Annals of the International Society of Dynamic Games
Volume 10

Series Editor
Tamer Başar

Advances in Dynamic Games and Their Applications

Analytical and Numerical Developments

Pierre Bernhard
Vladimir Gaitsgory
Odile Pourtallier
Editors

Birkhäuser
Boston • Basel • Berlin

Editors
Pierre Bernhard
Département de Mathématiques
Appliquées et Modélisation
Polytech'Nice-Sophia
930 route des Colles–B.P. 145
06903 Sophia Antipolis Cedex
France
Pierre.Bernhard@polytech.unice.fr

Vladimir Gaitsgory
School of Mathematics
and Statistics
University of South Australia
Mawson Lakes Campus,
OC Building
Mawson Lakes SA 5095
Australia
Vladimir.Gaitsgory@unisa.edu.au

Odile Pourtallier
INRIA–Sophia Antipolis
2004 route des Lucioles–B.P. 93
06902 Sophia Antipolis Cedex
France
Odile.Pourtallier@sophia.inria.fr

ISBN: 978-0-8176-4833-6 e-ISBN: 978-0-8176-4834-3
DOI: 10.1007/978-0-8176-4834-3

Library of Congress Control Number: 2009921134

Mathematics Subject Classification (2000): 91Axx, 91Bxx, 90Cxx, 49Mxx

Printed on acid-free paper.

www.birkhauser.com

Contents

PART III Numerical Approaches to Dynamic-Differential Games

PART IV Applications

PART V Search Games

Preface

Modern dynamic game theory has its roots in pioneering works on differential games by R. Isaacs and L.S. Pontryagin, in seminal papers on extensive form games by H.W. Kuhn, and in works on stochastic games by L.S. Shapley. Since those early developments, dynamic game theory has evolved enormously and has branched out in many directions, spanning such diverse disciplines as applied mathematics, economics, system theory, engineering, operations research, biology, ecology, and the environmental sciences.

This edited volume is an outgrowth of selected papers presented at the 12th International Symposium on Dynamic Games (ISDG) and Applications that was held July 3-6, 2006, at Sophia Antipolis, France. It provides state-of-the-art information about new developments in theoretical and numerical analysis of dynamic games and their applications

The papers selected for the volume cover a variety of topics ranging from purely theoretical game-theoretic developments, to numerical analysis of various dynamic games, and to dynamic games applications in economics, finance, and energy supply. The list of contributors contains both well-known names and names of young researchers from all over the world. All papers included in the volume went through a stringent reviewing process, and collected together, they represent a state-of-the-art of the theory of dynamic games and their applications.

The volume is divided into eight parts (chapters), each including papers devoted to a certain topic. Part I (five papers) is devoted to theoretical developments in general dynamic and differential games. The topic of Part II (three papers) is pursuit-evasion games. Part III (five papers) deals with numerical approaches to dynamic and differential games. Part IV (two papers) is on applications of dynamic games in economics and option pricing. Parts V, VI, VII and VIII (each containing two papers) are devoted to search games, evolutionary games, stopping games, and stochastic games and "large neighborhood" games, respectively.

The Editors are indebted to many colleagues involved in the reviewing process and to Miss Giang Nguyen for her help in editing and formatting the volume.

Sophia Antipolis, France — *Pierre Bernhard*
Adelaide, Australia — *Vladimir Gaitsgory*
Sophia Antipolis, France — *Odile Pourtallier*

October 2008

Contributors

Andrei Akhmetzhanov, IPM Russian Academy of Sciences, Vernadsky Ave, 101-1, Moscow 117526, Russia

Steve Alpern, London School of Economics, Houghton Street, London WC2A 2AE, UK

Martino Bardi, Dipartimento di Matematica Pura ed Applicata, Università di Padova, via Trieste 63, 35121 Padova, Italy

Pierre Bernhard, Ecole Polytechnique, 06903 Sophia Antipolis Cedex, France and Laboratory I3S, University of Nice Sophia Antipolis and CNRS Sophia, Antipolis, France

Michael H. Breitner, Institut für Wirtschaftsinformatik, Gottfried Wilhelm Leibniz Universität Hannover, Königsworther Platz 1, D-30167 Hannover, Germany

Pierre Cardaliaguet, Laboratoire de Mathématiques, Université de Bretagne Occidentale, UMR 6205, 6 Av. Le Gorgeu, BP 809, 29285 Brest

Dean Carlson, Mathematical Reviews, American Mathematical Society, Ann Arbor, MI, USA

Diana Cierpiał, Department of Mathematics and Computer Science, Wrocław University of Technology, Wrocław, Poland

Emiliano Cristiani, Dipartimento di Matematica, SAPIENZA - Università di Roma, Rome, Italy

David Ellison, Ecole Polytechnique, 91128 Palaiseau Cedex, France

Jacob Engwerda, Tilburg University Dept. of Econometrics and O.R., P.O. Box: 90153, 5000 LE Tilburg, The Netherlands

Maurizio Falcone, Dipartimento di Matematica, SAPIENZA - Università di Roma, Rome, Italy

Elżbieta Z. Ferenstein, Faculty of Mathematics and Information Science, Warsaw University of Technology, 00-661 Warsaw, Poland

Robbert Fokkink, Delft Institute of Applied Mathematics, Delft University, P.O. Box 5031, 2600 GA Delft, The Netherlands

Ryusuke Hohzaki, Department of Computer Science, National Defense Academy, Yokosuka, 239-8686, Japan

Naira Hovakimyan, Department of Aerospace and Ocean Engineering, Virginia Polytechnic, Institute and State University, Blacksburg, Virginia, USA

Keiko Ikeda, Bureau of Defense Policy, Ministry of Defense, Tokyo, 162-8801, Japan

Anna Krasnosielska, Faculty of Mathematics and Information Science, Warsaw University of Technology, 00-661 Warsaw, Poland

Roy Lindelauf, Faculty of Economics and Business Administration, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands

Guiomar Martín-Herrán, Departamento de Economía Aplicada (Matemáticas), Universidad de Valladolid, Valladolid, Spain

Arik Melikyan, Institute for Problems in Mechanics, Russian Academy of Sciences, Moscow, Russia

Hans-Jörg von Mettenheim, Institut für Wirtschaftsinformatik, Gottfried Wilhelm Leibniz Universität Hannover, Königsworther Platz 1, D-30167 Hannover, Germany

Andrzej Nowakowski, Faculty of Mathematics, University of Lodz, Lodz, Poland

Geert Jan Olsder, Delft Institute of Applied Mathematics, Delft University, P.O. Box 5031, 2600 GA Delft, The Netherlands

Meir Pachter, Department of Electrical and Computer Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH 45433, USA

Valerii S. Patsko, Institute of Mathematics and Mechanics, S. Kovalevskaya str. 16, 620219, Ekaterinburg, Russia

Jos A.M. Potters, Radboud University Nijmegen, Nijmegen, The Netherlands

T.E.S. Raghavan, University Illinois Chicago, Chicago (Il), USA

David Ramsey, Department of Mathematics and Statistics, University of Limerick, Limerick, Ireland

Michel Rudnianski, CEREM & IIM (CNAM), 2 rue Conté, Paris 75003, France

A.J. Shaiju, Laboratoire I3S, UNSA-CNRS, 06903 Sophia Antipolis Cedex, France

Wen Shen, Department of Mathematics, Penn State University, McAllister Building, University Park, PA 16802, USA

Igor Shevchenko, North Pacific Fisheries Research Center (TINRO-Center), Vladivostok, Russia

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

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

Stef H. Tijs, University Tilburg, Tilburg, The Netherlands

Varvara L. Turova, Technical University of Munich, Boltzmannstr. 3, 85747, Garching, Germany

Andrzej Wieczorek, Institute of Computer Science, Polish Academy of Sciences, Poland

Georges Zaccour, Chair in Game Theory and Management and GERAD, HEC Montréal, Montréal (Québec) Canada H3T 2A7

On Differential Games with Long-Time-Average Cost*

Martino Bardi
Dipartimento di Matematica Pura ed Applicata
Università di Padova
via Trieste 63, 35121 Padova, Italy
bardi@math.unipd.it

Abstract

The paper deals with the ergodicity of deterministic zero-sum differential games with long-time-average cost. Some new sufficient conditions are given, as well as a class of games that are not ergodic. In particular, we settle the issue of ergodicity for the simple games whose associated Isaacs equation is a convex-concave eikonal equation.

Key words. Ergodic control, differential games, viscosity solutions, Hamilton-Jacobi-Isaacs equations.

AMS Subject Classifications. Primary 49N70; Secondary 37A99, 49L25, 91A23.

Introduction

We consider a nonlinear system in $\mathbb{R}^m$ controlled by two players

$$\dot{y}(t) = f(y(t), a(t), b(t)), \qquad y(0) = x, \quad a(t) \in A, \quad b(t) \in B, \tag{1}$$

and we denote with $y_x(\cdot)$ the trajectory starting at x. We are also given a bounded, uniformly continuous running cost l, and we are interested in the payoffs associated to the *long-time-average cost* (briefly, LTAC), namely:

$$J^\infty(x, a(\cdot), b(\cdot)) := \limsup_{T\to\infty} \frac{1}{T} \int_0^T l(y_x(t), a(t), b(t))\, dt,$$

$$J_\infty(x, a(\cdot), b(\cdot)) := \liminf_{T\to\infty} \frac{1}{T} \int_0^T l(y_x(t), a(t), b(t))\, dt.$$

We denote with $u - \mathrm{val}\, J^\infty(x)$ (respectively, $l - \mathrm{val}\, J_\infty(x)$) the upper value of the zero-sum game with payoff J^∞ (respectively, the lower value of the game

*This work has been partially supported by funds from M.I.U.R., project "Viscosity, metric, and control theoretic methods for nonlinear partial differential equations", and G.N.A.M.P.A., project "Equazioni alle derivate parziali e teoria del controllo".

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_1,

with payoff J_∞) which the 1st player $a(\cdot)$ wants to minimize while the 2nd player $b(\cdot)$ wants to maximize, and the values are in the sense of Varaiya-Roxin-Elliott-Kalton. We look for conditions under which

$$u - \mathrm{val}\, J^\infty(x) = l - \mathrm{val}\, J_\infty(x) = \lambda \quad \forall\, x,$$

for some constant λ, a property that was called *ergodicity of the LTAC game* in [3]. The terminology is motivated by the analogy with classical ergodic control theory, see, e.g., [30,14,28,9,25,6,7,2]. Similar problems were studied for some games by Fleming and McEneaney [21] in the context of risk-sensitive control, by Carlson and Haurie [16] within the turnpike theory, and by Kushner [29] for controlled nondegenerate diffusion processes. There is a large literature on related problems for discrete-time games; see the survey by Sorin [35].

More recently, several sufficient conditions for the ergodicity of the LTAC game were given by Ghosh and Rao [24] and Alvarez and the author [3]. Among other things, these papers clarified the connections with the solvability of the stationary Hamilton-Jacobi-Isaacs equation associated to the problem and with the long-time behavior of the value functions of the finite horizon games with the same running cost. In particular, under the classical Isaacs' condition

$$\min_{b\in B}\max_{a\in A}\{-f(y,a,b)\cdot p - l(y,a,b)\} = \max_{a\in A}\min_{b\in B}\{-f(y,a,b)\cdot p - l(y,a,b)\} \tag{2}$$

for all $y, p \in \mathbb{R}^m$, the LTAC game is ergodic with value λ if the viscosity solution $u(t,x)$ of the evolutive Hamilton-Jacobi-Isaacs equation

$$\frac{\partial u}{\partial t} + \min_{b\in B}\max_{a\in A}\{-f(y,a,b)\cdot D_x u - l(y,a,b)\} = 0, \qquad u(0,x) = 0,$$

satisfies

$$\lim_{t\to+\infty}\frac{u(t,x)}{t} = \lambda, \qquad \text{locally uniformly in } x,$$

a property called *ergodicity of the lower game*. However, the results of the quoted papers do not give much information on some very simple games such as:

$$\begin{cases} \dot{y}^A(t) = a(t), & y^A(0) = x^A \in \mathbb{R}^{m/2}, \quad |a(t)| \le 1, \\ \dot{y}^B(t) = b(t), & y^B(0) = x^B \in \mathbb{R}^{m/2}, \quad |b(t)| \le \gamma, \end{cases} \tag{3}$$

with running cost $l = l(y^A, y^B)$ independent of the controls and $\mathbb{Z}^m$-periodic. This is related to the asymptotic behavior of the solution to the *convex-concave eikonal equation*

$$u_t + |D_{x^A} u| - \gamma |D_{x^B} u| = l(x^A, x^B), \quad u(0, x^A, x^B) = 0,$$

where $D_{x^A}u, D_{x^B}u$ denote, respectively, the gradient of u with respect to the x^A and the x^B variables. From [3] we can only say that the lower game and the LTAC game are ergodic if l has a saddle, namely

$$\min_{x^A}\max_{x^B} l(x^A, x^B) = \max_{x^B}\min_{x^A} l(x^A, x^B) =: \bar{l},$$

and then the ergodic value is $\lambda = \bar{l}$. Nothing seems to be known if, for instance, $l(x^A, x^B) = n(x^A - x^B)$.

