: Illustration: x indom y.

$(a-c, z_2)$, $y^3 = z = (z_1, z_2)$, and a sequence of firms $i^1 = 1$, $i^2 = 2$, $i^3 = 1$. Then $\pi_1(z) > \pi_1(y)$, $\pi_2(z) > 0 = \pi_2(a-c, y_2)$, and $\pi_1(z) > 0 \geq \pi_1(a-c, z_2)$ since $\pi_1(a-c, y_2) = 0$ and $\pi_1(a-c, z_2) \leq 0$. Therefore z indirectly dominates y.

Suppose next $y_2 \neq 0$. Take the same sequence of strategy pairs as above. Then $\pi_1(z) > \pi_1(y)$ and $\pi_2(z) \geq 0 > \pi_2(a-c, y_2)$. The latter holds since $y_2 \neq 0$ implies $\pi_2(a-c, y_2) < 0$. Recall Lemma 5.1. Moreover, $\pi_1(z) > \pi_1(a-c, z_2)$ holds. In fact, Lemma 5.1 shows that if $z_2 = 0$, then $\pi_1(z) > 0 = \pi_1(a-c, z_2)$, and that if $z_2 \neq 0$, then $\pi_1(z) \geq 0 > \pi_1(a-c, z_2)$. Therefore z indirectly dominates y.

Since y is in the stable set K, by the internal stability of K, we must have $z \notin K$. Hence by the external stability of K, there must exist $v = (v_1, v_2) \in K$ that dominates z. Since v dominates z, $\pi_1(v) > \pi_1(z)$ or $\pi_2(v) > \pi_2(z)$. W.l.o.g. we assume $\pi_1(v) > \pi_1(z)$, and thus $\pi_1(v) > \pi_1(y)$ since $\pi_1(z) > \pi_1(y)$. Since $\pi_1(z) \geq 0$, $\pi_1(v) > 0$ holds. Then if $v_2 = y_2$, v directly dominates y, which contradicts the internal stability of K since $v, y \in K$. Thus we suppose $v_2 \neq y_2$ in the following.

Suppose first $y_2 = 0$. Then Lemma 5.1 shows that $\pi_1(y) \geq 0$, $\pi_2(y) = 0$ and that $\pi_1(y) = 0$ if and only if $y_1 = 0$ or $a-c$. Since $\pi_1(v) > \pi_1(y) \geq 0$, v must be in region $A \cup C$ in Lemma 5.1. Since $v_2 \neq y_2 = 0$, v must be in region A. Hence $\pi_2(v) > 0$. Take a sequence of strategy pairs $y^0 = y = (y_1, y_2)$, $y^1 = (a-c, y_2)$, $y^2 = (a-c, v_2)$, $y^3 = v = (v_1, v_2)$, and a sequence of firms $i^1 = 1$, $i^2 = 2$, $i^3 = 1$. Then $\pi_1(v) > \pi_1(y)$, $\pi_2(v) > 0 = \pi_2(a-c, y_2)$, and $\pi_1(v) > 0 > \pi_1(a-c, v_2)$.

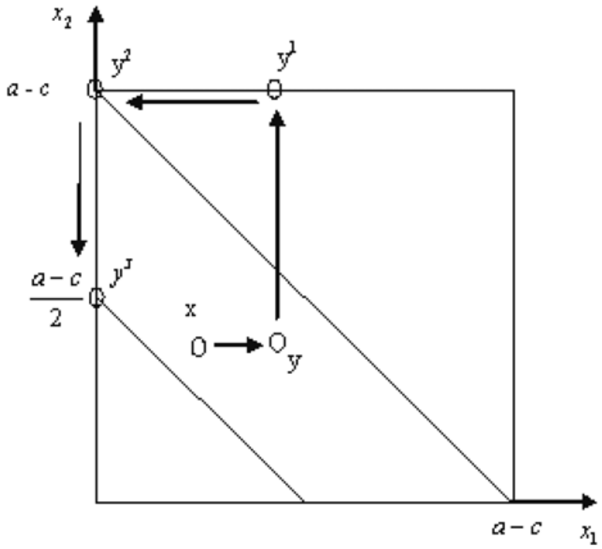


Figure 3: Illustration: y^3 indom y.

Note that $\pi_1(a-c, v_2) < 0$ since $v_2 \neq y_2 = 0$. Recall Lemma 5.1. Therefore v indirectly dominates y.

Suppose next $y_2 \neq 0$. Take the same sequence of strategy pairs as above. Then $\pi_1(v) > \pi_1(y)$, $\pi_2(v) \geq 0 > \pi_2(a-c, y_2)$, and $\pi_1(v) > 0 \geq \pi_1(a-c, v_2)$ hold. Therefore z indirectly dominates y. □

Proof of Theorem 4.2. We first show that L is a consistent set. To show the internal consistency of L, take any $x = (x_1, x_2) \in L$ and suppose firm 1 induces $y = (y_1, x_2)$. Note that $\pi_1(x) \geq 0$ and $\pi_2(x) \geq 0$. Recall Lemma 5.1. If $y_1 = 0$, then $y \in L$ and firm 1 is not better off since $\pi_1(y) = 0 \leq \pi_1(x)$. Suppose $y_1 \neq 0$. Take a sequence of strategy pairs $y^0 = y = (y_1, x_2)$, $y^1 = (y_1, a-c)$, $y^2 = (0, a-c)$, $y^3 = (0, (a-c)/2)$ and a sequence of firms $i^1 = 2$, $i^2 = 1$, $i^3 = 2$. See Figure 3. Then $\pi_1(y^3) = 0 > \pi_1(y^1)$. Note that $y_1 \neq 0$ implies that $\pi_1(y^1) = \pi_1(y_1, a-c) < 0$. Recall Lemma 5.1. Since $y^3 = (0, (a-c)/2)$ is on the joint profit maximization line and firm 2 gains the whole profit, $\pi_2(y^3) > \pi_2(y^0), \pi_2(y^2)$. Note that $y_1 \neq 0$ and thus $y^3 \neq y$. Therefore y^3 indirectly dominates y. Note that $y^3 \in L$. Since $\pi_1(y^3) = 0 \leq \pi_1(x)$, firm 1 is not better off in y^3 than in x. Thus the internal consistency of L holds.

To show the external consistency of L, take any $y = (y_1, y_2) \notin L$. Then $\pi_1(y) < 0, \pi_2(y) < 0$, and $y_1 > 0$. Recall Lemma 5.1. Suppose that firm 1 induces $z = (0, y_2)$. Then $z \in L$ and firm 1 is better off in z since $\pi_1(z) = 0 > \pi_1(y)$. Take any $x = (x_1, x_2) \in L$ that dominates z. Then since $x \in L$, $\pi_1(x) \geq 0 > \pi_1(y)$. Thus the external consistency of L holds.

We now show that L is the largest consistent set. Take any $y = (y_1, y_2) \notin L$ and suppose y is in a consistent set. Call this set L'. Note that $\pi_1(y) < 0$, $\pi_2(y) < 0$. Suppose that firm 1 induces $z = (0, y_2)$. Then since $\pi_1(z) = 0 > \pi_1(y)$ holds, by the internal consistency of L', there must exist $x \in L'$ such that x dominates z and firm 1 is not better off in x than in y. Suppose that firm 1 first deviates from z in the indirect domination. Then $\pi_1(x) > \pi_1(z) = 0$. Therefore firm 1 is better off in x than in y. Assume that firm 2 first deviates from z and induce $z' = (0, y_2')$. Then $\pi_1(z') = 0 > \pi_1(y)$; thus firm 1 is better off in z' than in y. If firm 1 deviates from z', it gains a positive profit at the end of the indirect domination sequence since $\pi_1(z) = 0$. Thus firm 1 is better off at the end than in y. Therefore the internal consistency of L' does not hold. □

6 Proofs of Theorems 4.3 and 4.4—Bertrand Duopoly

We first present a simple lemma on the firms' profits. Its proof is omitted since it is straightforward.

Lemma 6.1. *Let*

$$\begin{aligned} A &= \{(p_1, p_2) \in Y \mid c < p_1 = p_2 < a\}, \\ B &= \{(p_1, p_2) \in Y \mid c < p_2 < p_1 \le a\}, \\ C &= \{(p_1, p_2) \in Y \mid c < p_1 < p_2 \le a\} \\ D &= \{(p_1, p_2) \in Y \mid c = p_2 < p_1 \le a\}, \\ E &= \{(p_1, p_2) \in Y \mid c = p_1 < p_2 \le a\} \quad \textit{and} \\ F &= \{(p_1, p_2) \in Y \mid p_1 = p_2 = c \textit{ or } p_1 = p_2 = a\}. \end{aligned}$$

Then (1) $\rho_1(p_1, p_2) > 0, \rho_2(p_1, p_2) > 0$ if and only if $(p_1, p_2) \in A$, (2) $\rho_1(p_1, p_2) = 0$, $\rho_2(p_1, p_2) > 0$ if and only if $(p_1, p_2) \in B$, (3) $\rho_1(p_1, p_2) > 0$, $\rho_2(p_1, p_2) = 0$ if and only if $(p_1, p_2) \in C$, (4) $\rho_1(p_1, p_2) = 0$, $\rho_1(p_1, p_2) = 0$ if and only if $(p_1, p_2) \in D \cup E \cup F$. See Figure 4. Furthermore, the following hold. (5) In A, $\rho_1(p_1, p_2) = \rho_2(p_1, p_2)$ and $\rho_1(p_1, p_2)(= \rho_2(p_1, p_2))$ is maximized when $p_1 = p_2 = (a + c)/2$; it is monotone increasing in $p_1(= p_2)$ when $c < p_1 = p_2 < (a + c)/2$; and it is monotone decreasing in $p_1(= p_2)$ when $(a+c)/2 < p_1 = p_2 < a$. (6) In B, if $p_1 > (a+c)/2$, then $\rho_2(p_1, p_2)$ is maximized when $p_2 = (a+c)/2$, monotone increasing in p_2 when $p_2 < (a+c)/2$, and monotone decreasing in p_2 when $p_2 > (a + c)/2$. If $p_1 \le (a + c)/2$, then $\rho_2(p_1, p_2)$ is monotone increasing in p_2. (7) In C, if $p_2 > (a + c)/2$, then $\rho_1(p_1, p_2)$ is maximized when $p_1 = (a + c)/2$, monotone increasing in p_1 when $p_1 < (a + c)/2$, and monotone decreasing in p_1 when $p_1 > (a + c)/2$. If $p_2 \le (a + c)/2$, then $\rho_1(p_1, p_2)$ is monotone increasing in p_1.