In the present paper, we present some new conditions for ergodicity and a class of non-ergodic differential games. The sufficient conditions for ergodicity assume some form of controllability of each player on some state variables. Different from the controllability conditions in [3], they depend on the running cost l, that is assumed independent of the controls, and give an explicit formula for the ergodic value λ in terms of l. The result of non-ergodicity holds for systems of the form

$$\begin{cases} \dot{y}^A(t) = g(y(t), a(t)), & y^A(0) = x^A \in \mathbb{R}^{m/2}, \quad a(t) \in A, \\ \dot{y}^B(t) = g(y(t), b(t)), & y^B(0) = x^B \in \mathbb{R}^{m/2}, \quad b(t) \in B, \end{cases} \tag{4}$$

with $A = B$, and running cost $l(x) = n(x^A - x^B) + h(x^A, x^B)$ with a smallness assumption on h. As a special case we settle the issue of the game (3) with the running cost $l(x) = n(x^A - x^B)$ and of the convex-concave eikonal equation: it is ergodic if and only if $\gamma \neq 1$.

Undiscounted infinite horizon control problems arise in many applications to economics and engineering; see [17,14,28] and [16,21,35] for games. Our additional motivation is that ergodicity plays a crucial role in the theory of singular perturbation problems for the dimension reduction of multiple-scale systems [27,14,28,23,36,26,32] and for the homogenization in oscillating media [31,19,5]. A general principle emerging in the papers [8,1,2,4] is that an appropriate form of ergodicity of the fast variables (for frozen slow variables) ensures the convergence of the singular perturbation problem, in a suitable sense. The explicit applications of the results of the present paper to singular perturbations will be presented in a future article.

The paper is organized as follows. Section 1 recalls some definitions and known results. Section 2 gives two different sets of sufficient conditions for the ergodicity of the finite horizon games. Section 3 presents the non-ergodic games. Section 4 applies the preceding results to a slight generalization of the system (3) and of the convex-concave eikonal equation.

1 Definitions and preliminary results

About the system (1) and the cost we assume throughout the paper that $f : \mathbb{R}^m \times A \times B \mapsto \mathbb{R}^m$ and $l : \mathbb{R}^m \times A \times B \mapsto \mathbb{R}$ are continuous and bounded, A and B are compact metric spaces, and f is Lipschitz continuous in x uniformly in a, b.

We consider the cost functional

$$J(T, x) = J(T, x, a(\cdot), b(\cdot)) := \frac{1}{T} \int_0^T l(y_x(t), a(t), b(t))\, dt,$$

where $y_x(\cdot)$ is the trajectory corresponding to $a(\cdot)$ and $b(\cdot)$. We denote with $\mathcal{A}$ and $\mathcal{B}$, respectively, the sets of open-loop (measurable) controls for the first and second player, and with Γ and Δ, respectively, the sets of nonanticipating strategies for the first and the second player; see, e.g., [18,20,9] for the precise definition. Following Elliott and Kalton [18], we define the upper and lower values for the finite horizon game with average cost:

$$u - \mathrm{val}\, J(T,x) := \sup_{\beta \in \Delta} \inf_{a \in \mathcal{A}} J(T,x,a,\beta[a]),$$

$$l - \mathrm{val}\, J(T,x) := \inf_{\alpha \in \Gamma} \sup_{b \in \mathcal{B}} J(T,x,\alpha[b],b).$$

The player using nonanticipating strategies has an information advantage with respect to the other, so the inequality $l - \mathrm{val}\, J(T,x) \leq u - \mathrm{val}\, J(T,x)$ holds; see [18,20,9]. Moreover, all other reasonable notions of value are between $l - \mathrm{val}\, J$ and $u - \mathrm{val}\, J$; see [18] or Chapter 8 of [9] for a discussion. Therefore, when the game has a value, i.e., $l - \mathrm{val}\, J = u - \mathrm{val}\, J$, all notions of value coincide. For the LTAC game we define:

$$u - \mathrm{val}\, J^{\infty}(x) := \sup_{\beta \in \Delta} \inf_{a \in \mathcal{A}} \limsup_{T \to \infty} J(T,x,a,\beta[a]),$$

$$l - \mathrm{val}\, J_{\infty}(x) := \inf_{\alpha \in \Gamma} \sup_{b \in \mathcal{B}} \liminf_{T \to \infty} J(T,x,\alpha[b],b).$$

Note that we chose $\limsup_{T \to \infty}$ for the upper value and $\liminf_{T \to \infty}$ for the lower value, so we expect again that any other definition of ergodic value falls between them.

We say that the *the lower game is (locally uniformly) ergodic* if the long time limit of the finite horizon value exists, locally uniformly in x, and it is constant, i.e.,

$$l - \mathrm{val}\, J(T,\cdot) \to \lambda \quad \text{as } T \to \infty \text{ locally uniformly in } \mathbb{R}^m.$$

Similarly, *the upper game is ergodic* if

$$u - \mathrm{val}\, J(T,\cdot) \to \Lambda \quad \text{as } T \to \infty \text{ locally uniformly in } \mathbb{R}^m.$$

The next result gives the precise connection between these properties and the LTAC game.

Theorem 1.1. *[21,3] If the lower game is ergodic, then*

$$l - \mathrm{val}\, J_{\infty}(x) = \lim_{T \to \infty} l - \mathrm{val}\, J(T,x) = \lambda \quad \forall x \in \mathbb{R}^m; \tag{5}$$

if the upper game is ergodic, then

$$u - \mathrm{val}\, J^{\infty}(x) = \lim_{T \to \infty} u - \mathrm{val}\, J(T,x) = \Lambda \quad \forall x \in \mathbb{R}^m. \tag{6}$$

If the classical Isaacs' condition (2) holds then the finite horizon game has a value, which we denote with $\mathrm{val}\, J(T,x)$; see [20,9]. Therefore, we immediately get the following consequence of Theorem 1.1.

Corollary 1.1. *Assume* (2) *and that either the lower or the upper game is ergodic. Then the LTAC game is ergodic, i.e.,*

$$l - \text{val}\, J_\infty(x) = u - \text{val}\, J^\infty(x) = \lim_{T\to\infty} \text{val}\, J(T,x) = \lambda, \quad \forall x \in \mathbb{R}^m.$$

Remark 1.1. The ergodic value can also be characterized as the limit as $\delta \to 0$ of δw_δ where w_δ solves

$$\delta w_\delta + \min_b \max_a \{-f(y,a,b) \cdot Dw_\delta - l(y,a,b)\} = 0, \quad \text{in } \mathbb{R}^m,$$

and as the unique constant λ such that there exists a solution of

$$\lambda + \min_b \max_a \{-f(y,a,b) \cdot D\chi - l(y,a,b)\} = 0, \quad \text{in } \mathbb{R}^m;$$

see [2,3,24] for the precise statements. We will not use these properties in the present paper.

2 Sufficient conditions of ergodicity

In this section we prove two results on the ergodicity of the LTAC games. Both make controllability assumptions on at least one of the players, but they are weaker than those of Theorem 2.2 in [3]. On the other hand, here we assume the running cost $l = l(y)$ depends only on the state variables and the controllability assumptions are designed to get as value of the LTAC game a number depending explicitly on l. In the first result this is either $\min l$ or $\max l$.

We denote with $\mathcal{KL}$ the class of continuous functions $\eta : [0,+\infty) \times [0,+\infty) \to [0,+\infty)$ strictly increasing in the first variable, strictly decreasing in the second variable, and satisfying

$$\eta(0,t) = 0 \quad \forall t \geq 0, \qquad \lim_{t\to+\infty} \eta(r,t) = 0 \quad \forall r \geq 0. \tag{7}$$

Given a closed target $\mathcal{T} \subseteq \mathbb{R}^m$, we say that the system (1) is *(uniformly) asymptotically controllable to $\mathcal{T}$ in the mean by the first player* if the following holds: there exists a function $\eta \in \mathcal{KL}$ and for all $x \in \mathbb{R}^m$, there is a strategy $\widetilde{\alpha} \in \Gamma$ such that:

$$\frac{1}{T}\int_0^T \text{dist}(y_x(t), \mathcal{T})\, dt \leq \eta(\|x\|, T), \ \forall b \in \mathcal{B}, \tag{8}$$

where $y_x(\cdot)$ is the trajectory corresponding to the strategy $\widetilde{\alpha}$ and the control function b, i.e., it solves

$$\dot{y}(t) = f(y(t), \widetilde{\alpha}[b](t), b(t)), \qquad y(0) = x. \tag{9}$$

Here $\|x\| := |x|$ in the general case, whereas when the state space is the m-dimensional torus $\mathbb{T}^m = \mathbb{R}^m/\mathbb{Z}^m$ (i.e., all data are $\mathbb{Z}^m$-periodic)

$$\|x\| := \min_{k\in\mathbb{Z}^m} |x - k|,$$

and $\mathrm{dist}(z, \mathcal{T}) := \inf_{w \in \mathcal{T}} \|z - w\|$. The condition (8) means that the first player can drive asymptotically the state $y(t)$ near the target $\mathcal{T}$, in the sense that the average distance tends to 0, uniformly with respect to x and the control of the other player b.

Symmetrically, we say that the system (1) is *(uniformly) asymptotically controllable to $\mathcal{T}$ by the second player* if for all $x \in \mathbb{R}^m$, there is a strategy $\widetilde{\beta} \in \Delta$ such that:

$$\frac{1}{T}\int_0^T \mathrm{dist}(y_x(t), \mathcal{T})\, dt \le \eta(\|x\|, T),\ \forall a \in \mathcal{A},$$

where $y_x(\cdot)$ is the trajectory corresponding to the strategy $\widetilde{\beta}$ and the control function a, i.e., it solves

$$\dot{y}(t) = f(y(t), a(t), \widetilde{\beta}[a](t)), \qquad y(0) = x.$$

In the next result we will use as target either the set

$$\text{argmin}\, l := \{y \in \mathbb{R}^m \ :\ l(y) = \min l\}$$

or the set

$$\text{argmax}\, l := \{y \in \mathbb{R}^m \ :\ l(y) = \max l\}.$$

Proposition 2.1. *Assume the running cost is uniformly continuous and independent of the controls, i.e., $l = l(y)$.*
If the system (1) *is asymptotically controllable to $\mathcal{T} = \mathrm{argmin}\, l$ in the mean by the first player, then the lower game is ergodic with value $\lambda = \min l$.*
If the system (1) *is asymptotically controllable to $\mathcal{T} = \text{argmax}\, l$ in the mean by the second player, then the upper game is ergodic with value $\lambda = \max l$.*

Proof. We prove only the first statement because the proof of the second is analogous. Set $v(T, x) := l - \text{val}\ J(T, x)$.

Fix x and consider the strategy $\widetilde{\alpha} \in \Gamma$ from the asymptotic controllability assumption. If $y_x(\cdot) = y_x(\cdot, b)$ is the corresponding trajectory and $z(t)$ is its projection on the target, i.e.,

$$\mathrm{dist}(y_x(t), \mathcal{T}) = \|y_x(t) - z(t)\|, \quad z(t) \in \mathcal{T},$$

then the choice $\mathcal{T} = \mathrm{argmin}\, l$ gives

$$l(y_x(t)) \le \omega_l(\|y_x(t) - z(t)\|) + l(z(t)) = \omega_l(\mathrm{dist}(y_x(t), \mathcal{T})) + \min l,$$

where ω_l is the modulus of continuity of l. We recall that ω_l is defined by

$$|l(x) - l(y)| \le \omega_l(\|x - y\|),\ \forall x, y \in \mathbb{R}^m, \quad \lim_{r \to 0} \omega_l(r) = 0,$$

and it is not restrictive to assume its concavity. Therefore, Jensen's inequality and (8) imply, for all $b \in \mathcal{B}$,

$$\frac{1}{T}\int_0^T \omega_l(\mathrm{dist}(y_x(t),\mathcal{T}))\,dt \le \omega_l(\eta(\|x\|,T)).$$

Then:

$$v(T,x) \le \sup_{b\in\mathcal{B}} \frac{1}{T}\int_0^T l(y_x(t))\,dt \le \omega_l(\eta(\|x\|,T)) + \min l.$$

On the other hand, $v(T,x) \ge \min l$ by definition, thus

$$\lim_{T\to\infty} v(T,x) = \min l,$$

uniformly in x for $\|x\|$ bounded. □

An immediate consequence of this proposition and of Corollary 1.1 is the following.

Corollary 2.1. *Assume the Isaacs' condition* (2) *and that the system* (1) *is asymptotically controllable either to* $\mathrm{argmin}\, l$ *by the first player or to* $\mathrm{argmax}\, l$ *by the second player. Then the LTAC game is ergodic, i.e.,*

$$l-\mathrm{val}\, J_\infty(x) = u-\mathrm{val}\, J^\infty(x) = \lim_{T\to\infty} \mathrm{val}\, J(T,x) = \lambda, \quad \forall x \in \mathbb{R}^m.$$

Moreover, $\lambda = \min l$ *in the former case and* $\lambda = \max l$ *in the latter.*

Remark 2.1. The main improvement of this result with respect to Corollary 2.1 in [3] is that here we assume only the controllability in the mean to a target, instead of the bounded-time controllability to each point of the state space. On the other hand, here we must assume the independence of l from the controls a, b.

Remark 2.2. A sufficient condition for the asymptotic controllability in the mean is that the system (1) be *locally bounded-time controllable to* $\mathcal{T}$ *by the first player*, i.e., for each x there exist $S(\|x\|) > 0$ and a strategy $\widetilde{\alpha} \in \Gamma$ such that for all control functions $b \in \mathcal{B}$ there is a time $t^\# = t^\#(x, \widetilde{\alpha}, b, \mathcal{T})$ with the properties

$$t^\# \le S(\|x\|) \text{ and } y_x(t) \in \mathcal{T} \text{ for all } t \ge t^\#.$$

In other words, the first player can drive the system from any initial position x to some point of the target $\mathcal{T}$ within a time that is uniformly bounded for bounded x, and keep it forever on $\mathcal{T}$, for all possible behaviors of the second player. The proof of Proposition 2.1 shows that this strategy is optimal for the first player. This kind of behavior is called a *turnpike*; see [17,16].

The sufficient condition described above can be better studied by splitting it in two: reaching $\mathcal{T}$ and remaining in $\mathcal{T}$ afterwards. The first amounts to the local boundedness of the lower value of the generalized pursuit-evasion game with target $\mathcal{T}$. This occurs if such value function is finite and continuous, and a sufficient condition for it is the existence of a continuous supersolution U of the Isaacs equation for minimum-time problems

$$\min_{b\in B}\max_{a\in A}\{-f(y,a,b)\cdot DU\}\geq 1 \qquad \text{in } \mathbb{R}^m\setminus\mathcal{T},$$

such that $U=0$ on $\mathcal{T}$; see [12,33]. As for the second property, it is the viability of $\mathcal{T}$ by the first player against the second. This is well understood and has explicit characterizations; see [15,11].

Remark 2.3. Another sufficient condition for the asymptotic controllability in the mean is that the system (1) be *worst-case stabilizable to $\mathcal{T}$ by the first player*, i.e., there exists $\kappa\in\mathcal{KL}$ and for each x there exists a strategy $\widetilde{\alpha}\in\Gamma$ such that:

$$\operatorname{dist}(y_x(t),\mathcal{T})\leq\kappa(\|x\|,t), \qquad \forall b\in\mathcal{B},\ \forall t\geq 0, \tag{10}$$

where $y_x(\cdot)$ is the trajectory corresponding to the strategy $\widetilde{\alpha}$ and the control function b. In fact, it is enough to take

$$\eta(r,T)=\frac{1}{T}\int_0^T\kappa(r,t)\,dt.$$

This property was studied by Soravia [33,34] and the author and Cesaroni [10]. They characterized it in terms of the existence of a Lyapunov pair, that is, a lower semicontinuous W, continuous at $\partial\mathcal{T}$ and proper, and a Lipschitz h, both positive off $\mathcal{T}$ and null on $\mathcal{T}$, such that:

$$\min_{b\in B}\max_{a\in A}\{-f(y,a,b)\cdot DW\}\geq h(x) \qquad \text{in } \mathbb{R}^m,$$

in the viscosity sense. Related notions are known in the context of robust control [13,22].

The second result of this section concern systems of the form:

$$\begin{cases}\dot{y}^A(t)=f_A(y(t),a(t),b(t)), & y^A(0)=x^A\in\mathbb{R}^{m_A},\\ \dot{y}^B(t)=f_B(y(t),a(t),b(t)), & y^B(0)=x^B\in\mathbb{R}^{m_B},\\ y(t)=(y^A(t),y^B(t)).\end{cases} \tag{11}$$

We will assume the asymptotic controllability by the first player to the target

$$\mathcal{T}^*:=\left\{(z^A,z^B)\in\mathbb{R}^m\ :\ l(z^A,z^B)\leq\max_{y^B}\min_{y^A}l(y^A,y^B)\right\}. \tag{12}$$

Moreover, we will assume for some closed set $\mathcal{T}_B \subseteq \mathbb{R}^{m_B}$ that the state variables y^B are *(uniformly) asymptotically controllable to* $\mathcal{T}_B$ *in the mean by the second player* in the following sense: there exists a function $\eta : [0,\infty) \to [0,\infty)$ satisfying (7) and for all $x \in \mathbb{R}^m$ there is a strategy $\widetilde{\beta} \in \Delta$ such that:

$$\frac{1}{T}\int_0^T \mathrm{dist}(y_x^B(t), \mathcal{T}_B)\, dt \leq \eta(\|x\|, T),\ \forall a \in \mathcal{A}. \tag{13}$$

Proposition 2.2. *Assume the system* (1) *is of the form* (11)*,* $l = l(y^A, y^B)$*, and* (2) *holds. Suppose also that the system is asymptotically controllable in the mean by the first player to* $\mathcal{T}^*$ *and the state variables* y^B *are asymptotically controllable in the mean by the second player to:*

$$\begin{aligned}\mathcal{T}_B &= \mathrm{argmax}\ \min_{y^A} l(y^A, \cdot) \\ &:= \left\{ z^B \in \mathbb{R}^{m_B}\ :\ \min_{y^A} l(y^A, z^B) = \max_{y^B}\min_{y^A} l(y^A, y^B) \right\}.\end{aligned} \tag{14}$$

Then the LTAC game is ergodic and its value is:

$$l - \mathrm{val}\, J_\infty(x) = u - \mathrm{val}\, J^\infty(x) = \lambda := \max_{y^B}\min_{y^A} l(y^A, y^B), \quad \forall x \in \mathbb{R}^m. \tag{15}$$

Proof. The Isaacs conditions (2) implies the existence of the value of the finite horizon games and we set $v(T,x) := l - \mathrm{val}\ J(T,x) = u - \mathrm{val}\ J(T,x)$. By repeating the proof of Proposition 2.1 with the target $\mathcal{T}^*$ we obtain:

$$\begin{aligned} l(y_x(t)) &\leq \omega_l(\|y_x(t) - z(t)\|) + l(z(t)) \\ &\leq \omega_l(\mathrm{dist}(y_x(t), \mathcal{T})) + \max_{y^B}\min_{y^A} l(y^A, y^B), \end{aligned} \tag{16}$$

and then

$$v(T,x) \leq \omega_l(\eta(\|x\|, T)) + \lambda.$$

To get the opposite inequality fix x and consider the strategy $\widetilde{b} \in \Delta$ from the asymptotic controllability assumption on the y^B variables. Let $y_x(\cdot) = y_x(\cdot, a)$ be the corresponding trajectory and $z^B(t)$ the projection of its component $y_x^B(t)$ on the target $\mathcal{T}_B$, i.e.,

$$\mathrm{dist}(y_x^B(t), \mathcal{T}_B) = \|y_x^B(t) - z^B(t)\|, \quad z^B(t) \in \mathcal{T}_B.$$

Then the definition of $\mathcal{T}_B$ gives

$$\begin{aligned} l(y_x(t)) &\geq l(y_x^A(t), z^B(t)) - \omega_l(\|y_x^B(t) - z^B(t)\|) \\ &\geq \max_{y^B}\min_{y^A} l(y^A, y^B) - \omega_l(\mathrm{dist}(y_x^B(t), \mathcal{T}_B)), \end{aligned} \tag{17}$$

where the modulus of continuity ω_l was defined in the proof of Proposition 2.1. The concavity of ω_l, Jensen's inequality, and (13) imply, for all $a \in \mathcal{A}$,

$$\frac{1}{T}\int_0^T \omega_l(\mathrm{dist}(y_x^B(t), \mathcal{T}_B))\,dt \le \omega_l(\eta(\|x\|, T)).$$

Finally, the definition of upper value gives

$$v(T,x) \ge \inf_{a\in\mathcal{A}} \frac{1}{T}\int_0^T l(y_x(t))\,dt \ge \lambda - \omega_l(\eta(\|x\|, T)),$$

and therefore $\lim_{T\to\infty} v(T,x) = \lambda$ uniformly in x, for $\|x\|$ bounded. □

Remark 2.4. The main differences of this result with respect to Proposition 2.2 in [3] is that here we do not assume that the cost l has a saddle and we make different controllability assumptions that give an advantage to the first player.

Remark 2.5. By exchanging the assumptions on the two players and replacing maxmin with minmax in the targets, it is easy to give a symmetric result where the value of the LTAC game is:

$$\lambda = \min_{y^A\in\mathbb{R}^{m_A}} \max_{y^B\in\mathbb{R}^{m_B}} l(y^A, y^B).$$

Remark 2.6. No controllability assumption is necessary for ergodicity, and in general, the ergodic value λ can be any number between $\min l$ and $\max l$. For instance, by a classical result of Jacobi (see, e.g., [7,2]), the system $\dot{y}(t) = \xi$ with $\xi \cdot k \ne 0$ for all $k \in \mathbb{Z}^m$ is uniformly ergodic for all $\mathbb{Z}^m$-periodic l, and the ergodic value is $\lambda = \int_{[0,1]^m} l(y)\,dy$.

Remark 2.7. The results of this section can be extended to control systems driven by stochastic differential equations, as in Sec. 4 of [3]. We postpone this to a future paper.

3 Sufficient conditions for non-ergodicity

In this section we give some examples of games that are not ergodic. Let us first recall that a simple reason for non-ergodicity is the unboundedness of the trajectories, as shown in the next example.

Example 3.1. Consider the system $\dot{y} = y$ and running cost l such that there exist the limits $\lim_{x\to+\infty} l(x) = l_+$ and $\lim_{x\to-\infty} l(x) = l_-$. Then $\mathrm{val}\, J(T,x) = \frac{1}{T}\int_0^T l(xe^t)\,dt$ converges as $t \to +\infty$ to l_+ if $x > 0$, to $l(0)$ if $x = 0$, and to l_- if $x < 0$.
However, also on a compact state space such as $\mathbb{T}^2$, many systems are not ergodic, such as the next simple example.

Example 3.2. In $\mathbb{R}^2$ take the system $\dot{y} = (1,0)$ and l $\mathbb{Z}^2$-periodic. Then val $J(T,x) = \frac{1}{T}\int_0^T l(x_1+t,x_2)\,dt$ converges as $t \to +\infty$ to $\int_{[0,1]^2} l(s,x_2)\,ds$.

The main result of the section is about systems of the form (11) under assumptions that allow the controllability of the variables y^A by the first player and of y^B by the second player as in the ergodic games described in [3]. However, the running cost does not have a saddle and the system is completely fair, in the sense that both groups of variables have the same dynamics. Here are the precise assumptions. Suppose first that the vector field f_A is independent of b, f_B does not depend on a, and l depends only on the state y. Then the Isaacs condition (2) holds and the Hamiltonian takes the split form:

$$\begin{aligned} H(y,p) &:= \min_{b\in B}\max_{a\in A}\{-f(y,a,b)\cdot p - l(y,a,b)\} \\ &= \max_{a\in A}\{-f_A(y,a)\cdot p^A\} + \min_{b\in B}\{-f_B(y,b)\cdot p^B\} - l(y), \quad p=(p^A,p^B). \end{aligned}$$

Assume further that $A = B$, $m_A = m_B = m/2$, and $f_A = f_B =: g$, so the system takes the form (4). Then, if we define the reduced Hamiltonian

$$H_r(y,q) := \max_{a\in A}\{-g(y,a)\cdot q\}, \quad q \in \mathbb{R}^{m/2},$$

the Hamiltonian H becomes

$$H(y,p) = H_r(y,p^A) - H_r(y,-p^B) - l(y), \quad p=(p^A,p^B). \tag{18}$$

We will also take the running cost of the form

$$l(y) = n(y^A - y^B) + h(y^A,y^B), \quad y=(y^A,y^B), \tag{19}$$

and make assumptions of the functions $n : \mathbb{R}^{m/2} \to \mathbb{R}$ and $h : \mathbb{R}^m \to \mathbb{R}$.

Theorem 3.1. *Assume the Hamiltonian H has the form* (18) *with running cost of the form* (19) *and n, h bounded and uniformly continuous. If*

$$\sup h - \inf h < \sup n - \inf n, \tag{20}$$

then the lower and the upper game are not ergodic.

Proof. We explain first the idea in the special case $h \equiv 0$, $n \in C^1$. In this case, $u(t,y) := t\, n(y^A - y^B)$ solves the Hamilton-Jacobi-Isaacs equation:

$$u_t + H_r(y, D_{y^A}u) - H_r(y,-D_{y^B}u) = n(y^A - y^B).$$

If $v(t,y)$ is the value function of the finite horizon game, then $tv(t,y)$ solves the partial differential equation in viscosity sense [20,9] and takes the same initial value 0 at $t=0$. By the uniqueness of the viscosity solution to the Cauchy problem

[9], $tv(t,y) = u(t,y)$. Then $v(t,y) = n(y^A - y^B)$ does not converge to a constant as $t \to \infty$ because n is not constant.