Proof of Theorem 4.3. We will show the external stability of $\{p^*\}$ where $p^* = (p_1^*, p_2^*)$, $p_1^* = p_1^* = (a + c)/2$. Its internal stability is clear. Note that

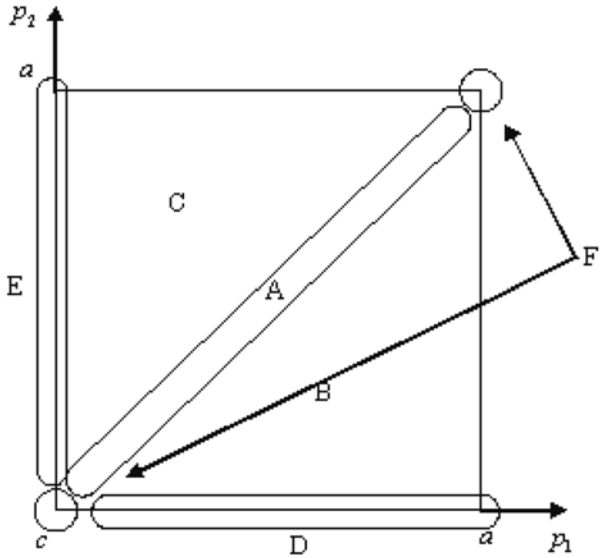


Figure 4: Regions A, B, C, D, E, F in Lemma 6.1.

$\rho_1(p^*) = \rho_2(p^*) > 0$ and that $\rho_1(p^*) = \rho_2(p^*) > \rho_1(p) = \rho_2(p)$ for all $p = (p_1, p_2)$ with $p_1 = p_2 \neq (a+c)/2$.

Take any strategy pair $q = (q_1, q_2) \in Y$, $q \neq p^*$. Then at least one firm's profit is lower than in p^*. W.l.o.g. we suppose $\rho_1(p^*) > \rho_1(q)$.

Assume first $q_1 \neq c$. Take a sequence of strategy pairs $p^0 = q = (q_1, q_2)$, $p^1 = (c, q_2)$, $p^2 = (c, p_2^*)$, $p^3 = (p_1^*, p_2^*)$ and a sequence of firms $i^1 = 1$, $i^2 = 2$, $i^3 = 1$. See Figure 5. Then for all $j = 1, 2, 3$, $p^{j-1} \rightarrow_{i^j} p^j$ and $\rho_{i^j}(p^3) > \rho_{i^j}(p^{j-1}), j = 1, 2, 3$. Therefore p^* indirectly dominates q; thus the external stability holds.

Assume next $q_1 = c$ and thus $\rho_1(q) = \rho_2(q) = 0$. Recall Lemma 6.1. If $q_2 = p_2^*$, then p^* dominates q. If $q_2 \neq p_2^*$, then take a sequence of strategy pairs $p^0 = q = (q_1, q_2) = (c, q_2)$, $p^1 = (c, p_2^*)$, $p^2 = p^* = (p_1^*, p_2^*)$ and a corresponding sequence of firms $i^1 = 2$, $i^2 = 1$. Then for j =1,2, $p^{j-1} \rightarrow_{i^j} p^j$ and $\rho_{i^j}(p^3) > \rho_{i^j}(p^{j-1}), j = 1, 2, 3$. Thus p^* indirectly dominates q.

We next show the uniqueness. It suffices to show that (p_1^*, p_2^*) is contained in any stable set. In fact, if (p_1^*, p_2^*) is contained in a stable set, the stable set must consist only of this point since (p_1^*, p_2^*) satisfies the external stability as shown above.

Suppose (p_1^*, p_2^*) is not contained in a stable set. Call this stable set K. Then by the external stability of K, there must exist $p = (p_1, p_2)$ in K which dominates (p_1^*, p_2^*). Note that $p_1 \neq p_2$ holds since if $p_1 = p_2$ then we must have $\rho_i(p) < \rho_i(p^*)$ for i =1,2; and thus indirect domination is impossible. Hence one firm gains zero profit in $p = (p_1, p_2)$.

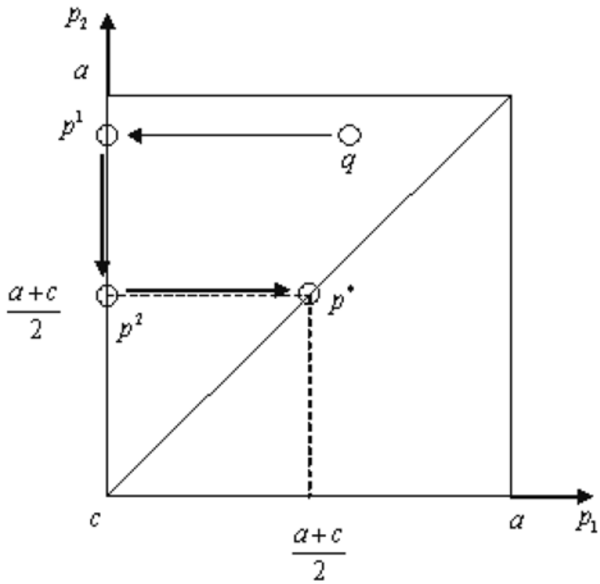


Figure 5: Illustration: p^* indom q.

W.l.o.g. suppose (p_1, p_2) indirectly dominates(p_1^*, p_2^*) starting with firm 1. This indirect domination never involves two firms since one firm gains zero profit in (p_1, p_2). Recall that zero is the minimum profit in the Bertrand duopoly game. See Lemma 6.1. Thus (p_1, p_2) directly dominates (p_1^*, p_2^*). Suppose (p_1, p_2) directly dominates (p_1^*, p_2^*) via firm 1; thus $p_2 = p_2^*$ and $\rho_1(p_1, p_2^*) > \rho_1(p_1^*, p_2^*)$. If $p_1 > p_1^*$, then $\rho_1(p_1, p_2^*) = 0 < \rho_1(p_1^*, p_2^*)$; thus we must have $p_1 < p_1^*$. We take p_1' such that $p_1 < p_1' < p_1^*$. See Figure 6. Then (p_1', p_2^*) directly dominates (p_1, p_2^*) via firm 1 since $p_1 < p_1' < p_1^*$ implies $\rho_1(p_1', p_2^*) > \rho_1(p_1, p_2^*)$. Since (p_1, p_2^*) is in K, (p_1', p_2^*) must not be in K by the internal stability. Then by the external stability, there must exist (p_1'', p_2'') in K, which dominates (p_1', p_2^*).

If this indirect domination involves more than one firm, then we must have $p_1'' = p_2''$; otherwise one firm's profit is zero and thus the indirect domination is impossible. Further, the indirect domination must start with firm 2 since $\rho_1(p_1', p_2^*) > \rho_1(p_1^*, p_2^*) > \rho_1(p_1'', p_2'')$. Thus $\rho_2(p_1'', p_2'') > \rho_2(p_1', p_2^*) = 0$.

Consider two cases: $p_1 = p_1''$ and $p_1 \neq p_1''$. Suppose first $p_1 = p_1''$. Then $(p_1'', p_2'')(= (p_1, p_1))$ directly dominates (p_1, p_2^*) via firm 2 since $\rho_2(p_1'', p_2'') > 0 = \rho_2(p_1, p_2^*)$. Suppose next $p_1 \neq p_1''$. Take a sequence of strategy pairs (p_1, p_2^*), (p_1, c), (p_1'', c), (p_1'', p_2''), and the corresponding sequence of firms $i^1 = 2$, $i^2 = 1$, $i^3 = 2$. See Figure 6. Then $\rho_2(p_1'', p_2'') > 0 = \rho_2(p_1'', c) = \rho_2(p_1, p_2^*)$ and $\rho_1(p_1'', p_2'') > 0 = \rho_1(p_1, c)$ hold. Recall Figure 4. Hence (p_1'', p_2'') dominates (p_1, p_2^*) starting with firm 2. Since (p_1'', p_2'') and (p_1, p_2^*) are both in K, this contradicts the internal stability of K.

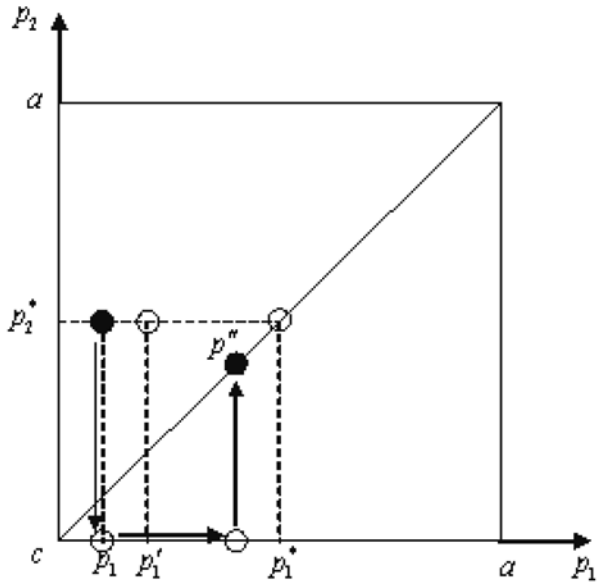


Figure 6: Points (p_1^*, p_2^*), (p_1, p_2^*), (p_1', p_2^*), (p_1'', p_2'').

Hence (p_1'', p_2'') directly dominates (p_1', p_2^*). If the domination is done via firm 1, then $\rho_1(p_1'', p_2'') > \rho_1(p_1', p_2^*) > \rho_1(p_1, p_2^*)$; and thus (p_1'', p_2'') directly dominates (p_1, p_2^*) via firm 1. This contradicts the internal stability of K.

Suppose that the domination is done via firm 2; thus $p_1'' = p_1'$. Note that (p_1', p_2'') must be on or below the main diagonal in Y. In fact, if (p_1', p_2'') is above the main diagonal and thus $p_1' < p_2''$, then $\rho_2(p_1', p_2'') = \rho_2(p_1', p_2^*) = 0$. Hence (p_1', p_2'') cannot indirectly dominate (p_1', p_2^*) via firm 2.

Take first the case where (p_1, p_2'') is below the main diagonal in the space Y, i.e., $p_1 > p_2''$. See Figure 7. Then (p_1, p_2'') directly dominates (p_1, p_2^*) via firm 2 since $\rho_2(p_1, p_2'') > 0 = \rho_2(p_1, p_2^*)$. Thus (p_1, p_2'') is not in K; and there must exist (p_1''', p_2''') in K which indirectly dominates (p_1, p_2''). This indirect domination must involve two firms. In fact, if (p_1''', p_2''') directly dominates (p_1, p_2'') via firm 1, (p_1''', p_2''') also directly dominates (p_1', p_2'') via firm 1 since $p_2''' = p_2''$ and $\rho_1(p_1, p_2'') = \rho_1(p_1', p_2'') = 0$. This contradicts the fact that (p_1''', p_2''') and (p_1', p_2'') are both in K. When (p_1''', p_2''') directly dominates (p_1, p_2'') via firm 2, (p_1''', p_2''') directly dominates (p_1, p_2^*), contradicting the fact that (p_1''', p_2''') and (p_1, p_2^*) are both in K. Therefore the indirect domination must involve two firms. Thus $\rho_1(p_1''', p_2''') = \rho_2(p_1''', p_2''') > 0$ and $p_1''' = p_2'''$.

Take a sequence of strategy pairs (p_1, p_2^*), (p_1, c), (p_1''', c), (p_1''', p_2'''), and the corresponding sequence of firms $i^1 = 2$, $i^2 = 1$, $i^3 = 2$. See Figure 7. Then (p_1''', p_2''') dominates (p_1, p_2^*) since $\rho_2(p_1''', p_2''') > 0 = \rho_2(p_1, p_2^*) = \rho_2(p_1''', c)$ and $\rho_1(p_1''', p_2''') > 0 = \rho_1(p_1, c)$.