The general case is a perturbation of the preceding one. Take a mollification $n^\varepsilon \in C^1$ of n such that $n^\varepsilon \to n$ as $\varepsilon \to 0$ uniformly in $\mathbb{R}^{m/2}$. Consider $u(t,y) := t\, n^\varepsilon(y^A - y^B) + tc$, for a constant c to be determined. Then:

$$u_t + H_r(y, D_{y^A} u) - H_r(y, -D_{y^B} u) = n^\varepsilon(y^A - y^B) + c \tag{21}$$

and the right-hand side is $\geq n(y^A - y^B) + h(y)$ for $c = \sup h + \delta$, $\delta > 0$, if ε is small enough. Therefore the comparison principle between viscosity sub- and supersolutions [9] gives:

$$v(t,y) \leq n^\varepsilon(y^A - y^B) + c, \quad \forall\, t, y,$$

and for y_1 such that $n(y_1^A - y_1^B)$ is close to $\inf n$ and ε small enough

$$v(t, y_1) \leq \inf n + \sup h + 2\delta. \tag{22}$$

On the other hand, the right-hand side of (21) is $\leq n(y^A - y^B) + h(y)$ for $c = \inf h - \delta$ and ε small enough. Then:

$$v(t,y) \geq n^\varepsilon(y^A - y^B) + c, \quad \forall\, t, y$$

and

$$v(t, y_2) \geq \sup n + \inf h - 2\delta, \tag{23}$$

if $n(y_2^A - y_2^B)$ is close to $\sup n$ and ε is small enough. By condition (20) we can choose δ so that the right-hand side of (22) is smaller than the right-hand side of (23). Then $v(t,y)$ cannot converge to a constant as $t \to \infty$. □

4 An example: the convex-concave eikonal equation

In this section we fix $g : \mathbb{R}^m \to \mathbb{R}$ Lipschitzean and such that $g(y) \geq g_o > 0$, and discuss the ergodicity of the games where the system is

$$\begin{cases} \dot{y}^A(t) = g(y(t))a(t), & y^A(0) = x^A \in \mathbb{R}^{m/2}, \quad |a(t)| \leq 1, \\ \dot{y}^B(t) = g(y(t))b(t), & y^B(0) = x^B \in \mathbb{R}^{m/2}, \quad |b(t)| \leq \gamma, \\ y(t) = (y^A(t), y^B(t)), \end{cases} \tag{24}$$

for all values of the parameter $\gamma > 0$. For a running cost l independent of the controls, the finite horizon game has a value $\mathrm{val}(t,x) := l - \mathrm{val}\, J(t,x) = u - \mathrm{val}\, J(t,x)$, and $u(t,x) = t\, \mathrm{val}(t,x)$ solves the Hamilton-Jacobi-Isaacs equation

$$u_t + g(x)|D_{x^A} u| - \gamma g(x)|D_{x^B} u| = l(x), \quad u(0,x) = 0, \tag{25}$$

that we call the convex-concave eikonal equation. As in the preceding section we take l of the form

$$l(y) = n(y^A - y^B) + h(y^A, y^B), \quad y = (y^A, y^B).$$

We also need a compact state space, so we assume for simplicity that all data g, n, h are $\mathbb{Z}^m$ periodic. We recall that the paper by Alvarez and the author [3] covers only the case that $n \equiv 0$ and h has a saddle point, and then the value of the LTAC game is $\lambda = \min_{y^A \in \mathbb{R}^{m/2}} \max_{y^B \in \mathbb{R}^{m/2}} h(y^A, y^B) = \max_{y^B \in \mathbb{R}^{m/2}} \min_{y^A \in \mathbb{R}^{m/2}} h(y^A, y^B)$.

Corollary 4.1. *Under the preceding assumptions, the upper, lower, and LTAC game are ergodic under either one of the following conditions:*

i) $\gamma < 1$ *and* $h = h(y^B)$ *is independent of* y^A*, and in this case*

$$\lim_{t \to \infty} v(t, x) = \min n + \max h;$$

ii) $\gamma > 1$ *and* $h = h(y^A)$ *is independent of* y^B*, and in this case*

$$\lim_{t \to \infty} v(t, x) = \max n + \min h.$$

If, instead, $\gamma = 1$ *and*

$$\sup h - \inf h < \sup n - \inf n,$$

then the upper and lower game are not ergodic.

Proof. If $\gamma < 1$, since the dynamics of the y^A and the y^B variables is the same, but the first player can drive y^A at higher speed, for any fixed $z \in \mathbb{R}^{m/2}$ the first player can drive the system from any initial position to $y^A = y^B + z$ in finite time for all controls of the second player. Since $\mathbb{T}^m$ is compact this can be done in a uniformly bounded time. In particular, the system is asymptotically controllable by the first player to the set

$$\mathcal{T}_n := \left\{ (y^A, y^B) \in \mathbb{R}^m \,:\, y^A - y^B \in \operatorname{argmin} n \right\}.$$

If $h \equiv 0$ we can conclude by Proposition 2.1. Note that in this case we do not need the controllability of y^B by the second player.

In the general case, we observe that $\mathcal{T}_n$ is a subset of the target $\mathcal{T}^*$ defined by (12), because here:

$$\max_{y^B \in \mathbb{R}^{m/2}} \min_{y^A \in \mathbb{R}^{m/2}} l(y^A, y^B) = \min n + \max_{y^B \in \mathbb{R}^{m/2}} h(y^B).$$

Therefore the system is asymptotically controllable to $\mathcal{T}^*$ by the first player.

On the other hand, the variables y^B are bounded time controllable by the second player to any point of $\mathbb{R}^{m/2}$; therefore they are also asymptotically controllable to $\mathcal{T}_B$. Then Proposition 2.2 gives the conclusion *i)*.

The statement *ii)* is proved in the same way by reversing the roles of the two players. In this case:

$$\min_{y^A \in \mathbb{R}^{m/2}} \max_{y^B \in \mathbb{R}^{m/2}} l(y^A, y^B) = \max n + \min_{y^A \in \mathbb{R}^{m/2}} h(y^A).$$

Finally, the case $\gamma = 1$ follows immediately from Theorem 3.1. □

References

[1] O. Alvarez and M. Bardi, *Viscosity solutions methods for singular perturbations in deterministic and stochastic control*, SIAM J. Control Optim. **40** (2001), 1159–1188.

[2] O. Alvarez and M. Bardi, *Singular perturbations of degenerate parabolic PDEs: a general convergence result*, Arch. Rational Mech. Anal. **170** (2003), 17–61.

[3] O. Alvarez and M. Bardi, *Ergodic problems in differential games*, in "Advances in Dynamic Game Theory", S. Jorgensen, M. Quincampoix, and T.L. Vincent, eds., pp. 131–152, Ann. Internat. Soc. Dynam. Games, vol. 9, Birkhäuser, Boston, 2007.

[4] O. Alvarez and M. Bardi, *Ergodicity, stabilization, and singular perturbations of Bellman-Isaacs equations*, Memoirs American Math. Soc., to appear.

[5] O. Alvarez, M. Bardi, and C. Marchi, *Multiscale problems and homogenization for second-order Hamilton-Jacobi equations*, J. Differential Equations **243** (2007), 349–387.

[6] M. Arisawa, *Ergodic problem for the Hamilton-Jacobi-Bellman equation II*, Ann. Inst. H. Poincaré Anal. Non Linéaire **15** (1998), 1–24.

[7] M. Arisawa and P.-L. Lions, *On ergodic stochastic control*, Comm. Partial Diff. Eq. **23** (1998), 2187–2217.

[8] Z. Artstein and V. Gaitsgory, *The value function of singularly perturbed control systems*, Appl. Math. Optim. **41** (2000), 425–445.

[9] M. Bardi and I. Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, Birkhäuser, Boston, 1997.

[10] M. Bardi and A. Cesaroni, Almost sure properties of controlled diffusions and worst case properties of deterministic systems, ESAIM Control Optim. Calc. Var. **14** (2008), 343–355.

[11] M. Bardi and P. Goatin, *Invariant sets for controlled degenerate diffusions: a viscosity solutions approach*, in “Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming”, W.M. McEneaney, G.G. Yin, and Q. Zhang, eds., pp. 191–208, Birkhäuser, Boston, 1999.

[12] M. Bardi and P. Soravia, Hamilton-Jacobi equations with singular boundary conditions on a free boundary and applications to differential games. Trans. Amer. Math. Soc. **325** (1991), 205–229.

[13] T. Başar and P. Bernhard, *H^∞-optimal control and related minimax design problems. A dynamic game approach*, 2nd edition, Birkhäuser, Boston, 1995.

[14] A. Bensoussan, *Perturbation methods in optimal control*, Wiley/Gauthiers-Villars, Chichester, 1988.

[15] P. Cardaliaguet, *A differential game with two players and one target*, SIAM J. Control Optim. **34** (1996), 1441–1460.

[16] D.A. Carlson and A.B. Haurie, *A turnpike theory for infinite horizon open-loop differential games with decoupled controls*, in “New Trends in Dynamic Games and Applications", G.J. Olsder, ed., pp .353–376, Ann. Internat. Soc. Dynam. Games, vol. 3, Birkhäuser Boston, 1995.

[17] D.A. Carlson, A.B. Haurie, and A. Leizarowitz, *Infinite horizon optimal control: Deterministic and stochastic systems*, Springer-Verlag, Berlin, 1991.

[18] R.J. Elliott and N.J. Kalton, *The existence of value in differential games*, Memoirs American Math. Soc., vol. 126. American Mathematical Society, Providence, RI, 1972.

[19] L. Evans, *Periodic homogenisation of certain fully nonlinear partial differential equations*, Proc. Roy. Soc. Edinburgh Sect. A **120** (1992), 245–265.

[20] L. Evans and P.E. Souganidis, *Differential games and representation formulas for solutions of Hamilton-Jacobi-Isaacs equations*, Indiana Univ. Math. J. **33** (1984), 773–797.

[21] W.H. Fleming and W.M. McEneaney, *Risk-sensitive control on an infinite time horizon*, SIAM J. Control Optim. **33** (1995), 1881–1915.

[22] R.A. Freeman and P.V. Kokotovic, *Robust nonlinear control design. State-space and Lyapunov techniques*, Birkäuser, Boston, 1996.

[23] V. Gaitsgory, *Limit Hamilton-Jacobi-Isaacs equations for singularly perturbed zero-sum differential games*, J. Math. Anal. Appl. **202** (1996), 862–899.

[24] M.K. Ghosh and K.S.M. Rao, *Differential Games with Ergodic Payoff* SIAM J. Control Optim. **43** (2005), 2020–2035.

[25] L. Grüne, *On the relation between discounted and average optimal value functions*, J. Diff. Eq. **148** (1998), 65–99.

[26] Y. Kabanov and S. Pergamenshchikov, *Two-scale stochastic systems. Asymptotic analysis and control*, Springer-Verlag, Berlin, 2003.

[27] P.V. Kokotović, H.K. Khalil, and J. O'Reilly, *Singular perturbation methods in control: analysis and design*, Academic Press, London, 1986.

[28] H.J. Kushner, *Weak convergence methods and singularly perturbed stochastic control and filtering problems*, Birkhäuser, Boston, 1990.

[29] H.J. Kushner, *Numerical approximations for stochastic differential games: the ergodic case*, SIAM J. Control Optim. **42** (2004), 1911–1933.

[30] P.-L. Lions, *Neumann type boundary conditions for Hamilton-Jacobi equations*, Duke Math. J. **52** (1985), 793–820.

[31] P.-L. Lions, G. Papanicolaou, and S.R.S. Varadhan, *Homogenization of Hamilton-Jacobi equations*, Unpublished, 1986.

[32] M. Quincampoix and F. Watbled, *Averaging method for discontinuous Mayer's problem of singularly perturbed control systems*, Nonlinear Anal. **54** (2003), 819–837.

[33] P. Soravia, *Pursuit-evasion problems and viscosity solutions of Isaacs equations*, SIAM J. Control Optim. **31** (1993), 604–623.

[34] P. Soravia, *Stability of dynamical systems with competitive controls: the degenerate case*, J. Math. Anal. Appl. **191** (1995), 428–449.

[35] S. Sorin, *New approaches and recent advances in two-person zero-sum repeated games*, in "Advances in Dynamic Games", A.S. Nowak and K. Szajowski, eds., pp. 67–93, Ann. Internat. Soc. Dynam. Games, vol. 7, Birkhäuser, Boston, 2005.

[36] N.N. Subbotina, *Asymptotics for singularly perturbed differential games*, in "Game Theory and Applications", vol. VII, pp. 175–196, Nova Science Publishers, Huntington, NY, 2001.

Fields of Extremals and Sufficient Conditions for a Class of Variational Games

Dean Carlson
Mathematical Reviews
American Mathematical Society
Ann Arbor, MI, USA
dac@ams.org

Abstract

In this paper, using G. Leitmann's direct sufficiency method, extended for differential games, we present sufficient conditions for open loop strategies for a class of continuous time variational games. The result presented here is an extension of the classical Weierstrass sufficiency results for free problems in the calculus of variations.

Key words. Differential games, fields of extremals, calculus of variations.

AMS Subject Classifications. Primary 91A23; Secondary 49N70, 49K10.

1 Introduction

The study of calculus of variations is mainly focused on "single-player" games and has not concerned itself, for the most part, with the case of n-player games. The first exception to this case is the papers of C. F. Roos in the 1920's, which through a series of papers explored what we refer to here as variational games. In particular, we mention his 1925 paper in which he considers the mathematical theory of competition in the context of a dynamic economic growth model [13]. The culmination of these works appears to be his 1927 paper "Generalized Lagrange Problems in the Calculus of Variations" that appeared in the Transactions of the American Mathematical Society [14]. In this latter paper, he considers a general two-player game in the form of two coupled Lagrange problems in the calculus of variations. He presents a rather complete theory for what we now call a Nash equilibrium (some 25 years before Nash's theorem!) providing analogues of the classical necessary conditions, including the Euler-Lagrange equations, Weierstrass necessary conditions, Legendre conditions, as well as transversality conditions. The paper concludes with analogues of the classical sufficient conditions in the spirit of Weierstrass and Legendre for weak and strong local minimizers. Unfortunately, he includes no examples illustrating the application of his work. Moreover, with the advent of optimal control and differential games, much of this classical

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_2,

theory has been neglected. Indeed, the only reference I am aware of regarding this early work is a reference to the 1925 paper in the textbook by Kamien and Schwartz [9], although I am sure there are others.

In the work presented here, we consider a class of variational games with the idea of presenting a version of Weierstrass's sufficiency theorem and the classical field theory. Our approach is to present, via the direct method of G. Leitmann, a theorem which allows us to conclude that when a solution of the Euler-Lagrange equations can be embedded in a family of extremals then it is indeed an open loop Nash equilibrium. The direct method of G. Leitmann is a non-variational method which utilizes coordinate transformations to construct an "equivalent variational problem" for which the solution may be obtained easily (often by inspection). Once the equivalent problem is solved, the inverse coordinate transformation gives a solution to the original problem. The original formulation by Leitmann has been generalized to dynamic games and additionally has incorporated the notion of an equivalent problem arising in the work of C. Carathéodory for single-player games.

The plan of our paper is as follows. In Sec. 2 we introduce the class of variational games we are concerned with. Section 3 introduces the direct method presenting the fundamental lemma. In Sec. 4 we introduce the notion of a field of extremals and present our sufficient conditions for an open loop Nash equilibrium. Section 5 we give a presentation of the "classical presentation" of field theory illustrating the connections between the results of Sec. 4 with those of the classical variational theory. Finally, in Sec. 6 we conclude with two elementary examples illustrating our result.

2 The Class of Games Considered

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

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

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

Clearly, the strategies of the other players influences the decision of the j-th player and so each player is unable to minimize independently of the other players. As a consequence, the players seek to play a Nash equilibrium instead.

To introduce this concept we first introduce the following notation. For each fixed $j = 1, 2, \ldots, N$, $\mathbf{x} = (x_1, x_2, \ldots, x_N) \in \mathbb{R}^N$, and $y_j \in \mathbb{R}$ we use the notation,

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

With this notation we have the following definitions.

Definition 2.1. We say a function $\mathbf{x}(\cdot) = (x_1(\cdot), x_2(\cdot), \ldots, x_N(\cdot)) : [a, b] \to \mathbb{R}^N$ is admissible if it is continuous, has a piecewise continuous first derivative, satisfies the fixed initial and terminal conditions

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

and $(t, x_j(t), \dot{x}_j(t)) \in A_j$ for all $t \in [a, b]$.

Definition 2.2. Given an admissible function $\mathbf{x}(\cdot)$ we say a function $y_j(\cdot) : [a, b] \to \mathbb{R}$ is admissible for player j relative to $\mathbf{x}(\cdot)$ if the function $[\mathbf{x}^j, y_j](\cdot)$ is admissible.

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

Definition 2.3. An admissible function $\mathbf{x}^*(\cdot) : [a, b] \to \mathbb{R}^N$ is a Nash equilibrium if for each player $j = 1, 2, \ldots N$ and each function $y_j(\cdot) : [a, b] \to \mathbb{R}$ that is admissible for player j relative to $\mathbf{x}^*(\cdot)$ one has:

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

Remark 2.1. From the above definition it is clear that when all of the players "play" a Nash equilibrium, then each player's strategy is his best response to that of the other players. In other words, if player j applies any other strategy, than the equilibrium strategy, his cost functional will not decrease.

Remark 2.2. The above dynamic game clearly is not the most general structure one can imagine, even in a variational framework. In particular, the cost functionals are coupled only through their state variables and not through their strategies. While not the most general, one can argue that this form is general enough to cover many cases of interest since in a "real-world setting," an individual player will not know the strategies of the other players (see e.g., Dockner and Leitmann [8]).

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

Theorem 2.1. *If $\mathbf{x}^*(\cdot)$ is a Nash equilibrium, then the following system of Euler-Lagrange equations are satisfied:*

$$\frac{d}{dt}\frac{\partial L_j}{\partial \dot{x}_j}\left(t, \mathbf{x}^*(t), \dot{x}_j^*(t)\right) = \frac{\partial L_j}{\partial x_j}\left(t, \mathbf{x}^*(t), \dot{x}_j^*(t)\right), \quad t \in [a,b]. \tag{4}$$

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

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

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

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

3 Leitmann's Direct Method

In this section we briefly outline a coordinate transformation method originally developed for single-player games in [10] and [11] (see also [2]), and further extended to N-player games in [8] and [3] which will enable us to derive our results. In particular, in Leitmann [12] (see also Carlson and Leitmann [3]) we have the following theorem.

Lemma 3.1. *For $j = 1, 2, \dots, N$ let $x_j = z_j(t, \tilde{x}_j)$ be a transformation of class C^1 having a unique inverse $\tilde{x}_j = \tilde{z}_j(t, x_j)$ for all $t \in [a,b]$ such that there is a one-to-one correspondence $\mathbf{x}(t) \Leftrightarrow \tilde{\mathbf{x}}(t)$, for all admissible trajectories $\mathbf{x}(\cdot)$ satisfying the boundary conditions* (2) *and for all $\tilde{\mathbf{x}}(\cdot)$ satisfying:*

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

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

$$\begin{aligned} L_j(t, [\mathbf{x}^*(t)^j, x_j(t)], \dot{x}_j(t)) \;&-\; \tilde{L}_j(t, [\mathbf{x}^*(t)^j, \tilde{x}_j(t)], \dot{\tilde{x}}_j(t)) \\ &= \frac{d}{dt} G_j(t, \tilde{x}_j(t)) \end{aligned} \tag{5}$$

hold on $[a,b]$*. If* $\tilde{x}_j^*(\cdot)$ *yields an extremum of* $\tilde{I}_j([\mathbf{x}^{*j}(\cdot),\cdot])$ *with* $\tilde{x}_j^*(\cdot)$ *satisfying the transformed boundary conditions, then* $x_j^*(\cdot)$ *with* $x_j^*(t) = z_j(t, \tilde{x}^*(t))$ *yields an extremum for* $I_j([\mathbf{x}^{*j}(\cdot),\cdot])$ *with the boundary conditions* (2).

Moreover, the function $\mathbf{x}^*(\cdot)$ *is an open-loop Nash equilibrium for the variational game.*

Proof. Leitmann [12]. □

Remark 3.1. This result has been successfully applied to a number of examples having applications in mathematical economics and we refer the reader to the references. Additionally, this lemma has been extended to include various classes of control systems (e.g., affine in the strategies) [6], infinite horizon models [5], as well as multiple integral problems [4].

Two immediate and useful corollaries are as follows.

Corollary 3.1. *The existence of* $G_j(\cdot,\cdot)$*,* $j = 1, 2\ldots, N$*, in* (5) *imply that the following identities hold for* $(t, \tilde{x}_j) \in (a,b) \times \mathbb{R}$ *and* $\tilde{q}_j \in \mathbb{R}$ *for* $j = 1, 2, \ldots, N$*:*

$$L_j\Big(t, [\mathbf{x}^{*j}(t), z_j(t,\tilde{x}_j)], \frac{\partial z_j(t,\tilde{x}_j)}{\partial t} + \frac{\partial z_j(t,\tilde{x}_j)}{\partial \tilde{x}_j}\tilde{q}_j\Big) - \tilde{L}_j(t, [\mathbf{x}^{*j}(t), \tilde{x}_j], \tilde{q}_j) \equiv \frac{\partial G_j(t,\tilde{x}_j)}{\partial t} + \frac{\partial G_j(t,\tilde{x}_j)}{\partial \tilde{x}_j}\tilde{q}_j. \qquad (6)$$

Corollary 3.2. *For each* $j = 1, 2, \ldots, N$ *the left-hand side of the identity,* (6) *is linear in* $\tilde{q}_j$*, that is, it is of the form,*

$$\theta_j(t,\tilde{x}_j) + \psi_j(t,\tilde{x}_j)\tilde{q}_j$$

and,

$$\frac{\partial G_j(t,\tilde{x}_j)}{\partial t} = \theta_j(t,\tilde{x}_j) \quad \textit{and} \quad \frac{\partial G_j(t,\tilde{x}_j)}{\partial \tilde{x}_j} = \psi(t,\tilde{x}_j)$$

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

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

4 A Direct Method Proof of a Classical Sufficiency Condition

In this section we demonstrate how the direct method described above can be used to extend a classical result from the calculus of variations to the class of variational games considered here. Moreover, the approach presented here further provides an elementary proof of this result in the single-player case. We begin by extending the notion of a field of extremals beginning with the following definition.

Definition 4.1. For $j = 1, 2, \ldots N$, let $\xi_j(\cdot,\cdot) : [a, b] \times \mathbb{R} \to \mathbb{R}$ be a set of twice continuously differentiable functions and put $\xi(\cdot,\cdot) = (\xi_1(\cdot,\cdot), \xi_2(\cdot,\cdot), \ldots, \xi_N(\cdot,\cdot)) : [a, b] \times \mathbb{R}^N \to \mathbb{R}^N$. We say that $\xi(\cdot,\cdot)$ is a family of extremals for the differential game if for each constant vector $\beta = (\beta_1, \beta_2, \ldots, \beta_N) \in \mathbb{R}^N$ one has that the functions $t \to \xi_j(t, \beta_j)$ satisfies the Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L_j}{\partial \dot{x}_j}(t, \xi(t,\beta), \frac{\partial \xi_j(t,\beta_j)}{\partial t}) = \frac{\partial L_j}{\partial x_j}(t, \xi(t,\beta), \frac{\partial \xi_j(t,\beta_j)}{\partial t}), \quad t \in [a, b], \quad (7)$$

and that $(t, \xi_j(t, \beta_j), \frac{\partial \xi_j(t,\beta_j)}{\partial t}) \in A_j$ for $t \in [a, b]$ and $j = 1, 2 \ldots N$.

When given a game to solve, one naturally makes an effort to solve the corresponding necessary conditions to obtain candidates of optimality. For the case considered here, these necessary conditions are the system of Euler-Lagrange equations (4). Consequently, suppose that we have a candidate for a Nash equilibrium, say $\mathbf{x}^*(\cdot)$ that satisfies the Euler-Lagrange equations as well as the fixed end conditions (2). The following theorem gives additional conditions that allow us to determine when this candidate is an open-loop Nash equilibrium

Theorem 4.1. *In addition to the hypotheses given in Sec. 1, assume that for each $j = 1, 2, \ldots N$ the functions $p_j \to L_j(t, \mathbf{x}, p_j)$ are convex on $A_j(t, \mathbf{x}) = \{p_j \in \mathbb{R} : (\mathbf{x}, p_j) \in A_j(t)\}$. Furthermore, let $\mathbf{x}^*(\cdot)$ be an admissible function for the dynamic game which satisfies the Euler-Lagrange equation* (4) *and assume that there exists a family of extremals, $\xi(\cdot,\cdot)$ having the following properties.*

(1) The transformations $x_j = \xi_j(t, \beta_j)$ are of class C^2 with a unique inverse $\beta_j = \tilde{\xi}_j(t, x_j)$.
(2) There exists a constant vector $\beta^* \in \mathbb{R}^N$ such that $x_j^*(t) = \xi_j(t, \beta_j^*)$ for $t \in [a, b]$.
(3) The partial derivatives $\frac{\partial}{\partial \beta_j}\xi_j(t, \beta_j)$ are nonzero for all (t, β_j).

Then $\mathbf{x}^*(\cdot)$ is an open-loop Nash equilibrium for the dynamic game.

Remark 4.1. To put the above theorem into a classical context, in the one-player game the family of extremals, assumed to exist in the above theorem, is referred to as a “field of extremals” in the classic book by Bolza [1] and in this case we have a direct comparison to the classical sufficiency condition due to Weierstrass and Hilbert. In the next section we will explore this relationship further by presenting these results in a more classical manner.

Remark 4.2. A family of extremals satisfying the conditions indicated above defines a one-to-one correspondence between the admissible trajectories of the dynamic game considered, say $\mathbf{x}(\cdot)$, and a set of functions $\tilde{\mathbf{x}}(\cdot)$ satisfying the boundary conditions:

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

To see this we notice that if $\mathbf{x}(\cdot)$ is admissible for the original problem then we have the function $\tilde{\mathbf{x}}(\cdot) : [a,b] \to \mathbb{R}^N$ defined component-wise by $\tilde{x}_j(t) = \tilde{\xi}_j(t, x_j(t))$, for $j = 1, 2, \ldots N$, satisfies the fixed-end conditions

$$\tilde{x}_j(a) = \tilde{\xi}_j(a, x_{aj}) \quad \text{and} \quad \tilde{x}_j(b) = \tilde{\xi}_j(b, x_{bj}).$$

Further, since $\beta_j^* = \xi(t, x_j^*(t))$ for all $t \in [a,b]$, it follows immediately that $\beta_j^* = \tilde{\xi}(a, x_{aj}) = \tilde{\xi}(b, x_{bj})$ as desired.