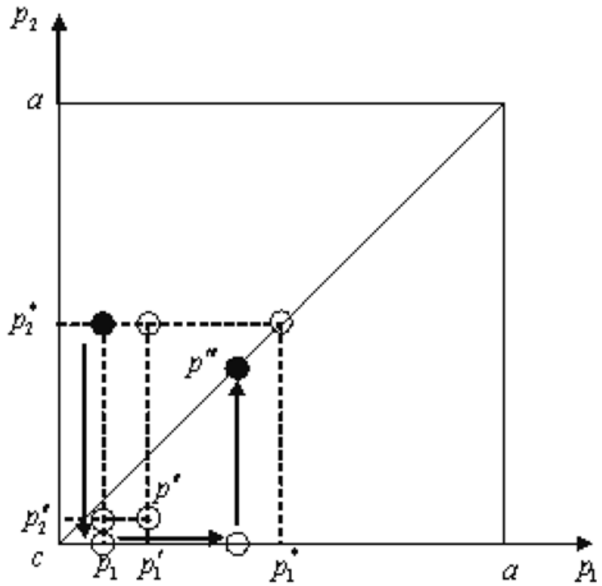


Figure 7: Illustration: Case (p_1, p_2'') is below the main diagonal.

When (p_1, p_2'') is on or above the main diagonal in Y, i.e., $p_1 \leq p_2''$, (p_1, p_2'') directly dominates (p_1', p_2'') via firm 1 since $\rho_1(p_1, p_2'') > 0 = \rho_1(p_1', p_2'')$. Thus (p_1, p_2'') is not in K; and there must exist (p_1''', p_2''') in K which indirectly dominates (p_1, p_2''). Then a proof similar to the above applies. Figure 8, similar to Figure 7, illustrates indirect domination, which leads to a contradiction. $\square$

Proof of Theorem 4.4. It suffices to show the internal consistency. Take any strategy combination $(p_1, p_2) \in Y$. W.l.o.g. assume $p_1 \leq p_2$.

First consider the case $p_1 = c$. Thus $\rho_1(p_1, p_2) = \rho_2(p_1, p_2) = 0$. Even if firm 2 moves, 2's profit remains zero since $p_1 = c$. Recall Lemma 6.1. Suppose that firm 1 moves to p_1', $c < p_1' < a$. If $p_2 = c$, then 1's profit remains zero. Thus suppose $p_2 > c$. If $p_1' > p_2$, then 1's profit is zero. Thus suppose $c < p_1' \leq p_2$. If $p_2 = a$, then suppose $c < p_1' < p_2$. Then $\rho_1(p_1', p_2) > 0$ and 1's profit increases. But then 2 moves to p_2' such that $c < p_2' < p_1$. See Figure 9.

Firm 2's profit increases and 1's profit goes down to zero. Thus for each of 1's moves, there exists 2's move, which decreases 1's profit down to zero.

Next suppose $c < p_1 < p_2$; then $\rho_1(p_1, p_2) > 0 = \rho_2(p_1, p_2)$. Suppose that first firm 1 moves to p_1'. If $p_1' > p_2$ or $p_1' = c$, then 1's profit decreases to zero. If $c < p_1' \leq p_2$, firm 2 then lowers his price to p_2', $c < p_2' < p_1'$, which gives a positive profit to firm 2 and makes firm 1's profit zero. Next suppose that firm 2 moves. If $p_2' > p_1$ or $p_2' = c$, then firm 2's profit remains zero. If $c < p_2' \leq p_1$,

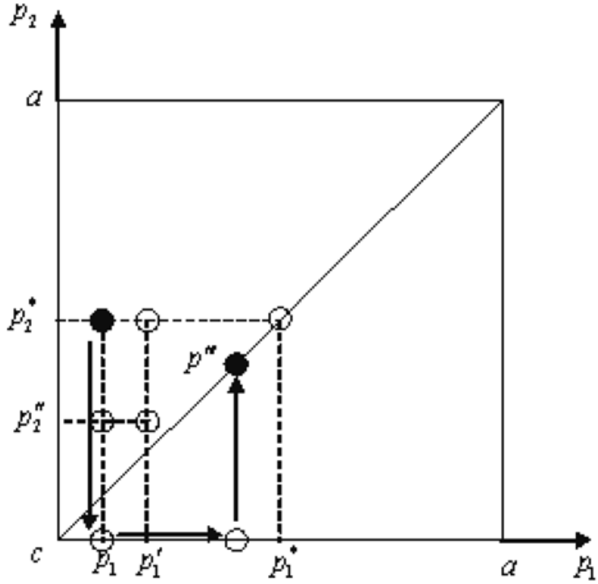


Figure 8: Illustration: Case (p_1, p_2'') is on or above the main diagonal.

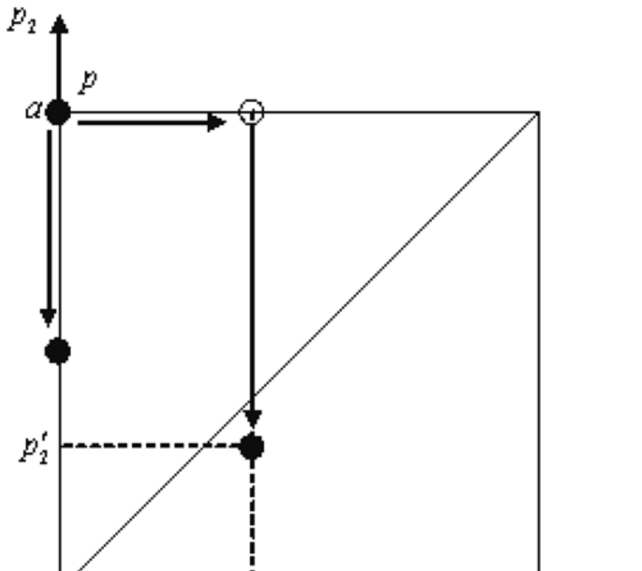


Figure 9: Case: $c < p_1' \leq p_2$.

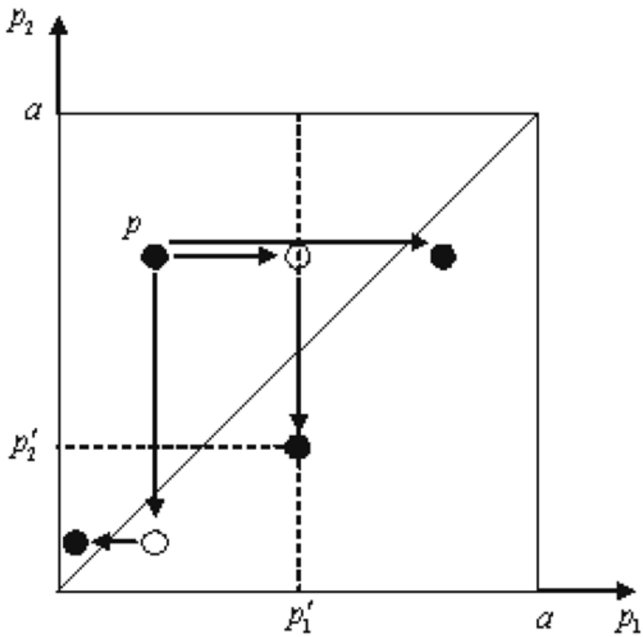


Figure 10: Case: $c < p_1' \leq p_2$.

then firm 1 moves to p_1' such that $c < p_1' < p_2'$. See Figure 10. This move gives a positive profit to 1 and makes 2's profit zero.

Finally consider the case $c < p_1 = p_2$. A proof similar to the above applies. A deviating firm becomes worse off or even if it temporarily becomes better, its profit goes down to zero by its rival's successive move. □

7 Concluding remarks

In this chapter we studied stable outcomes in duopoly markets when firms acted with farsightedness. The solution concepts that we used were the stable set and the (largest) consistent set. We showed that only efficient outcomes (from the standpoint of firms) were stable (in the sense of a stable set) when firms acted with farsightedness, even though they acted independently. Our analysis also revealed that the largest consistent set of Chwe was too large and gave no sharp prediction of stable outcomes.

The result obtained in a Bertrand duopoly reminds us of the result of Bhaskar [1], who considered an alternating-move preplay of two firms in which both firms' interests were only in a final outcome in which neither firm deviated. He showed that the Markov-perfect equilibrium with no dominating strategy in any subgame gave us a unique outcome, that is, the pair of monopoly prices.

It would be interesting to study more in detail relations between equilibrium analyses of alternating-move games such as those of Bhaskar [1] and Maskin and Tirole [5–7] and the stable set analysis. An extension to a general n-person case would also be interesting. As for the extension, Harsanyi's idea of interpreting a stable set in coalitional form games as an equilibrium in a certain bargaining game [4] would be helpful.

REFERENCES

[1] Bhaskar V., Quick responses in duopoly ensure monopoly pricing, *Economics Letters*, 29, 103–107, 1989.

[2] Chwe M.S.-Y., Farsighted coalitional stability, *Journal of Economic Theory*, 63, 299–325, 1994.

[3] Greenberg J., *The Theory of Social Situations: An Alternative Game Theoretic Approach*, Cambridge University Press, London, 1990.

[4] Harsanyi J., An equilibrium-point interpretation of stable sets and a proposed alternative definition, *Management Science*, 20, 1472–1495, 1974.

[5] Maskin E. and Tirole J., A theory of dynamic oligopoly I: Overview and quantity competition with large fixed costs, *Econometrica*, 56, 549–569, 1988a.

[6] Maskin E. and Tirole J., A theory of dynamic oligopoly II: Price competition, kinked demand curves, and edgeworth cycles, *Econometrica*, 56, 571–599, 1988b.

[7] Maskin E. and Tirole J., A theory of dynamic oligopoly III: Cournot competition, *European Economic Review*, 31, 947–969, 1987.

[8] Muto S. and Okada D., Von Neumann-Morgenstern stable sets in a price-setting duopoly, *Economy and Economics*, Faculty of Economics, Tokyo Metropolitan University, 81, 1–14, 1996.

[9] Muto S. and Okada D., Von Neumann-Morgenstern stable sets in Cournot competition, *Economy and Economics*, Faculty of Economics, Tokyo Metropolitan University, 85, 37–57, 1998.

[10] von Neumann J. and Morgenstern O., *The Theory of Games and Economic Behavior*, 3rd ed., Princeton University Press, Princeton, NJ, 1953.

A Stochastic Game Model of Tax Evasion

T.E.S. Raghavan
Department of Mathematics
Statistics and Computer Science
University of Illinois at Chicago, USA
ter@uic.edu

Abstract

The problem of tax evasion is modelled as a zero-sum two-person generalized stochastic game with incomplete information. This model incorporates the classical statistical classification procedures used in classifying a random observation from a mixed population. The model incorporates the secrecy of the tax office and lack of information about the past history of the taxpayer. With full information, the model is closer to certain structured classes of stochastic games that admit efficient algorithms for optimal solutions.