Proof of Theorem 4.1. To begin, we note that by Taylor's theorem, for each $t \in [a,b]$, $\beta \in \mathbb{R}^N$, $\tilde{\mathbf{p}} = (\tilde{p}_1, \tilde{p}_2, \ldots \tilde{p}_N) \in \mathbb{R}^N$ and $j = 1, 2, \ldots, N$ there exists $\gamma_j(t, \beta, \tilde{p}_j) \in [0,1]$ such that:

$$\begin{aligned} & L_j\left(t, \xi(t,\beta), \frac{\partial \xi_j}{\partial t}(t,\beta_j) + \frac{\partial \xi_j}{\partial \beta_j}(t,\beta_j)\tilde{p}_j\right) = \\ & L_j\left(t, \xi(t,\beta), \frac{\partial \xi_j}{\partial t}(t,\beta_j)\right) + \frac{\partial L_j}{\partial p_j}\left(t, \xi(t,\beta), \frac{\partial \xi_j}{\partial t}(t,\beta_j)\right)\frac{\partial \xi_j}{\partial \beta_j}(t,\beta_j)\tilde{p}_j \\ & + \frac{1}{2}\frac{\partial^2 L_j}{\partial p_j^2}\bigg|_{\left(t,\xi(t,\beta),\frac{\partial \xi_j}{\partial t}(t,\beta_j)+\gamma_j(t,\beta,\tilde{p}_j)\frac{\partial \xi_j}{\partial \beta_j}(t,\beta_j)\tilde{p}_j\right)}\left(\frac{\partial \xi_j}{\partial \beta_j}(t,\beta_j)\right)^2 \tilde{p}_j{}^2. \end{aligned} \tag{8}$$

Now define $\tilde{L}_j(\cdot,\cdot,\cdot) : [a,b] \times \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}$ by the formula,

$$\tilde{L}_j(t,\beta,\tilde{p}_j) = \frac{1}{2}\frac{\partial^2 L_j}{\partial p_j^2}\left(\frac{\partial \xi_j}{\partial \beta_j}(t,\beta_j)\right)^2 \tilde{p}_j{}^2,$$

in which the second partial derivative of L_j in the above is evaluated at

$$\left(t, \xi(t,\beta), \frac{\partial \xi_j}{\partial t}(t,\beta_j) + \gamma_j(t,\beta,\tilde{p}_j)\frac{\partial \xi_j}{\partial \beta_j}(t,\beta_j)\tilde{p}_j\right),$$

and let $\theta(\cdot,\cdot) : [a,b] \times \mathbb{R}^N \to \mathbb{R}$ and $\psi(\cdot,\cdot) : [a,b] \times \mathbb{R}^N \to \mathbb{R}$ be given by

$$\theta(t,\beta) = L_j\left(t, \xi(t,\beta), \frac{\partial \xi_j}{\partial t}(t,\beta_j)\right)$$

and

$$\psi(t,\beta) = \frac{\partial}{\partial p_j}L_j\left(t, \xi(t,\beta), \frac{\partial \xi_j}{\partial t}(t,\beta_j)\right)\frac{\partial \xi_j}{\partial \beta_j}(t,\beta_j).$$

Observe that (8) may now be compactly written as:

$$\begin{aligned} L_j(t, \xi(t,\beta), \frac{\partial \xi_j(t,\beta_j)}{\partial t} + \frac{\partial \xi_j(t,\beta_j)}{\partial \beta_j}\tilde{p}_j) - \tilde{L}_j(t,\beta,\tilde{p}_j) & \\ = \ \theta(t, [\beta^j, \beta_j]) + \psi(t, [\beta^j, \beta_j])\tilde{p}_j. & \end{aligned} \tag{9}$$

To apply the direct method we must show that the right-hand side in (9) is exact, when viewed as a function of the variables (t, β_j). To do this it is sufficient to show that we have:

$$\frac{\partial}{\partial \beta_j}\theta(t,[\beta^j,\beta_j]) = \frac{\partial}{\partial t}\psi(t,[\beta^j,\beta_j]).$$

To this end we observe that since $\xi_j(\cdot,\cdot)$, $j = 1, 2, \ldots N$, is a family of extremals we have:

$$\begin{aligned}
\frac{\partial}{\partial \beta_j}\theta(t,[\beta^j,\beta_j]) &= \frac{\partial L_j}{\partial x_j}(t,\xi(t,\beta),\frac{\partial \xi_j}{\partial t}(t,\beta_j))\frac{\partial \xi_j}{\partial \beta_j}(t,\beta_j) \\
&\quad + \frac{\partial L_j}{\partial p_j}(t,\xi(t,\beta),\frac{\partial \xi_j}{\partial t}(t,\beta_j))\frac{\partial^2 \xi_j}{\partial \beta_j \partial t}(t,\beta_j) \\
&= \frac{\partial}{\partial t}\left[\frac{\partial L_j}{\partial p_j}(t,\xi(t,\beta),\frac{\partial \xi_j}{\partial t}(t,\beta_j))\right]\frac{\partial \xi_j}{\partial \beta_j}(t,\beta_j) \\
&\quad + \frac{\partial L_j}{\partial \xi_j}(t,\xi(t,\beta),\frac{\partial \xi_j}{\partial t}(t,\beta_j))\frac{\partial^2 \xi_j}{\partial \beta_j \partial t}(t,\beta_j) \\
&= \frac{\partial}{\partial t}\left[\frac{\partial}{\partial \xi_j}L_j(t,\xi(t,\beta),\frac{\partial \xi_j}{\partial t}(t,\beta_j))\right]\frac{\partial \xi_j}{\partial \beta_j}(t,\beta_j) \\
&\quad + \frac{\partial}{\partial \xi_j}L_j\left(t,\xi(t,\beta),\frac{\partial \xi_j}{\partial t}(t,\beta_j)\right)\frac{\partial^2 \xi_j}{\partial t \partial \beta_j}(t,\beta_j) \\
&= \frac{\partial}{\partial t}\psi(t,[\beta^j,\beta_j]),
\end{aligned}$$

where we have used the fact that the family of extremals is twice continuously differentiable. From this we can conclude that there exists a function $G_j(\cdot,\cdot)$: $[a, b] \times \mathbb{R}^N \to \mathbb{R}$ such that:

$$\begin{aligned}
L_j(t,\xi(t,\beta),\frac{\partial \xi_j}{\partial t}(t,\tilde{x}_j) &+ \frac{\partial \xi_j}{\partial x_j}(t,\tilde{x}_j)\tilde{p}_j) - \tilde{L}_j(t,\beta,\tilde{p}_j) \\
&= \frac{\partial G_j}{\partial t}(t,[\beta^j,\tilde{x}_j]) + \frac{\partial G_j}{\partial \tilde{x}_j}(t,[\beta^j,\tilde{x}_j])\tilde{p}_j.
\end{aligned}$$

Thus, for the trajectory $\mathbf{x}^*(\cdot) = \xi(\cdot,\beta^*)$ and any function $x_j(\cdot)$ that is admissible for player j relative to $\mathbf{x}^*(\cdot)$ for player j, we have:

$$L_j(t,[\mathbf{x}(t)^{*j},x_j(t)],\dot{x}_j(t)) - \tilde{L}_j(t,[\beta^{*j},\tilde{x}_j(t)],\dot{\tilde{x}}_j(t)) = \frac{d}{dt}G(t,[\beta^{*j},\tilde{x}_j(t)]),$$

in which $\tilde{x}_j(\cdot) : [a, b] \to \mathbb{R}$ is defined by $\tilde{x}_j(\cdot) = \tilde{\xi}_j(\cdot, x_j(\cdot))$. This is precisely Eq. (5) so that we can apply the direct method. To this end we consider the following N auxiliary problems of minimizing the functionals:

$$\tilde{I}_j(\tilde{x}_j(\cdot)) = \int_a^b \tilde{L}_j(t,[\beta^{*j},\tilde{x}_j(t)],\dot{\tilde{x}}_j(t))\,dt \tag{10}$$

subject to the fixed (constant) end conditions

$$\tilde{x}_j(a) = \tilde{x}_j(b) = \beta_j^*. \tag{11}$$

Observe that since L_j is convex in its last argument we have $(t, \tilde{x}_j, \tilde{p}_j) \to \tilde{L}_j(t, [\beta^{*j}, \tilde{x}_j], \tilde{p}_j)$ is non-negative and assumes the value zero whenever $\tilde{p}_j = 0$. This means that minimizers for each of the N auxiliary problems are given by $\tilde{x}_j^*(t) \equiv \beta_j^*$ for $j = 1, 2, \ldots N$. Thus an application of the direct method gives us that the trajectory $\mathbf{x}^*(\cdot)$ is an open-loop Nash equilibrium for the original dynamic game. $\square$.

5 The Classical Approach to Sufficient Conditions

As indicated earlier, Bolza [1] referred to a family of extremals satisfying the hypotheses of Theorem 2 as a field of extremals. In other treatments, the notion of a field of extremals (see e.g., Cesari [7]) is defined in terms of a "slope function." To see how this concept arises, suppose that we have a candidate for a Nash equilibrium, say $\mathbf{x}^*(\cdot) = (x_1^*(\cdot), \ldots, x_N^*(\cdot))$, and a family of extremals $\{\xi_j(\cdot,\cdot)\}_{j=1}^N$ satisfying the following properties.

(i) There is a vector $\beta^* = (\beta_1^*, \ldots \beta_N^*) \in \mathbb{R}^N$ such that for each $j = 1, 2, \ldots N$ and all for $t \in [a, b]$ one has $x_j^*(t) = \xi_j(t, \beta_j^*)$.

(ii) For each $j = 1, 2 \ldots N$, there exists a set S_j that is a subset of a set of the form $\{(t, x_j) : t \in [a, b],\ x_j^*(t) - k < x_j < x_j^*(t) + k,\ k \in \mathbb{R}_+\}$ such that the equation $x_j = \xi_j(t, \tilde{x}_j)$ implicitly defines a function $\tilde{x}_j : S_j \to \mathbb{R}$ having continuous first partial derivatives on S (i.e. $\tilde{x}_j = \tilde{\xi}_j(t, x_j)$).

(iii) For each $j = 1, 2, \ldots N$ the functions $\xi_j(\cdot,\cdot)$ are twice continuously differentiable in an interval of the form $[a, b] \times [A_j, B_j]$ such that β_j^* is an interior point of $[A_j, B_j]$.

For each j, we view the function $\xi_j(\cdot, \beta_j)$ as a set of curves that cover S_j with exactly one curve through each point of S_j.

Now define a family of slope functions $\pi_j : S_j \to \mathbb{R}$, $j = 1, 2, \ldots N$ by the equations:

$$\pi_j(t, x_j) = \frac{\partial \xi_j}{\partial t}(t, \tilde{\xi}_j(t, x_j)), \quad (t, x_j) \in S_j, \quad j = 1, 2, \ldots N. \tag{12}$$

Observe that at a point $(t, x_j) \in S_j$, $\pi_j(t, x_j)$ is a tangent vector of the unique curve $\xi_j(\cdot, \beta_j)$ that passes through the point (t, x_j). Hence, one has

$$\pi_j(t, \xi_j(t, \beta_j)) = \frac{\partial \xi_j}{\partial t}(t, \beta_j), \tag{13}$$

which, as a consequence of (iii), implies that π_j has continuous first partial derivatives. Differentiating (13) yields:

$$\frac{\partial \pi_j}{\partial t}(t, \xi_j(t, \beta_j)) + \frac{\partial \pi_j}{\partial x_j}(t, \xi_j(t, \beta_j)) \frac{\partial \xi_j}{\partial t}(t, \beta_j) = \frac{\partial^2 \xi_j}{\partial t^2}(t, \xi_j(t, \beta_j)) \tag{14}$$

which becomes, via (13),

$$\frac{\partial \pi_j}{\partial t}(t, x_j) + \pi_j(t, x_j)\frac{\partial \pi_j}{\partial x_j}(t, x_j) = \frac{\partial^2 \xi_j}{\partial t^2}(t, x_j). \tag{15}$$

Letting $\mathbf{S} = S_1 \times \cdots S_N$ and $\pi(\cdot,\cdot) = (\pi_1(\cdot,\cdot),\ldots \pi_N(\cdot,\cdot))$ (viewed as a function from $[a,b] \times \mathbf{S}$ into $\mathbb{R}^N$) we refer to the pair $(\mathbf{S}, \pi)$ as a field $\mathcal{F}$ about $\mathbf{x}^*(\cdot)$ and say that the trajectory $\mathbf{x}^*(\cdot)$ is embedded in the field $\mathcal{F}$. To continue this further, and to recapture Hilbert's theory we assume that the integrands L_j are smooth enough so that all the partial derivatives taken below exist and are continuous. Let $\mathbf{x}^*(\cdot)$ be a twice continuously differentiable function that is embedded in field $\mathcal{F}$ and consider the functionals, for $j = 1, 2, \ldots N$, $\mathcal{J}_j$ defined on $\mathcal{Y}_j = \{y \in C^2([a,b];\mathbb{R}) : \ (t, [\mathbf{x}^{*j}(t), y(t)], \dot{y}(t)) \in A_j, \ a \le t \le b\}$ by

$$\begin{aligned}
\mathcal{J}_j(y(\cdot)) &= \int_a^b \Big\{ L_j(t, [\mathbf{x}^*(t)^j, y(t)], \pi_j(t, y(t))) \\
&\qquad + [\dot{y}(t) - \pi_j(t, y(t))] \frac{\partial L_j}{\partial p_j}(t, [\mathbf{x}^*(t)^j, y(t)], \pi_j(t, y(t))) \Big\}\, dt \\
&= \int_\sigma \Big\{ L_j(t, [\mathbf{x}^*(t)^j, y], \pi_j(t, y)) \\
&\qquad - \pi_j(t, y) \frac{\partial L_j}{\partial p_j}(t, [\mathbf{x}^*(t)^j, y], \pi_j(t, y)) \Big\}\, dt \\
&\qquad + \frac{\partial L_j}{\partial p_j}(t, [\mathbf{x}^*(t)^j, y], \pi_j(t, y))\, dy
\end{aligned} \tag{16}$$

in which $\sigma = \{(t, y(t)) : a \le t \le b\} \subset \mathbb{R}^2$ is viewed as a curve in the plane and the last integral is interpreted as a line integral. From classical calculus, this line integral is path independent (i.e., depends only endpoints $(a, y(a)), (b, y(b)) \in \mathbb{R}^2$ if and only if one has:

$$\begin{aligned}
\frac{\partial}{\partial x_j}\Big\{ L_j(t, [\mathbf{x}^*(t)^j, y], \pi_j(t, y)) \ &- \ \pi_j(t, y) \frac{\partial L_j}{\partial p_j}(t, [\mathbf{x}^*(t)^j, y], \pi_j(t, y)) \Big\} \\
&= \ \frac{\partial}{\partial t}\Big\{ \frac{\partial L_j}{\partial p_j}(t, [\mathbf{x}^*(t)^j, y], \pi_j(t, y)) \Big\}.
\end{aligned} \tag{17}$$

Expanding the left side of the above gives us:

$$\begin{aligned}
\frac{\partial}{\partial x_j}\Big\{ L_j(t, [\mathbf{x}^*(t)^j, y], \pi_j(t, y)) \ &- \ \pi_j(t, y) \frac{\partial L_j}{\partial p_j}(t, [\mathbf{x}^*(t)^j, y], \pi_j(t, y)) \Big\} \\
&= \ \frac{\partial L_j}{\partial x_j} - \pi_j \left[\frac{\partial^2 L_j}{\partial x_j \partial p_j} + \frac{\partial^2 L_j}{\partial p_j^2} \frac{\partial \pi_j}{\partial x_j} \right]
\end{aligned} \tag{18}$$

in which $(t, [\mathbf{x}^*(t)^j, y], \pi_j(t, y))$ are the arguments of L_j and its partial derivatives (t, y) are the arguments π_j and its partial derivatives. Similarly, expanding the right side gives us:

$$\frac{\partial}{\partial t}\Big\{\frac{\partial L_j}{\partial p_j}(t,[\mathbf{x}^*(t)^j,y],\pi_j(t,y))\Big\} = \frac{\partial^2 L_j}{\partial t \partial p_j} + \sum_{i\neq j}\frac{\partial^2 L_j}{\partial x_i \partial p_j}\dot{x}_j^*(t) + \frac{\partial^2 L_j}{\partial p_j^2}\frac{\partial \pi_j}{\partial t}, \tag{19}$$

with the same convention regarding arguments as above. In view of the computations regarding the field, for each $(t,y) \in S_j$ there is a parameter β so that $y = \xi_j(t,\beta)$ and $\pi_j(t,y) = \frac{\partial \xi_j}{\partial t}(t,\beta)$. Moreover, we also have for each $i \neq j$ that $\dot{x}_i^*(t) = \pi_i(t, x*_i(t))$. Thus, using these facts, by equating (18) and (19) and arranging terms we arrive at the fact that (17) is equivalent to having,

$$\begin{aligned}\frac{\partial L_j}{\partial x_j} &= \frac{\partial^2 L_j}{\partial t \partial p_j} + \sum_{i\neq j}\frac{\partial^2 L_j}{\partial x_i \partial p_j}\frac{\partial \xi_i}{\partial t}(t,\beta_i^*) + \frac{\partial^2 L_j}{\partial x_j \partial p_j}\frac{\partial \xi_j}{\partial t}(t,\beta) \\ &\qquad + \frac{\partial^2 L_j}{\partial p_j^2}\frac{\partial^2 \xi_j}{\partial t^2}(t,\beta) \\ &= \frac{d}{dt}\frac{\partial L_j}{\partial p_j}(t,[\mathbf{x}^*(t)^j,\xi_j(t,\beta)],\frac{\partial \xi_j}{\partial t}(t,\beta)),\end{aligned} \tag{20}$$

which holds since we are assuming that $\xi(\cdot,\cdot)$ is a field of extremals. In the above computations we have used Eq. (15). Thus, in this setting we have shown the following version of the Hilbert invariant integral theorem.

Theorem 5.1. *Let $\mathbf{x}^*(\cdot)$ be any trajectory that can be embedded in a field $\mathcal{F}$, and suppose that for $j = 1, 2, \ldots$, fixed we have two piecewise smooth functions $y_1(\cdot)$ and $y_2(\cdot)$ such that the plane curves $(t, y_i(t)) \in S_j$ for $i = 1, 2$ have common endpoints. Then it is the case that $\mathcal{J}_j(y_1(\cdot)) = \mathcal{J}_j(y_2(\cdot))$.*

Now define, for a fixed trajectory $\mathbf{x}^*(\cdot)$ embedded in a field $\mathcal{F}$, the “Weierstrass excess functions” $E_j(\cdot,\cdot,\cdot,\cdot) : S_j \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by the formula:

$$\begin{aligned}E_j(t,y,q,p) = L_j(t,[\mathbf{x}^*(t)^j,y],p) - L_j(t,[\mathbf{x}^*(t)^j,y],q) \\ -(p-q)\frac{\partial L_j}{\partial p_j}(t,[\mathbf{x}^*(t)^j,y],q).\end{aligned} \tag{21}$$

Using this notation one can easily show the following result.

Theorem 5.2. *Let $\mathbf{x}^*(\cdot)$ be a smooth admissible trajectory embedded in a field $\mathcal{F} = (\mathbf{S}, \pi)$ and for each $j = 1, 2, \ldots, N$ let $y_j(\cdot)$ be any function defined on $[a, b]$ such that $(t, y_j(t)) \in S_j$ for all $t \in [a, b]$ and such that $y_j(a) = x_j^*(a)$ and $y_j(b) = x_j^*(b)$ (i.e., $y_j(\cdot)$ is admissible for player j relative to $\mathbf{x}^*(\cdot)$). Then one has:*

$$I_j([\mathbf{x}^*(\cdot)^j, y_j(\cdot)]) - I_j(\mathbf{x}^*(\cdot)) = \int_a^b E_j(t, y_j(t), \pi_j(t, y_j(t)), \dot{y}_j(t))\, dt.$$

Proof. We first observe that from the definition of the Weierstrass excess functions we immediately have for each $j = 1, 2, \ldots N$ that $\mathcal{J}_j(x_j^*(\cdot)) = I_j(\mathbf{x}^*(\cdot))$ since, $\dot{x}_j^*(\cdot) = \pi_j(\cdot, x_j^*(\cdot))$. Further, from the above theorem we also have that $\mathcal{J}_j(y_j(\cdot)) = \mathcal{J}_j(x_j^*(\cdot))$. Thus we have:

$$\begin{aligned}
I_j([\mathbf{x}^*(\cdot)^j\,,\, y_j(t)]) - I_j(\mathbf{x}^*(\cdot)) &= I_j([\mathbf{x}^*(\cdot)^j, y_j(t)]) - \mathcal{J}_j(x_j^*(\cdot)) \\
&= I_j([\mathbf{x}^*(\cdot)^j, y_j(t)]) - \mathcal{J}_j(y_j(\cdot)) \\
&= \int_a^b L_j(t, [\mathbf{x}^*(t)^j, y_j(t)], \dot{y}_j(t))\, dt - \\
&\quad \int_a^b \Big\{ L_j(t, [\mathbf{x}^*(t)^j, y(t)], \pi_j(t, y(t))) \\
&+ [\dot{y}(t) - \pi_j(t, y(t))] \frac{\partial L_j}{\partial p_j}(t, [\mathbf{x}^*(t)^j, y(t)], \pi_j(t, y(t))) \Big\}\, dt \\
&= \int_a^b E_j(t, y_j(t), \pi_j(t, y_j(t)), \dot{y}_j(t))\, dt,
\end{aligned}$$

as desired. □

From this last theorem we immediately have the following game-theoretic version of the classical Weierstrass-Hilbert sufficiency theorem.

Theorem 5.3. *If $\mathbf{x}^*(\cdot)$ is a smooth admissible trajectory of the variational game (1)–(2) which can be embedded in a field $\mathcal{F} = (\mathbf{S}, \mathbf{p})$, and if for each $j = 1, 2, \ldots, N$ and $y_j : [a, b] \to \mathbb{R}$ that is admissible for player j relative to $\mathbf{x}^*(\cdot)$ one has*

$$E_j(t, y_j(t), \pi_j(t, y_j(t)), \dot{y}_j(t)) \geq 0 \quad \textit{for all} \quad t \in [a, b],$$

then $I_j(\mathbf{x}^*(\cdot)) \leq I_j([\mathbf{x}^*(\cdot)^j, y_j(t)])$, which implies that $\mathbf{x}^*(\cdot)$ is an open-loop Nash equilibrium for the game (1)–(2).

Proof. The proof follows immediately from the previous theorem. □

Remark 5.1. To compare the Weierstrass-Hilbert sufficiency theorem to our result in the previous section we observe that Eq. (8) may be rewritten, for $\beta_i = \beta_i^*$ for $i \neq j$ as:

$$\begin{aligned}
&E_j\left(t, \xi_j(t, \beta_j), \frac{\partial \xi_j}{\partial t}(t, \beta_j), \frac{\partial \xi_j}{\partial t}(t, \beta_j) + \frac{\partial \xi_j}{\partial \beta_j}(t, \beta_j)\tilde{p}_j\right) \\
&= \frac{1}{2} \frac{\partial^2 L_j}{\partial p_j^2}\bigg|_{\left(t, \xi(t,\beta), \frac{\partial \xi_j}{\partial t}(t,\beta_j) + \gamma_j(t,\beta)\frac{\partial \xi_j}{\partial \beta_j}(t,\beta_j)\tilde{p}_j\right)} \left(\frac{\partial \xi_j}{\partial \beta_j}(t, \beta_j)\right)^2 \tilde{p}_j^2 \\
&\geq 0
\end{aligned}$$

due to our convexity assumption on the integrand $p_j \to L_j(t, \mathbf{x}, p_j)$.

Remark 5.2. Generally speaking, embedding a candidate for a Nash equilibrium, say $\mathbf{x}^*(\cdot)$, satisfying the Euler-Lagrange equations into a field $\mathcal{F}$ can only be guaranteed locally. As a consequence, the resulting sufficient condition only yields a local open-loop Nash equilibrium. In practice, for elementary examples it is possible to construct "global fields" which allows one to conclude that the candidate is a global Nash equilibrium. Sufficient conditions that insure the existence of an appropriate field for open-loop dynamic games are presently unavailable. We conjecture that theorems similar to those from the classical variational calculus can be appropriately modified to provide such conditions. One concept which will need to be addressed in this regard will be the notion of conjugate points. We leave these ideas for future research.

6 Examples

In this section we present several examples illustrating the above, theory beginning first with a one-player game (i.e., a calculus of variations problem).

6.1 Example

Consider the problem:

$$\text{minimize } J(x(\cdot)) = \int_a^b \dot{x}(t)^4 \, dt \tag{22}$$

over all piecewise smooth functions $x(\cdot) : [a, b] \to \mathbb{R}$ satisfying the fixed-end conditions:

$$x(a) = x_a \quad \text{and} \quad x(b) = x_b. \tag{23}$$

The Euler-Lagrange equation in this case is:

$$\frac{d}{dt} 4\dot{x}(t)^3 = 0$$

which gives rise to the extremals $x(t) = \alpha t + \beta$ for constants $\alpha, \beta \in \mathbb{R}$. The candidate for optimality is given as $x^*(t) = \alpha^* t + \beta^*$ in which α^* and β^* are chosen so that $x^*(\cdot)$ satisfies the end conditions (22). Explicitly, it is easy to check that $\alpha^* = (x_a - x_b)/(a - b)$ and that $\beta^* = x_a - \alpha^* a = b(x_b - x_a)/(a - b) = -b\alpha^*$. From this two-parameter family of extremals we must choose a one-parameter family. The one to choose is to let $\xi(t, \beta) = \alpha^* t + \beta$. Geometrically, the family of equations $x = \alpha^* t + \beta$ is a family of parallel lines that cover the tx-plane with a unique line (i.e., extremal) passing through each point of the plane. Furthermore, this gives rise to the desired one-to-one correspondence for the trajectories of the original problem (i.e., those that satisfy the end conditions (23)) and the piecewise trajectories $\tilde{x}(\cdot) : [a, b] \to \mathbb{R}$ that satisfy the fixed-end conditions:

$$\tilde{x}(a) = \beta^* \quad \text{and} \quad \tilde{x}(b) = \beta^*. \tag{24}$$

To see, this observe that if $x(\cdot)$ is a trajectory for the original problem then define $\tilde{x}(\cdot)$ by the formula

$$\tilde{x}(t) = x(t) - \alpha^* t, \quad t \in [a, b],$$

and observe that we have $\tilde{x}(a) = x_a - \alpha^* a = \beta^*$ and $\tilde{x}(b) = x_b - \alpha^* b = \beta^*$. Conversely, if $\tilde{x}(\cdot)$ satisfies (24) then we have that:

$$x(t) = \alpha^* t + \tilde{x}(t)$$

satisfies $x(a) = \alpha^* a + \tilde{x}(b) = \alpha^* a + \beta^* = x_a$ and $x(b) = \alpha^* b + \tilde{x}(b) = \alpha^* b + \beta^* = x_b$ as desired.