1 Historical Introduction

Tax evasion as a topic for theoretical investigation was first suggested by J.A. Mirrlees in a paper prepared for the International Economic Associations Workshop in Economic Theory, in Bergen, Norway in 1971 (Allingham and Sandmo [2]). Independently Allingham and Sandmo [2],[1] and Srinivasan [45] considered static models which were almost identical but still different in terms of tax function and taxpayer's aim. The following is their model: Suppose the income y of a taxpayer when reported results in a tax $T(y)$. Let λ be a proportion by which y is understated. Since the government can find it only when the return is audited, the government does not know y, but only the reported $(1-\lambda)\,y$. Let π be the chance for being audited (of course π can depend on y). Let $P(\lambda)$ be the penalty multiplier, i.e., $P(\lambda)\lambda y$ is the penalty on the undeclared income λy. Let us assume that the individual chooses λ that minimizes his expected income (or suitable expected utility function of income). If the taxpayer is risk averse and if his expected payment on unreported income is less than what he/she has to pay otherwise, he/she will declare less when $\pi(y) = \pi$. When the probability of audit increases, the optimal proportion λ^* by which income is understated decreases. If the audit chances are independent of the income level, then, richer income is underreported at greater λ's. However, this will not be the case when π depends on income levels. When the actual income varies, the fraction declared increases or decreases according to the relative risk aversion in the sense of Arrow [4] is an increasing or decreasing function of income. Allingham and Sandmo [2] also considered a dynamic model for a constant tax rate with t periods. They assumed that by audit, the government could recover all the dues up to that day from the remote past. They showed that the optimal strategy for the taxpayer would be to choose a period T and evade until period T and report fully after period T. This line of research presupposes that *while the government is ignorant of a taxpayer, the taxpayer is fully informed of their audit chances!* Based on a survey conducted in Belgium, Frank and Dekeyser-Meulders [17] calculated certain tax discrepancy coefficients. They found that wage earners and salaried persons gained the least by tax evasion and that certain types of evasions could not be caught even after an audit. Balbir Singh [6] observed in Srinivasan's model that with fixed chance π for audit, if $\pi < 1/3$, taxpayers could even evade income tax completely. Kolm [21] pointed out that their models never involved any auditing costs. Based on a survey, Monk [24] suggested that greater resources should be allocated to auditing higher income groups. Spicer and Lundstedt [44] pointed out that tax evasion was more than just gambling. A psychological survey conducted by Spicer and Lundstedt [44] (also see Spicer [42]) revealed the following phenomena.

(1) Evasion is less likely when sanctions against evasion are perceived to be severe.
(2) Evasion is less likely when probability of detection is perceived to be high.
(3) Evasion is more likely when a taxpayer perceives that his terms of trade with the government are inequitable compared to others.

Vogel, in a survey conducted in Sweden [47], observed that taxpayers were vulnerable for tax evasion when their aspirations were not matched by the government's services. They also observed that direct cash flow resulted in greater tax evasion. Cross and Shaw [9] corroborated the same view on many professionals who were self-employed. Allingham, by a simple model, pointed out [1] that progressive taxation need not be a solution for removing inequities.

The models by Reinganum and Wilde [34], [35] and Erard and Feinstein [13] were clearly game theoretic and allowed strategic behavior by the Internal Revenue Service (IRS) against taxpayers. Reinganum and Wilde [34] through a simple model showed that an audit cut-off policy would be more desirable as it would dominate any random audit policy. Erard and Feinstein [13] expanded on the model of Reinganum and Wilde [34] and showed that unlike the model of Allingham and Sandmo [2] where honest taxpayers had no influence on the rest of the population indulging in tax evasion strategies, in their extended [13] model, in equilibrium, honest taxpayers had indirect peer pressure on tax evaders. Mookherjee and Png [25] develop a model and find sufficient conditions for random audits to be optimal.

There is a small amount of recent empirical work on what determines tax compliance (see [8], [11]). By partitioning the set of all taxpayers into three distinct classes, called 1. *honest*, 2. *susceptible*, and 3. *evading* types, Davis, Hecht and Perkins [10] study the problem via an explicit law of motion and its solution. For example, they assume that the rate of change of the population of honest taxpayers with respect to time is a negative proportion of the product of honest and evasive taxpayers. Justification for this assumption of their model is based on the empirical observation by Vogel [47] and Spicer and Lundstedt [44] that even honest people can become evaders in the future when their colleagues are evaders. One can contrast this with the assertions of Erard and Feinstein [13].

While many of these models are static, tax evasion and tax compliance are dynamic phenomena. One of the earliest dynamic game models of tax evasion was initiated by Greenberg [19]. See also Landsberger and Meiljison [22]. Greenberg, who formulated the problem as a repeated game with absorbing states, imposed some strong assumptions on the law of motion in order to achieve an elegant characterization of the optimal strategies. These were all zero-sum models.

The model proposed here is a generalized zero-sum stochastic game but with incomplete information. For the case when past history and immediate payoffs and transitions are common knowledge, these games reduce to tractable classes admitting efficient algorithms for computing good strategies (see Parthasarathy and Raghavan [28], Filar and Vrieze [15], Raghavan and Syed [30],[31]). The model is capable of incorporating empirical evidences via the immediate payoffs and transition probabilities.

Tax agencies like the IRS will show greater interest in the game theoretic approach only when the suggested solutions are further refinements that are closer to their current audit procedures developed in cooperation with their electronic data processing (EDP) units. Even popular books by IRS agents and supervisors (see Murphy [27], Monk [24] and informative articles by tax agency directors (Pond [29], Smith [39]) agree on the power and usefulness of discriminant analysis. Our models here complement and refine the discriminant function approach. We will still need the valuable and ingenious techniques of conducting sample surveys as in Frank and Dekeyser [17], Monk [24], Strumpel [46] to gather information about psychological behavior patterns of taxpayers in forming immediate payoffs. In this context the psychological studies in simulating income tax evasion by Friedland, Maital, and Rutenberg [18] and Spicer and Becker [43] will be very useful.

2 Secrecy and Lack of Information

An essential feature of taxation is the secrecy behind auditing procedures implemented by the tax office and the lack of full information about any taxpayer and his possible tax evasion strategies. These aspects have not been effectively incorporated into the models of tax evasion considered thus far in public economics literature. Often, in order to characterize equilibrium strategies and optimal strategies in closed form, model builders tend to make drastic assumptions. Our approach to modelling tax evasion is certainly not to look for closed form solutions but to look for models that retain the notions of *incomplete information* and *secrecy of actions* intrinsic to the tax evasion problem. At the same time, these models are quite close to existing stochastic game models where efficient solution techniques have already been developed. The notions of capturing secrecy and asymmetry in information have both been part of substantial research in the area of game theory known as *games with incomplete information*. Existence theorems are much harder to obtain in many such games with incomplete information. Here we propose a dynamic game theoretic approach to the study of the tax compliance problem that incorporates the dual secrecy inherent in the problems of tax evasion and auditing. The problem is viewed as a multistage game between the IRS (player I) and a taxpayer in a socioeconomic group (player II). The taxpayer adopts, either by choice or by ignorance of tax laws, a strategy to evade taxes on certain selected taxable items. Based on the particular socioeconomic group of the person, the IRS has a prior perception about the taxpayer with respect to his methods and modes. This perception is modified from year to year based on the tax returns and the dictates of the discriminant function and the norms of the IRS. This is modelled as a stochastic game with transition laws and states unknown to the taxpayer (player II) but known to the IRS (player I).

3 Detecting Tax Evasion via Discriminant Analysis

Fisher, in his seminal paper on taxonomic problems [16], suggested an ingenious procedure to classify any observation drawn randomly from a mixture of populations into one of them, based on the densities of the sub-populations. For many practical applications see [26], [3], [32], [33]. The procedure is easily adaptable to problems involving classifications in many other areas including bankers lending credit facilities for small businesses, taxation, and credit card approvals. Intuitively, we can describe this procedure for tax evasion problems as follows.

Although the tax paying population is quite heterogeneous, people in each professional group are relatively homogeneous. They tend to associate with people in the same professional group and inherit similar socioeconomic patterns of life. Thus, the population can be made more homogeneous by stratifying according to profession. Having stratified the population into sub-populations, such as executives, doctors, lawyers, salesmen etc., the next problem is to further divide each sub-population into two types, namely those filing legally correct and honest returns and those filing legally incorrect or manipulated tax returns. Apparently, in the early 1940s nearly 25% of the tax returns belonged to the second type [27]. While deductions in income tax returns accounted for less than 12% in 1947, a decade later the same deductions were almost 15% of the reported gross income [27]. Apparently, tax loopholes and manipulations were used in the process.

The classification of all members of a professionally homogeneous group into the above two distinct types is much more complex. This could only be achieved when persons in the profession were targeted earlier with a foolproof audit. As a first step one needs norms for auditing, so that tax items violating these norms conspicuously can be considered as candidates for auditing. Only expert tax inspectors can be relied upon to come to grips with this initial data classification problem.

Assume that there is data available from the past for this classification. Our hunch is that the IRS will know from past data the chance that a random tax return from a specified professional group is legally correct and honest. Needless to say, this chance will vary from profession to profession. It is known that many self-employed professionals and especially those who deal exclusively with cash transactions are often involved in tax evasions. Given all tax returns, the main statistical approach is to partition data into two disjoint sets where data in one set is classified as honest and requires no prima facie reason for auditing and the data from the complement is classified as incorrect or dishonest reports that need auditing. There are two costs associated with any such classification. If an honest return is audited, the cost of auditing time is wasted on the return. If a manipulated return is not audited, then the cost is the loss in taxes properly due. We have to convert all costs into money for proper comparisons. Now any tax return x is simply a vector whose coordinates correspond to tax items such as 1) married or single, 2) gross income, 3) dividend or interest income, 4) employee business expense, 5) real estate taxes paid, etc. Thus, in general $x = (x_1, x_2, \dots, x_p)$ is a vector of observations. Some are qualitative and some quantitative. For simplicity, we will assume all are quantitative. Then the past data collected from the two sub-populations adjusted for inflation will give mean values and variances and covariances for each sub-population. Let $f_1(x)$ and $f_2(x)$ be the densities that represent the populations. If the populations are normal, f_1 and f_2 are uniquely determined by the mean vectors μ_1 and μ_2 and variance covariance matrices Σ_1 and Σ_2. Thus, an optimal procedure is one that minimizes expected costs of misclassification. For example, if population 1 corresponds to the honest and correct tax returns, then c_{12} = cost of classifying 1 into 2 for auditing = audit costs. Now a tax audit procedure has to decide which observations x have to be audited. Let $(\mathbb{R}, \mathbb{R}^c)$ be a partition of all observations into don't audit, audit classifications. Then given the prior $\xi = (\xi_1, \xi_2)$ and $\mathbb{R}$ the expected cost is simply

$$c_{12}\xi_1 \int_{\mathbb{R}^c} f_1(x)\, dx + c_{21}\xi_2 \int_{\mathbb{R}} f_2(x)\, dx.$$

We could rewrite the same as

$$c_{12}\xi_1 + \int_{\mathbb{R}} (c_{21}\xi_2 f_2 - c_{12}\xi_1 f_1)\, dx.$$

Thus, when the integrand is less than zero on $\mathbb{R}$ the expected cost is minimized, which is the same as saying that the optimal classification procedure $\mathbb{R}^*$ satisfies

$$\mathbb{R}^* = \left\{ x : \frac{c_{21}\xi_2}{c_{12}\xi_1} \leq \frac{f_1(x)}{f_2(x)} \right\}.$$

Equivalently, by taking logs, the procedure $\mathbb{R}^*$ reduces to

$$\mathbb{R}^* = \left\{ x : \log \frac{f_1(x)}{f_2(x)} \geq c \right\},$$

where c is known since $c_{21}, c_{12}, \xi_1, \xi_2$ are known.

If $\Sigma_1 = \Sigma_2$, $\log(f_1/f_2)$ to within some constant factor reduces to $\varrho(x) = (\mu_1 - \mu_2)^T \Sigma^{-1} x$. This is the famous *linear discriminant function* of Fisher ([3], [33]). The function $\varrho(x)$ is simply a linear combination of the x_i's for some suitable weights w_i's. We have the following intuitive interpretation of the discriminant function.

Each tax item i with reported x_i is given a weight w_i. The return is not audited if $\sum w_i x_i > c$, otherwise an audit is suggested.

Of course, what is mathematically easily said is quite hard to implement. Even statistical problems with cost coefficients, prior distributions, etc., are quite difficult to compute exactly.

For example, when certain professions are hard hit by federal regulations, the changes in the pattern of expenditures may not come through immediately. Say that doctors and hospitals are being pressured to charge only a fixed amount for a certain diagnostic treatment, then clearly the income of the profession is much affected. The life style cannot be changed and the temptation to get away from tax payments increases. One needs to study such complex phenomena with suitable models. As the priors (ξ_1, ξ_2) will also change, we need to find suitable models to analyze them.

4 A Need for Further Game Theoretic Refinement of the Discriminant Function Approach

In the discriminant function approach, though the individual returns are classified into one of two sub-populations within each professional category, *the dynamics of tax evasion from tax year to tax year and the strategic audit manipulations to curb the evasions are not at all captured by such a purely statistical model. Straightforward discriminant analysis ignores the strategic manipulations of individual taxpayers, a key element in tax returns.*

As a further refinement of the statistical discriminant function approach we propose to formulate various game theoretic models of multistage games that conceptually capture the essence of tax games between the IRS and individual returns.

Before modelling in full generality, we will introduce the notion of a zero-sum two-person stochastic game with two states and two actions for each player.

A Stochastic Game with Two States and Two Actions Consider two players playing one of the games A or B. In both games players secretly select one of the numbers 1 or 2. Depending on their choices an immediate reward is received by player I from player II. Their choices and the current game they play determine which game will be played next time. The following is a simple example of such a game:

$$\begin{array}{cc} \text{A} & \text{B} \\ \begin{matrix} 5/A & 0/B \\ 0/B & 3/A \end{matrix} & \begin{matrix} 4/B & 0/A \\ 2/A & 1/B \end{matrix} \end{array}$$

Let player I secretly select one of the rows and II secretly select one of the columns. If in A row 1 and column 2 are chosen, player I receives nothing and the game moves to playing B next round. If in B row 2, column 1 are their choices the game moves to A after a reward of 2 to player I from player II. The payoff accrues and future payoffs are discounted at a fixed discount rate β. The aim of player I is to maximize the total discounted payoff. The aim of player II is to minimize the same. If x_n is the payoff on the nth day, $\sum_{n=0}^{\infty} \beta^n x_n$ is the total payoff where $0 < \beta < 1$.

Shapley [38] proved the remarkable theorem that these games can be intelligently played by locally randomizing the selection of rows in each matrix independant of the history of the play leading to the given game. For example if $\beta = .8$, the game value starting in A is approximately 6.79; in B it is approximately 5.43. A good strategy for I is to choose row 1 in matrix A with a chance 0.223 and to choose row 1 all the time in matrix B. Player II should choose column 1 in A with chance 0.223 and column 2 all the time in matrix B.

5 A Simple Model of a Tax Return-Audit Game

Consider the population of professional engineers employed by engineering firms. Suppose that from past auditing the IRS has a hunch that 10% of them manipulate returns, while 90% are honest. Given a tax return x, the IRS can compute the discriminant function which could decide whether to audit or not. *However, the IRS may have an initial perception on a return, which may cause the agency to audit, even though the discriminant function may indicate the opposite.* Namely, besides the two actions available to the IRS, the perception of the IRS is a variable which could vary from year to year depending on the years past. *This year's data may not reveal it. Last year's perception alone could give some clue.* Thus, the perception of the IRS can be thought of as states of the game which, for example, can also vary between the two states: honest and manipulating. Only an audit can make perceptional changes. Even if the discriminant function favors auditing, it cannot be immediately implemented for want of staff. One may have to manage with existing staff, which means limiting thee auditing facility. In such a case, strategic selection of auditing may be the only alternative. Thus, we can think of a tax return as a game with the following interpretation.

- **Players:** I - IRS, II - individual or firm filing tax return
- **Pure strategies:**
  For player I: 1. audit 2. don't audit
  For player II: 1. honest return 2. cheat
- **States:**
  A. IRS perceives a return as honest
  B. IRS perceives a return as manipulating.
- **Law of motion or transition probabilities:**
  If a return is not audited, then the perception of the IRS is the same as it was the previous year. If an audit finds someone guilty of manipulation, the perception changes from honest to manipulating. This is our stochastic game. The other situations are given below as in our mathematical example of a stochastic game. The perception of the IRS in states A and B is given below:

$$\text{state A} = \text{(honest)}$$
$$\begin{array}{ll} & \quad\text{Honest} \quad \text{Manipulate} \\ \begin{array}{l}\text{Audit}\\ \text{Don't audit}\end{array} & \begin{bmatrix} a_1/A & b_1/B \\ c_1/A & d_1/A \end{bmatrix} \end{array}$$

$$\text{State B} = \text{(manipulating)}$$
$$\begin{array}{ll} & \quad\text{Honest} \quad \text{Manipulate} \\ \begin{array}{l}\text{Audit}\\ \text{Don't audit}\end{array} & \begin{bmatrix} a_2/.5A & b_2/B \\ c_2/B & d_2/B \end{bmatrix} \end{array}$$

  For example, in state B, the IRS could perceive a taxpayer as being susceptible for manipulations even if the current audit finds no tax evasion on the items audited. As a measure of deterrence, the IRS continues to view any past tax violators with a 50:50 suspicion even after a current audit finds them honest.
- **Rewards:**
  In parlor games the immediate rewards are well defined simply by the rules of the game. In modelling real problems as games the most thorny issue is to define meaningful payoffs. In the case of tax returns, the actual tax collected with or without audit can be taken to be the immediate payoff corresponding to independent choices by the tax office and taxpayer. This immediate payoff can be defined as the expected tax paid when not audited and the expected tax collected with suitable fines imposed when audited finds someone guilty or not guilty less audit costs.

First one needs to estimate the prior perception probabilities. The IRS will have on k random persons, data $x_1, x_2, \ldots, x_k$ on tax item i and $y_1, y_2, \ldots, y_k$

on tax item j where i and j are independent deduction items and where the k persons were audited for the first time. Let $\xi_1, \xi_2, \ldots, \xi_k$ and $\eta_1, \eta_2, \ldots, \eta_k$ be the revised amount for the same two items after audit. Let $\bar{x}, \bar{y}, \bar{\xi}$, and $\bar{\eta}$ be the averages.

$$\frac{1}{k}\left[\left|i : x_i - \bar{x} > 0, y_i - \bar{y} > 0\right| - \left|i : \xi_i - \bar{\xi} < 0 \text{ or } \eta_i - \bar{\eta} < 0\right|\right]$$

is a rough overestimate of the proportion of people who would have manipulated.

The credibility is maintained until an audit proves otherwise. The threat of audit should always be on any person to discourage any future manipulation. This is incorporated in the first row first column entry in matrix B. The game is played as an ordinary stochastic game with discounted payoff. The above model, though completely in line with a model of an ordinary stochastic game, misses an important ingredient of our tax return problem.

Suppose a tax officer has two file cabinets to store all tax returns. Depending on the current perception of the tax officer that a taxpayer is honest or cheating, he stores honest ones in cabinet A and the rest in cabinet B. Thus the actual cabinet in which one's current tax return is saved will be known only to the tax officer. A taxpayer can assume that his file is in file cabinet A when he has never been audited. If a taxpayer was, after an audit, found cheating some time in the past, even if he is found honest by later audits, the taxpayer cannot be sure where his file will be stored by the officer. Thus, the taxpayer is often ignorant of the current state (*perception of the officer*) of the stochastic game. Similarly, if the taxpayer cheats, the tax office will not know this without auditing. Thus the tax office is in general not fully informed about the past actions of the taxpayer. Full information about the current state and past actions, namely the partial history of the game is not fully known to both players. Currently, all the standard existence theorems for zero-sum stochastic games assume full information about past history of actions for both players. See [41].

6 Generalized Stochastic Game

A population Π is partitioned into n sub-populations $\pi_1, \pi_2, \ldots, \pi_n$. Player II selects secretly a sub-population π_j and chooses a random observation x from π_j. Only the observation x is revealed to player I. Independent of the observation revealed, player I has a fixed prior distribution $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ on the sub-population selected by player II. Initially player I selects secretly a P_i according to ξ. Given the data x from π_j unknown to player I, based on the observation x revealed he computes a set $A_i = i(x)$, a finite set of actions available in P_i. Now he chooses secretly an action $a \in i(x)$ and receives from player II an immediate reward $r(x, i, a)$ and the game moves to P_k from P_i with chance $q(k/i, x, a)$. Player II secretly chooses a new j and a random x' from π'_j and the x' is revealed to player I. He computes possible actions $A_k = k(x')$ and selects an action $a' \in k\,(x')$. Again he receives a reward $r\,(x', k, a')$ and so on. The payoff accrues each time and the future payoffs are discounted at a fixed discount rate $\beta, 0 < \beta < 1$. The aim of player I is to maximize the expected discount reward. The aim of player II is to minimize the same.

It will be convenient to motivate our above generalized stochastic game both from the point of view of statistical decision theory (Wald [49], Blackwell and Girshick [7], Ferguson [14] and multistage game theory (Filar and Vrieze [15]). First we will set up the correspondence between our generalized stochastic game and tax return-audit game. This is shown in Table 1.

Let A be the maximal finite set of all possible actions for all possible tax returns. Suppose that the set A has ϱ elements $\{a_1, a_2, \ldots, a_\varrho\}$. For each x one can associate a probability distribution $\left\{\phi_1^s(x), \phi_2^s(x), \ldots, \phi_\varrho^s(x)\right\}$ on the exhaustive action space $A = \{a_1, a_2, \ldots, a_\varrho\}$, where $\sum\limits_{i=1}^{\varrho} \phi_i^s(x) = 1$. Here s is the current perception of the IRS. Thus, we can associate a stationary strategy $\left\{\phi_1^s, \phi_2^s, \ldots, \phi_\varrho^s\right\}$ on the action space A for each x and current perception s. Let $\psi_1, \psi_2, \ldots, \psi_N$ be a probability distribution on $\{\pi_1, \pi_2, \ldots, \pi_N\}$. We are now ready to state some open problems.

Problem 6.1. Does the generalized β-discounted stochastic game admit a stationary optimal strategy for player I *assuming the following* conditions (1)–(4)?

(1) The partition $\pi = (\pi_1, \pi_2, \ldots, \pi_N)$ is the same for both players.
(2) $q\,(k/x, i, a_i)$ is known to both players.
(3) The discriminant function $i(x)$ and the associated norm violation resulting in possible audit action set $A_i = i(x)$ is known to player II, for each data x.
(4) The perception s of player I about player II is also known to player II.

Problem 6.2. When the data x comes from continuous densities corresponding to $\pi_1, \pi_2, \ldots, \pi_n$ and when the conditions of Problem 6.1 are satisfied, can one replace the stationary strategies $\{\phi_1^s(x), \phi_2^s(x), \ldots, \phi_r^s(x)\}$ by a pure strategy? That is, given data x do we have a single action for each perception P_s which is equivalent to $\{\phi_1^s(x), \phi_2^s(x), \ldots, \phi_r^s(x)\}$ in the sense of equivalent rewards?

In this context we want to recall the theorem of Dvoretsky, Wald, and Wolfowitz [12] in statistical decision theory.

Theorem 6.1. *Let Ω be a finite set of parameters representing states of nature. Let A be a finite subset of $\mathbb{R}^n$ representing actions of a statistician. Let f_ϖ for each ϖ be a continuous density function. Let D be the space of decisions where each $d \in D$ is a map $d : X \to A$ where X is the sample space. Let $L\,(\varpi, a)$ be a bounded measurable loss function. Then any randomized decision $\phi : x \to \{\phi_1(x), \phi_2(x), \ldots, \phi_k(x)\}$, where $\phi_i(x) =$ the chance action i is taken*

Table 1: Correspondence between the two games.

Player I	IRS
Player II sub-populations $\pi_1, \pi_2, \ldots, \pi_n$	Individual taxpayer Those within the professional group who manipulate a specific set of tax items in the tax return.
$P_1, P_2, \ldots, P_n$	In the eyes of IRS possible sets of items that are being manipulated by various types of persons in the profession. (Partition according to perception based on past data.)
A_i	For perception P_i, the set of actions available to IRS. (For example, if IRS suspects on, say, items 1) moving expenses and 2) charitable contributions. They may choose to audit on item 1, 2, both or none. These are 4 possible actions.)
$x = (x_1, x_2, \ldots, x_p)$	A vector of items filled in the tax return with some of the x_i's as qualitative variables such as marital status, filing status, etc.
$i(x)$	With each tax data vector x, a set A_i of audit actions that are needed to correct the *norm violations, as found in the data* via discriminant analysis. Whereas the perception of the IRS is based on *past tax returns* by the taxpayer, the actions at the current time are dictated by the current discriminant function. Thus the importance of the discriminant function lies in not just classifying the observation, but also suggesting possible audit actions, based on the *current data*.
$q(k/x, i, a)$	The chance that the tax return x, with an initial perception P_i and an action $a \in A_i = i(x)$ by IRS, changes the perception of the IRS from P_i to the new perception P_k.
$r(x, i, a)$	The actual tax collected when the return reports data x, when the perception by IRS is P_i and when action a is taken by the IRS.
β	Discount factor accounting for inflation rate, etc.

when x is observed, can be equivalently replaced by a pure decision $d : x \to A$ in the sense that the expected risk $r(\varpi, d) = r(\varpi, \phi)$ for all $\varpi \in \Omega$.

Our Problem 6.2 is to extend this theorem in the context of our generalized stochastic games.

7 Stochastic Games with Incomplete Information

From the point of view of our actual tax return-audit problem we need to handle the more difficult problem of lack of information from either side.[1] For example, the perception P_i of the IRS is rarely known to player II, the taxpayer. Also the law of motion $q(k/s, x, a_i)$ is unknown to the taxpayer. Actually, the theory of stochastic games with incomplete information has few computable solutions. The theory of structured stochastic games has many existence theorems and efficient algorithms to compute value and optimal or equilibrium strategies ([28],[48],[15],[30],[31]), and for structured repeated games (a very special class of stochastic games) with incomplete information of a special type, one has some existence theorems. However, there are very few computational tools. See [23], [40], [41], [36], and [37] (this volume) in recent years. The researches in the area of stochastic games with incomplete information that are close to our model are the ones by Melolidakis [23] and Rosenberg, Solon, and Vieille [36]. We could call our tax return-audit problem a statistical extension of stochastic games with incomplete information. In the next section, we will briefly discuss the notation of stochastic games with lack of information and show what our generalized stochastic game is with reference to this setup.

8 Games with Lack of Information on One Side

Games with incomplete information were pioneered by Harsanyi [20], and later formulated in some precise mathematical models for certain special kinds of information lags by Aumann and Maschler [5]. For more recent developments on stochastic games with incomplete information, see Sorin [41]. For our tax model what we will need is a certain subclass of games called stochastic games with lack of information on one side (SGLIOS) in the sense of [23], an adaptation of the Aumann–Maschler model for discounted and undiscounted stochastic games.

SGLIOS Model: A stochastic game with lack of information on one side consists of:

(i) A set of $m \times n$ matrices $S = \left\{A^1, A^2, \ldots, A^N\right\}$ called the "states" of the game. We identify A^s with state s.

[1] This critical aspect of the problem was first pointed out to the author by Professor Ritzburger, of the Institute of Advanced Study, Vienna, Austria.

(ii) A prior distribution $\xi^0 = \left(\xi_1^0, \ldots, \xi_N^0\right)$ on S.
(iii) A law of motion $q\,(t/s, i, j)$ where the game moves to state t from state s, when row i and column j are chosen secretly by the players in state s resulting in an immediate payoff $(A^s)_{ij} = a_{ij}^{(s)}$.
(iv) Player I alone knows the true state. Each time the choices i, j are revealed to both players after the choices are made.
(v) The immediate payoff $a_{ij}^{(s)}$ is kept secret from player II, though player I knows the same.
(vi) The payoff is evaluated by discounting each time with a discount factor β $(0 < \beta < 1)$.

The following is the main theorem.

Theorem 8.1 (Melolidakis)**.** *Let* Γ *be a* β*-discounted stochastic game of the above type SGLIOS.* We can associate an ordinary stochastic game*,* Γ^* *where player I has a pure optimal stationary strategy* $f^*\,(\xi)$ *and player II has a stationary optimal strategy* $g^*\,(\xi)$*. Here the game* Γ^* *is played as follows. Let player I, as in SGLIOS, use his usual information in selecting his behavioral strategy. Unlike in* Γ, *here player II is informed of the posterior distribution at each stage based on the state of the game, the actions of the player, and the law of motion. One of the main observations of Melolidakis is that the value* $v\,(\Gamma) = v\,(\Gamma^*)\,,$ *for player I loses nothing by revealing the posterior.*

As player I knows all about the law of motion, prior, state of the game, etc., one can consider the following stochastic game.
Game Γ^{**}*:* Let the action space of I be $\{f(s) : s \in S\}$ where $f(s)$ is a mixed strategy on the rows of A^s. Let the action space of II be the set $\{1, 2, \ldots, N\}$. Let the state space be all probability vectors in $\mathbb{R}^N$. Let the law of motion Q be $q\,(./\xi, f, j)$ where the new prior at t is the posterior given f, j, To clearly understand the new prior, as the posterior given the actions of the players in the original game, we evaluate the posterior probability of the game to be in state s, given the actions i, j of the two players. This is $\eta\,(s/f, j, \xi)$ where for i fixed the entry is

$$\begin{aligned}
\eta\,(t/i, j, \xi) &= \sum_s q\,(t/s, i, j)\, f_i(s)\xi(s) / \sum_s f_i(s)\xi(s). \\
r\,(\xi, f, j) &= \sum_s \sum_i \xi(s) f_i(s) r\,(s, i, j)\,.
\end{aligned}$$

Here $(A^s)_{ij} = r\,(s, i, j)$. This is the stochastic game induced by SGLIOS. An important observation is that the value of this stochastic game coincides with the value of the original stochastic game.

9 Similarities and Differences between the SGLIOS and Our Tax Return-Audit Game

In our tax return game player I (the IRS) knows about the actual state (the perception). The prior is also known to the IRS (based on past data). Instead of a finite set of actions for player II, the selection involves a partition and the selection of a random observation from a partitioned sub-population. While the actual actions are revealed in SGLIOS, here only the data x and action i, corresponding to the audit decision by the IRS, are known to each other. Whereas the actual payoff is unknown to player I but the law of motion $q\,(t/s, i, j)$ is known in SGLIOS, here the actual payoff is known but the law of motion $q\,(t/s, x, j)$ corresponding to data x, action j and change of perception to t from s is kept secret by the IRS. Intuitively the change of perception of the IRS about a taxpayer is kept secret even though the audit resulting in tax payments is known to both sides.

10 Some Thoughts on Norms for Audits and Some Questions on Revealing Audit Policies

In a penetrating paper on deriving norms for income tax audits, Pond [29] makes the following remarks: "There are always some, who through inadvertence or design, minimize their tax liability. Deductions offer one of the greatest avenues for minimizing tax liability and it is evident that to be most successful with the available staff, the audit program should concentrate on taxpayers whose deductions are excessive in relation to others in a comparable income classification. The first step in deriving the norms is to determine a frequency distribution for each deduction "and" calculate the ratio of net income to gross income! The carefully chosen tax returns for closer audit saves audit time without losing tax arrears. Such carefully chosen returns represent 5/6 of the total on all the cases. Since the audit agency was handling only 2/3 of all cases on a nonselected basis and therefore only was producing 4/6 of the potential, the application of norm method has an imputed gain of 25%."

In a sense many statistically heuristic procedures are already perhaps adopted by the IRS! As part of the deduction of Pond's paper the first question raised was whether it was possible for a taxpayer to become familiar with the selection criteria and thus become able to evade taxes and be sure of escaping detection. Two factors were seen mitigating against this: 1) The norms are kept secret, and 2) they are constantly being reevaluated. There was no general agreement on how the norms were to be evaluated. Fault was found with putting emphasis on assessment/cost ratio; the failure to audit in such cases might reduce voluntary compliance within those groups.

In our opinion, the verbal language above and the discussions pertaining to the problem of tax are in the spirit of multistage games. Thus a proper analysis of the problem via models of stochastic games is only desirable from both the theoretical and practical points of view. Perhaps solutions to such models might reveal answers to unanswered questions like the following.

Problem 10.1. Can the norms for auditing tax returns be made public?

An answer to Problem 10.1, though volatile, might still be valuable to look into for suitable models. With reference to our formulated model, this is the same as the following mathematical interpretation.

Problem 10.2. In our game with incomplete information if the state (IRS's perception) alone is kept secret, but not the actual law of motion of the game and the discriminant function and the partition $P_1, P_2, \dots, P_M$ by the IRS, will the value of the stochastic game change?

We could conceptually understand Problem 10.1 by modelling the game as the following single controller stochastic game with incomplete information.

11 A Single Controller Game with Incomplete Information

Players I and II know that a population Π is a mixture of sub-populations $\Pi_1, \Pi_2, \dots, \Pi_n$ with densities $f_1(x), f_2(x), \dots, f_n(x)$, respectively. Player II chooses secretly a $j \in \{1, 2, \dots, n\}$ and then selects a random observation x from Π_j. The choice j is not revealed to player I. However, the randomly selected observation x from density f_j is revealed to player I. Player I has prior $\xi_1, \xi_2, \dots, \xi_n$ on his perceptions about the current choice of player II. We call player I's current perception the state of the game. The perception of player I remains unchanged and stays at state s with probability $1-\phi(s)$, unknown to player II (here $0 \leq \phi(s) \leq 1$). With probability $\phi(s)$ the perception of player I changes to a new state k taking into account the posterior dictated by the data x. Based on x, he selects an action $i \in \{1, 2, \dots, n\}$ with probability $\psi_i(x)$. In the case $i \neq j$, player I receives a reward $c_i(s)$ from player II. In the case $i = j$ he receives an amount $u_j(x)$ from player II. The play continues with the posterior as the new prior. The payoff accrues at a fixed discount rate β. The aim of player I is to maximize the total discounted reward. The aim of player II is to minimize the same.

We could convert the above model into the following single controller stochastic game with incomplete information. We need to define immediate rewards and transition probabilities.

Let $q\left(k/s, \psi, j\right)$ = expected transition probability of the perception of player I to move from perception s to perception k given the strategies ψ and j by players I and II, respectively. Since the transition depends only on the posterior and the preassigned norms for remaining in status quo or following the decision based on the data, we get

$$q\,(k/s,\psi,j)=\begin{cases}\phi(s)\cdot\int\left[\frac{\xi_k f_k(x)}{\sum_t \xi_t f_t(x)}\right]\cdot f_j(x)\;dx, & k\neq s\\ [1-\phi(s)]+\phi(s)\cdot\int\left[\frac{\xi_s f_s(x)}{\sum_t \xi_t f_t(x)}\right]\cdot f_j(x)\;dx, & k=s.\end{cases}$$

Notice that $\xi_k f_k(x)/[\sum_t \xi_t f_t(x)]$ is the posterior probability given the data and q gives the expected transition probability.

An important observation is that this transition probability $q\,(k/s,\psi,j)$ depends only on the action of player II. Since $\phi(s)$ is unknown to player II, the actual law of motion is unknown to player II, although he controls the law of motion! The expected immediate reward $r\,(s,i,j)$ to player I can be written as

$$r\,(s,\psi,j)=\int_\Omega\left(\sum_{i\neq j}\psi_i(x)c_i(s)\right)f_j(x)dx+\int_\Omega\psi_j(x)u_j(x)f_j(x)dx.$$

Thus, our problem reduces to an ordinary one player control game if the perceptions of player I, u_j, c_i, ξ_i, etc., are common knowledge. Even if c_i's and ϕ's are known to player II, as long as the actual perception of player I about player II is kept secret, the game will still be a single controller stochastic game with lack of information on the law of motion for player II.

Thus we are led to the following problem.

Problem 11.1. Let Γ be a single controller stochastic game with reward $r\,(s,\psi,j)$, transition probabilities $q\,(k/s,j)$, and discount factor β, $0<\beta<1$. A prior distribution ξ on the states is chosen. Though the prior is known to both players, the actual state is known only to player I.

The law of motion $q\,(k/s,j)$ will be known to player I if action j of II is known. Even if $\phi(s)$ is revealed to player II, only the law of motion will be known to player II, but the true state s of the game will still be unknown to player II. The reward $r\,(s,\psi,j)$ is unknown to player I as he does not know j, the choice of player II. If $c_i(s)$ is independent of s, the reward in each state is $r\,(\psi,j)$, which depends only on the actions of players I and II. Even in this case the immediate payoff is unknown to either player as each one's choice remains secret in each round. A final case is when player II chooses a fixed j once and for all and all he does from one round to the next is choose an independent observation from the same density f_j. This is the closest to the single control games considered by Rosenberg, Solon, and Vieille [36]. In this case we are led to the one player control game with known reward, but unknown law of motion for the controlling player. The main problem is to find whether such games admit value and, given the data, to solve for the value and good strategies if any.

This game captures the spirit of tax return-audit in the following sense. The perceptions about a taxpayer by the IRS as honest, moderate, cheats on moving expenses, etc., have to be solely based on the data x the taxpayer submits and the action he chooses to get x. Thus, his own actions essentially contribute towards any changes in the perception of the IRS. The threat to keep him obedient to tax laws needs the secrecy of the perception (state). The norms governing the status of an audited taxpayer are captured by the function $\phi(s)$. We will briefly summarize the research findings of some earlier models that are found in the literature on public finance.

Generalized stochastic games: We have already described these games earlier. To focus just on the mathematical formulation of these games, we need only to modify our preceding formulation on stochastic games.

Generalized stochastic game: Players I and II play the following game. The game has a finite number of states $1, 2, \ldots, S$. In each state player I has m actions and player II has n actions. Player II selects an action j in state s. This is revealed to a referee. The referee picks a random observation x according to the density function $f_j(x)$ and reveals x to player I. Not knowing j, but knowing x, player I selects an action i among his m actions. Then he receives an amount $r(s, i, x)$, and the game moves to a state k with chance $q(k/s, i, j)$, and so on. The payoff as before is the total discounted payoff. Here a stationary strategy for player II is the same as before. However, a stationary strategy for player I is of the type $\phi_i^s(x)$ where $\phi_i^s(x) =$ the chance action i is selected in state s when observation x is given. It is a pure stationary strategy if $\phi_i^s(x) = 0$ or 1 for each x, s. The law of motion is common knowledge and the reward is known to both players.

Our problem is to check whether the game has optimal stationary strategies: Also, we are interested in the situation where the optimal stationary strategies are replaceable by pure stationary optimals. We have already discussed the special subclasses of such single controller games with incomplete information as a model of our tax problem.

12 A Model of Tax Evasion as a Stochastic Game with Incomplete Information on the States

Consider a population of taxpayers all belonging to a single professional category. A tax return is simply a p-vector $X = (X_1, X_2, \ldots, X_p)$ with X_1 as the adjusted gross income and X_p as the tax due, *as reported by the taxpayers* in their tax returns. Based on their past tax returns and past audit actions by the IRS, the current tax returns are stratified and stored in S distinct file cabinets $1, 2, \ldots, S$. In the *perception of the IRS*, based on most recent audits, the returns stored in file cabinet $j > i$ are viewed as higher-order tax violations than those in file cabinet i. We will assume that, by an audit, the tax office can always find out the true values $(Y_1, Y_2, \ldots, Y_p)$ for a taxpayer's tax return. If a taxpayer resorted to, say, the kth level of tax violation, then his reported tax return will be taken to be a vector function of the true values defined by

$$X_t = \phi_t^k(Y_1, Y_2, \dots, Y_p), \quad t = 1 \dots, p.$$

Let $f^1(x), f^2(x), \dots, f^S(x)$ be the joint density functions corresponding to the reported tax dues of the population of tax returns in file cabinets $1, \dots, S$. We will often use the random variables (X_1, X_p) with marginal joint densities $g^j(x_1, x_p)$, $j = 1, \dots, S$. In general the tax office maintains secrecy of the locations of individual files, and information about the marginal densities and joint densities of the returns stored in various file cabinets. When the IRS decides to audit a tax return from file cabinet k, the taxpayer will be notified about the current location k from where the return was chosen for audit.

Since the number of auditors is fixed, the tax office has to allocate the available auditing time A efficiently. An intuitive policy would be to rearrange the files in each file cabinet i from the smallest to the largest values of R where

$$R = \frac{\text{Adjusted Gross income}}{\text{Tax due}} = \frac{X_1}{X_p}$$

and target the upper end among them, namely those for whom $R > \rho$ for some ρ chosen secretly by the tax office. Obviously, there should be many deductions of various kinds to arrive at a relatively small tax, and the auditing hours will be longer on such tax returns. Just because R is large one cannot immediately conclude that the person is a cheater. The deductions could be genuine and the person could be honest. It could have been a bad year for the taxpayer with large hospital bills beyond insurance coverage. However, this ratio R is more likely to exceed the given value ρ in a population of higher-order tax violators than in a population of lower-order tax violators and in particular for honest taxpayers. Therefore, for any random tax return X^j from file cabinet j, let R^j denote the above-mentioned R value. Then for any random tax return X^i from file cabinet $i, P(R^j > \rho) \geq P(R^i > \rho)$ if $j > i$.

A strategy for the tax office is to choose a threshold value ρ and to target for audit all tax returns with $R > \rho$. Thus by the stochastic ordering assumption, a greater proportion of tax returns from file cabinet j will be targeted than from file cabinet $i < j$. For simplicity let us suppose the audit time to audit a tax return with value R is cR. Let $h_i(r)$ be the density of the reported value of R for the tax returns in the file cabinet i. Thus the expected audit time for file cabinet i is given by

$$c \int_\rho^\infty r h_i(r) dr = q_i.$$

This will also immediately fix the total audit time for all file cabinets as

$$\sum_i c \int_\rho^\infty r h_i(r) dr = \sum_i q_i.$$

Let $u^i(x_p)$ denote the density of the random variable X_p representing the tax amount on any random tax return from file cabinet i.

If by an audit the IRS comes to know that a taxpayer has indulged in a tax violation of order $k > i$, then the IRS classifies the current and future returns of the taxpayer in file cabinet k. If the audit reveals that the tax violation is of order $k < i$, the IRS classifies the current return and future returns of this taxpayer in file cabinet k. The following is the intuition for such a transition. If a person's return, either based on prior allocation or on recent audit was found to be a tax violation of a certain order, when audited currently is found to be one of a higher order, he/she deserves to be watched with immediate reclassification with necessary caution. The persons who are found from current audit to be improved with lower levels of violation are recognized for their acceptance of law and order with a slight bit of reservation. When a tax return is not audited the IRS loses tax on a taxpayer when he/she becomes a tax evader of a higher order. However when a tax return from a taxpayer who has considerably toned down from his/her original level of tax evasion is audited, the tax office incurs higher cost due to unnecessarily prolonged auditing. The transition probability based on the preceding intuitive principles can be defined as follows.

Let the transition probability be $q(j/s,k,\rho)$ where s is the current location (file cabinet) of the return from where the return with data X was picked for audit using ρ strategy and found to be of violation level k. In case $k \geq s$ the file is immediately transferred to file cabinet k with probability one. Suppose that the audit reveals a violation level $k < s$; then with a small probability α it is kept in the same file cabinet and with probability $1-\alpha$ it is moved to file cabinet k. Thus if a tax return X from file cabinet s is audited and if the taxpayer has chosen a tax violation level k currently, then the tax return moves to state j with transition probabilities given by

$$\begin{aligned} q(j/s,k,\rho) &= 1 \quad \text{if} \quad j=k \text{ and } k \geq s \\ &= \alpha \quad \text{if} \quad j=s \text{ and } k < s \\ &= 1-\alpha \quad \text{if} \quad j=k \text{ and } k < s. \end{aligned}$$

Suppose that the audit strategy ρ is chosen by the IRS. If a taxpayer has never been audited, then he can assume that his tax return is located in file cabinets $1, 2, \ldots, S$ with respective priors $\xi_1,\ \xi_2,\ \xi_S$. The priors are known to all taxpayers. If a taxpayer was audited in the past, based on the most recent audit he can evaluate the posterior probabilities for the current location of his tax return.

The revenue for the IRS from a taxpayer will depend on the following:

- Was he ever audited and if so what was the violation level of the most recent audit?
- Is he currently being audited?
- What is the current level of violation of the taxpayer?

Suppose that the taxpayer chooses currently a level $j = j(X)$ for tax violation. Including the current year suppose he has never been audited. Since his return could be from file cabinet i with stationary prior probability ξ_i it would have escaped the current audit if the calculated R value $X_1/X_p < \rho$. When it is not audited, he pays only X_p. Since he has chosen level j, the tax return can be thought of as a random observation from file cabinet j with density $f^j(x)$ and with $R < \rho$. Thus the conditional expected payoff to the tax office given that the tax return was in file cabinet i, and the current choice was j by the taxpayer, and it escaped audit currently is given by

$$\int\int_{\{(x_1,x_p):x_1<x_p\rho\}} x_p g^j(x_1, x_p) dx_1 dx_p.$$

Thus the expected income to the IRS from such a never-audited tax return is given by

$$\sum_j \xi_j \int\int_{\{(x_1,x_p):x_1<x_p\rho\}} x_p g^j(x_1, x_p) dx_1 dx_p.$$

Suppose that audit costs are w dollars per hour. The IRS charges a suitable penalty for tax violations depending on the level of tax violation when audited. Let each dollar due be multiplied by a penalty factor θ_k for tax returns audited from file cabinet i found to be a tax violation of level k. If the IRS charges a penalty proportional to the difference between the true tax due and reported tax amount X_p specified in the tax return, then the net expected income to the IRS from an audited tax return from file cabinet i with violation level $k > i$ is $r(i,k) = \theta_k(\mu_1 - \mu_k) + \mu_k - cw\int_\rho^\infty rg_i(r)dr$. Here $g_i(r)$ is the density of the statistic R from file cabinet i and μ_i = expected tax from file cabinet i for all files that escaped audit. Also $\theta_k > \theta_{k-1} \cdots > \theta_1 = 1$. In the case $k < i$, and if the taxpayer is audited then the tax office finds that the taxpayer is relatively reformed and the expected income to the tax office is $\theta_i(\mu_1 - \mu_i) + \mu_i - cw\int_\rho^\infty rg_i(r)dr$. The expected income to the tax office with the ρ strategy when the taxpayer in file cabinet i wants to choose tax violation level k is given by

$$\begin{aligned} &P_i(R < \rho)\theta_i(\mu_1 - \mu_i) + \mu_i - cw\int_\rho^\infty rg_i(r)dr \\ &\qquad + P_i(R > \rho)\theta_k(\mu_1 - \mu_k) + \mu_k - cw\int_\rho^\infty rg_i(r)dr. \end{aligned}$$

If the tax form for a taxpayer has p items to fill in with numerical values, any subset S among those items can be misrepresented by the taxpayer by deviating from the true value. Suppose that the tax office can identify the deviated items by audit; then the set of such deviators will constitute a sub-population with a density function f_S. In our model we assume that their tax returns are to be stored in a file cabinet labelled S. When a tax office calls a taxpayer for audit, and spells out where they have doubts on the tax return, they essentially reveal the label of the file cabinet from which this return is chosen for audit on the labelled items. Given the information that a taxpayer's file was stored in file cabinet S, after the most recent audit, a simple class of pure strategies for the taxpayer who wants to act like a random person from file cabinet A can be generated by any scale vector $a = (a_1, a_2, \ldots, a_p)$ that is used to fudge the true data X and report it as $Y = (Y_1, \ldots, Y_p)$ where $Y_i = a_i X_i, i = 1, \ldots, p$ and $A = \{i : a_i \neq 1\}$. Similarly, a simple pure strategy for the tax office is a choice of ρ that selects in the first round all tax data whose R value exceed ρ. If there are too many selected this way, a suitable stratified random sampling scheme can be used to select the size that is manageable with existing audit resources.

REFERENCES

[1] Allingham, M., Inequality and progressive taxation, *J. Public Economics,* 11, 273–274 (1979).

[2] Allingham, M. and A. Sandmo, Income tax evasion: A theoretical analysis, *J. Public Economics,* 1, 323–338 (1972).

[3] Anderson, T.W., *An Introduction to Multivariate Statistical Analysis*, Wiley, New York (1972).

[4] Arrow, K.J., *Essays in the Theory of Risk Bearing*, Markham, Chicago (1971).

[5] Aumann, R.J. and M. Maschler, *Repeated Games with Incomplete Information*, M.I.T Press, Cambridge, MA (1995).

[6] Balbir, S., Making honesty the best policy, *J. Public Economics,* 2, 257–263 (1973).

[7] Blackwell, D. and G.A. Girschick, *Theory of Games and Statistical Decisions*, Wiley, New York (1955).

[8] Clotfelter, C., Tax evasion and tax rates: An analysis of individual returns, *Review of Economics and Statistics,* 65, 363–373 (1983).

[9] Cross, R.B. and G.K. Shaw, The evasion-avoidance choice: A suggested approach, *Nat'l Tax Journal,* 34, 487–491 (1981).

[10] Davis, J.S., Hecht, G., and J.D. Perkins, Social behaviors, enforcement and tax compliance, *The Accounting Review,* 78, 39–69 (2003).

[11] Dubin, J.A., Graetz, M.J., and L.L. Wilde, Are we a nation of tax cheaters? New econometric evidence on tax compliance, *The American Economic Review,* 77, 240–245 (1987).

[12] Dvoretzky, A., Wald, A., and J. Wolfowitz, Elimination of randomization in certain statistical decision procedures and zero-sum two person games, *Ann. Math. Stat.,* 22, 1–21 (1951).

[13] Erard, B. and J. S. Feinstein, Honesty and evasion in the tax compliance game, *Rand J. Economics,* 25, 1–19 (1994).

[14] Ferguson, T.S., *Mathematical Statistics: A Decision Theoretic Approach*, Academic Press, New York (1967).

[15] Filar, J.A. and O.J. Vrieze, *Competitive Markov Decision Processes*, Springer-Verlag, Berlin (1996).

[16] Fisher, R.A., The use of multiple measurements in taxonomic problems, *Ann. Eugenics,* 7, 179–188 (1936).

[17] Frank, M. and M. Dekeyser, A tax discrepancy coefficient resulting from tax evasion or tax expenditures, *J. Public Economics,* 8, 67–78 (1977).

[18] Friedland, N., Maital, S., and A. Rutenberg, A simulation study of income tax evasion, *J. Public Economics,* 10, 107–116 (1978).

[19] Greenberg, J., Avoiding tax avoidance: A (repeated) game theoretic approach, *J. Economic Theory*, 32, 1–13 (1984).

[20] Harsanyi, J., Games with incomplete information played by Bayesian players, I–III, *Management Science,* 14, 159–182; 320–334; 486–502 (1967–68).

[21] Kolm, S.C., A note on tax evasion, *J. Public Economics,* 2, 265–270 (1973).

[22] Landsberger, M. and I. Meiljison, Incentive generating state dependent penalty system: The case on income tax evasion, *J. Public Economics,* 19, 333–352 (1982).

[23] Melolidakis, C.A., Stochastic games with lack of information on one side and positive stop probabilities, pp. 113–126, in: *Stochastic Games and Related Topics*, Raghavan, T.E.S., Ferguson, T.S., Parthasarathy, T., and O.J. Vrieze (eds.), Kluwer Publishers, Dordrecht, The Netherlands (1991).

[24] Monk, K.A., Income tax evasion: Some empirical evidence, *J. Public Finance*, 30, 70–76 (1975).

[25] Mookherjee, D. and I. Png, Optimal auditing, insurance and redistribution, *The Quarterly J. Economics,* 104, 399–415 (1989).

[26] Morrison, D.F., *Mutivariate Statistical Methods*, Fourth Edition, Duxbury Press, Pacific Grove, CA (2002).

[27] Murphy, J.H., EDP and tax administration in New York, *Nat'l Tax Journal*, 14, 223–227 (1961).

[28] Parthasarathy, T. and T.E.S. Raghavan, An order field property for stochastic games when one player controls transition probabilities, *J. Optimization Theory and Appl.*, 33, 375–392 (1981).

[29] Pond, C.B., Derivation and application of norms in selecting audit, *Nat'l Tax Journal,* 14, 227–231 (1984).

[30] Raghavan, T.E.S. and Z. Syed, A policy-improvement type algorithm for solving zero-sum two-person stochastic games of perfect information, *Math. Program.,* 95, 513–532 (2003).

[31] Raghavan, T.E.S. and Z. Syed, Computing stationary Nash equilibria of undiscounted single-controller stochastic games, *Math. Oper. Res.,* 27, 384–400 (2002).

[32] Rao, C.R., *Advanced Statistical Methods in Biometric Research*, Wiley, New York (1952).

[33] Rao, C.R., Discrimination among groups and assigning new individuals: The role of methodology of classification in psychiatry and psychopathology, 229–240, U.S. Dept. of Health Education and Welfare (1966).

[34] Reinganum, J. and L. Wilde, Income tax compliance in a principal agent framework, *J. Public Economics,* 26, 1–18 (1985).

[35] Reinganum, J. and L. Wilde, Equilibrium verification and reporting policies in a model of tax compliance, *International Economic Review,* 27, 739–760 (1986).

[36] Rosenberg, D., Solon, E., and N. Vieille, Stochastic games with a single controller and incomplete information, *SIAM Journal on Control and Optimization,* 43, 86–110 (2004).

[37] Rosenberg, D., Solon, E., and N. Vieille, Stochastic games with imperfect monitoring (this volume).

[38] Shapley, L.S., Stochastic games, *Proc. Nat'l Acad. Sci. U.S.A.,* 39, 1095–1100 (1953).

[39] Smith, W.H., Developing a new technique in selecting returns for audit, *J. Accountancy,* 123, 22–23 (1967).

[40] Sorin, S., Big Match with lack of information on one side, *Internat. J. Game Theory,* 14, 173–204 (1985).

[41] Sorin, S., *A First Course on Zero-Sum Repeated Games*, Springer-Verlag, New York (2002).

[42] Spicer, M.W., A behavioral model of tax evasion, Ph.D. Thesis, Ohio State University, Columbus, OH (1974).

[43] Spicer, M.W. and L.A. Becker, Fiscal inequity and tax evasion: An experimental approach, *Nat'l Tax Journal*, 33, 171–175 (1980).

[44] Spicer, M.W. and S.B. Lundstedt, *Public Finance*, 31, 295–305 (1976).

[45] Srinivasan, T.N., Tax evasion: A model, *J. Public Economics,* 2, 339–346 (1973).

[46] Strumpel, B., Contributions of survey research, in: A. Peacock (ed.), *Quantitative Analysis in Public Finance*, 29–32, Praeger Publ., New York (1969).

[47] Vogel, J., Taxation and public opinion in Sweden: An interpretation of recent data, *Nat'l Tax Journal,* 27, 499–514 (1974).

[48] Vrieze, O.J., Stochastic games with finite state and action spaces, Ph.D. Thesis, Free University of Amsterdam, The Netherlands (1983).

[49] Wald, A., *Statistical Decision Functions*, Wiley, New York (1950).