Now the Weierstrass sufficiency theorem tells us that $x^*(\cdot)$ is an optimal solution. What is new here is our proof of that fact, which is based on the direct method. To see how this works out in the example, observe that for $L(t, x, p) = p^4$ and $\xi(t, \beta) = \alpha^* t + \beta$ we have by Taylor's formula:

$$\begin{aligned} L(t, \xi(t,\beta), \xi_t(t,\beta) + \xi_\beta(t,\beta)\tilde{p}) &= L(t, z(t,\beta), \xi_t(t,\beta)) \\ &+ \frac{\partial L}{\partial z}(t, \xi(t,\beta), \xi_t(t,\beta))\xi_\beta(t,\beta)\tilde{p} \\ &+ \frac{1}{2}\frac{\partial^2 L}{\partial z^2}(t, \xi(t,\beta), \xi_t(t,\beta) + \gamma(t,\beta,\tilde{p})\xi_\beta(t,\beta)\tilde{p})\xi_\beta(t,\beta)^2\tilde{p}^2, \end{aligned} \tag{25}$$

where $\gamma(t, \beta, \tilde{p}) \in [0, 1]$. Here the subscripts $\xi_t(\cdot, \cdot)$ and $\xi_\beta(\cdot, \cdot)$ are partial derivatives. From the above we want to choose $\tilde{L}(\cdot, \cdot, \cdot)$ to be

$$\tilde{L}(t, \beta, \tilde{p}) = \frac{1}{2}\frac{\partial^2 L}{\partial p^2}(t, \xi(t,\beta), \xi_t(t,\beta) + \gamma(t,\beta,\tilde{p})\xi_\beta(t,\beta)\tilde{p})\xi_\beta(t,\beta)^2\tilde{p}^2$$

so that the auxiliary problem becomes:

$$\text{minimize } \tilde{J}(\tilde{x}(\cdot)) = \int_a^b \tilde{L}(t, \tilde{x}(t), \dot{\tilde{x}}(t))\, dt$$

over all trajectories $\tilde{x}(\cdot) : [a, b] \to \mathbb{R}$ satisfying the end conditions (24). Further from the above we have the following:

$$\begin{aligned} L(t, \xi(t,\beta), \xi_t(t,\beta) + \xi_\beta(t,\beta)\tilde{p}) \; - \; \tilde{L}(t, \beta, \tilde{p}) = \\ L(t, \xi(t,\beta), \xi_t(t,\beta)) + \frac{\partial L}{\partial p}(t, \xi(t,\beta), \xi_t(t,\beta))\xi_\beta(t,\beta)\tilde{p}. \end{aligned}$$

To apply the direct method we must show that there exists a function $G(t, \beta)$ such that

$$G_t(t, \beta) = L(t, \xi(t,\beta), \xi_t(t,\beta))$$

and

$$G_\beta(t, \beta) = \frac{\partial L}{\partial p}(t, \xi(t,\beta), \xi_t(t,\beta))\xi_\beta(t,\beta)$$

both hold. A necessary and sufficient condition for this to hold is that $G_{t\beta}(t,\beta) = G_{\beta t}(t,\beta)$ or equivalently that

$$\frac{\partial}{\partial \beta} L(t, \xi(t,\beta), \xi_t(t,\beta)) = \frac{\partial}{\partial t}\left(\frac{\partial L}{\partial p}(t, \xi(t,\beta), \xi_t(t,\beta))\xi_\beta(t,\beta)\right).$$

To see that this is the case we observe that (where we have suppressed the (t,β) arguments for brevity):

$$\begin{aligned}\frac{\partial}{\partial \beta} L(t, \xi(t,\beta), \xi_t(t,\beta)) &= \frac{\partial L}{\partial x}(t,\xi,\xi_t)\cdot \xi_\beta + \frac{\partial L}{\partial p}(t,\xi,\xi_t)\cdot \xi_{t\beta} \\ &= \frac{\partial}{\partial t}\left(\frac{\partial L}{\partial x}(t,\xi,\xi_t)\right)\xi_\beta + \frac{\partial L}{\partial p}(t,\xi,\xi_t)\cdot \xi_{t\beta} \\ &= \frac{\partial}{\partial t}\left(\frac{\partial L}{\partial \xi}(t, \xi(t,\beta), \xi_t(t,\beta))\xi_\beta(t,\beta)\right).\end{aligned}$$

The second line in the above is a consequence of the fact that $t \to \xi(t,\beta)$ solves the Euler-Lagrange equation for each β and the last line is just the product rule. This means we do indeed have the desired function $G(\cdot,\cdot)$ so that we can write

$$L(t, \xi(t,\beta), \xi_t(t,\beta) + \xi_\beta(t,\beta)\tilde{q}) - \tilde{L}(t,\beta,\tilde{p}) = G_t(t,\beta) + G_\beta(t,\beta)\tilde{p}$$

which gives the fundamental identity

$$\begin{aligned}&L(t, x(t), \dot{x}(t)) + \tilde{L}(t, \tilde{x}(t), \dot{\tilde{x}}(t)) \\ &\qquad = G_t(t,\tilde{x}(t)) + G_\beta(t,\tilde{x}(t))\dot{\tilde{x}}(t) = \frac{d}{dt}G(t,\tilde{x}(t)),\end{aligned}$$

for each pair of trajectories $\{x(\cdot), \tilde{x}(\cdot)\}$ related through $x(t) = \xi(t,\tilde{x}(t))$. For the simple example considered here we can explicitly determine all of these functions. Indeed, since $\xi_t = \alpha^*$, $\xi_\beta = 1$, $L_p = 4p^3$, and $L_{pp} = 12p^2$ we have that (25) becomes

$$(\alpha^* + \tilde{p})^4 = (\alpha^*)^4 + 4(\alpha^*)^3\tilde{p} + \frac{1}{2}12(\alpha^* + \gamma(t,\beta,\tilde{p})\tilde{p})^2\tilde{p}^2.$$

Expanding this last statement and collecting terms it is easy to see that $\gamma = \gamma(t,\beta,\tilde{p})$ satisfies the quadratic equation

$$6\tilde{p}\gamma^2 + 12\alpha^*\tilde{p}\gamma - \tilde{p}^2 - 4\alpha^*\tilde{p} = 0$$

from which γ is easily determined by the quadratic formula. In addition, we also must have $G(\cdot,\cdot)$ satisfy $G_t(t,\beta) = (\alpha^*)^4$ and $G_\beta(t,\beta) = 4(\alpha^*)^3$ so that from the first we must have $G(t,\beta) = (\alpha^*)^4 t + g(\beta)$ so that $g'(\beta) = G_\beta(t,\beta) = 4(\alpha^*)^3$ which implies that $g(\beta) = 4(\alpha^*)^3\beta + C$ where C is an arbitrary constant. Thus, we have $G(t, \beta) = (\alpha^*)^4 t + 4(\alpha^*)^3\beta + C$ for any constant C. Furthermore, we have that the auxiliary problem consists of minimizing

$$\tilde{I}(\tilde{x}(\cdot)) = \int_a^b 6(\alpha^* + \gamma(t, \tilde{x}(t), \dot{\tilde{x}}(t))\dot{\tilde{x}}(t))^2 \dot{\tilde{x}}(t)^2 \, dt$$

over all piecewise smooth trajectories $\tilde{x}(\cdot) : [a, b] \to \mathbb{R}$ satisfying the fixed-end conditions

$$\tilde{x}(a) = \beta^* = \tilde{x}(b).$$

Clearly, the integrand is non-negative for all choices of $\tilde{x}(\cdot)$ and, moreover, is identically zero when $\dot{\tilde{x}}(t) \equiv 0$. Thus, $\tilde{x}^*(t) \equiv \beta^*$ is an optimal solution and by the direct method we have $x^*(t) = \alpha^* t + \beta^*$ is an optimal solution of the original problem.

6.2 Example

In this example we consider a two-player game in which the objective of player $j = 1, 2$ is to maximize the objective functional

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

with the fixed-end conditions

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

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

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

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

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

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

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

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

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

which has the general solution

$$x_1(t) + x_2(t) = (\alpha_1 + \alpha_2)te^{2t} + Ce^{2t},$$

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

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

giving us

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

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

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

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

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

From this it is easy to see that if we choose

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

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

References

[1] O. Bolza, *Calculus of variations*, 3rd ed., Chelsea Publishing Company, New York, 1973.

[2] D.A. Carlson, *An observation on two methods of obtaining solutions to variational problems*, Journal of Optimization Theory and Applications **114** (2002), 345–362.

[3] D.A. Carlson and G. Leitmann, *An extension of the coordinate transformation method for open-loop Nash equilibria*, Journal of Optimization Theory and its Applications **123** (2004), no. 1, 27–47.

[4] ______, *A coordinate transformation method for the extremization of multiple integrals*, Journal Optimization Theory and Applications **127** (2005), 523–533.

[5] ______, *The direct method for a class of infinite horizon dynamic games*, Optimal Control and Dynamic Games, Applications in Finance, Management Science and Economics (New York) (Christophe Deissenberg and Richard F. Hartl, eds.), Springer, 2005.

[6] ______, *A direct method for open-loop dynamic games for affine control systems*, Dynamic Games: Theory and Applications (New York) (Alain Haurie and Georges Zaccour, eds.), Springer, 2005, pp. 37–55.

[7] L. Cesari, *Optimization-theory and applications: Problems with ordinary differential equations*, Applications of Applied Mathematics, vol. 17, Springer-Verlag, New York, 1983.

[8] E.J. Dockner and G. Leitmann, *Coordinate transformation and derivation of open-loop Nash equilibrium*, Journal of Optimization Theory and Applications **110** (2001), no. 1, 1–16.

[9] M.I. Kamien and N. L. Schwartz, *Dynamic optimization: The calculus of variations and optimal control in economics and management sciences*, second ed., Advanced Textbooks in Economics, vol. 31, North-Holland, Amsterdam, 1991.

[10] G. Leitmann, *A note on absolute extrema of certain integrals*, International Journal of Non-Linear Mechanics **2** (1967), 55–59.

[11] ______, *On a class of direct optimization problems*, Journal Optimization Theory and Applications **108** (2001), no. 3, 467–481.

[12] ______, *A direct method of optimization and its applications to a class of differential games*, Dynamics of Continuous, Discrete and Impulsive Systems, Series A **11** (2004), 191–204.

[13] C. F. Roos, *A mathematical theory of competition*, American Journal of Mathematics **46** (1925), 163–175.

[14] ______, *Generalized Lagrange problems in the calculus of variations*, Transactions of the American Mathematical Society **30** (1928), no. 2, 360–384.

Linear Quadratic Differential Games: An Overview

Jacob Engwerda
Department of Econometrics and O.R.
Tilburg University
P.O. Box: 90153, 5000 LE Tilburg, The Netherlands
engwerda@uvt.nl

Abstract

In this paper, we review some basic results on linear quadratic differential games. We consider both the cooperative and non-cooperative case. For the non-cooperative game we consider the open-loop and (linear) feedback information structure. Furthermore, the effect of adding uncertainty is considered. The overview is based on [9]. Readers interested in detailed proofs and additional results are referred to this book.

Key words. Linear-quadratic differential games, Nash equilibrium, affine systems, solvability conditions, Riccati equations.

AMS Subject Classifications. Primary 49N70; Secondary 49N90, 49N10, 49K15.

1 Introduction

Many situations in economics and management are characterized by multiple decision makers/players and enduring consequences of decisions. The theory which conceptualizes problems of this kind is dynamic games. Dynamic game theory tries to arrive at appropriate models describing the process. Depending on the specific problem this model sometimes can be used by an individual decision maker to optimize his performance. In other cases, it may serve as a starting point to introduce new communication lines which may help to improve upon the outcome of the current process. Furthermore it is possible by the introduction as “nature" as an additional player in games, which is trying to work against the other decision makers in a process, to analyze the robustness of strategies of players w.r.t. worst-case scenarios.

Examples of dynamic games in economics and management science can be found, e.g., in [8], [19], and [27].

In this paper, we consider a special class of dynamic games. We study games where the process can be modeled by a set of linear differential equations and the preferences are formalized by quadratic utility functions. The so-called linear quadratic differential games. These games are very popular in literature and a

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_3,

recent exposition (and additional references) of this theory can be found in [9]. The popularity of these games is caused, on the one hand, by practical considerations. To some extent these kinds of differential games are analytically and numerically solvable. On the other hand, this linear quadratic setting naturally appears if the decision makers' objective is to minimize the effect of a small perturbation of their optimally controlled nonlinear process. By solving a linear quadratic control problem, and using the optimal actions implied by this problem, players can avoid most of the additional cost incurred by this perturbation.

In a dynamic game, information available to the players at the time of their decisions plays an important role and, therefore, has to be specified before one can analyze these kind of games appropriately. We will distinguish two cases: the open-loop and feedback information case, respectively. In the open-loop information case, it is assumed that all players know just the initial state of the process and the model structure. More specifically, it is assumed that players simultaneously determine their actions for the whole planning horizon of the process before it starts. Next they submit their actions to some authority who then enforces these plans as binding commitments. So players cannot react on any deviations occurring during the evolution of the process. In the feedback information case, it is assumed that all players can observe at every point in time the current state of the process and determine their actions based on this observation.

In this paper, we will review some main results for linear quadratic differential games. Both the case when players cooperate in order to achieve their objectives and the case when players do not cooperate with each other are considered. The reason that players do not cooperate may be caused by individual motivations or physical reasons. We will take these reasons for granted. In the case when players do not cooperate, it seems reasonable that all players individually will try to play actions which are optimal for themselves. That is, for actions they cannot improve upon themselves. If there exists a set of actions such that none of the players has an incentive to deviate from his action (or stated otherwise, given the actions of the other players his choice of action is optimal), we call such a set of actions a Nash[1] equilibrium of the game. In general, a game may either have none, one, or more than one Nash equilibrium. This leads, on the one hand, to the question under which conditions these different situations will occur, and on the other hand, in case there is more than one equilibrium solution, whether there are additional motivations to prefer one equilibrium outcome to another.

As already indicated above, linear quadratic differential games have been studied a lot in the past. We will review some basic results and algorithms to compute equilibria, for the cooperative case as well as the open-loop and feedback information case. The outline of the rest of the paper is as follows. Sec. 2 considers the cooperative case, Sec. 3 the open-loop information case, and Sec. 4 the feedback information case. Section 5 recalls some results for the case when the system

[1]This after J.F. Nash who proved in a number of papers from 1950-1953 the existence of such equilibria.

is corrupted by noise. Finally, Sec. 6 reviews some extensions that can be found elsewhere in literature.

For ease of exposition we will just deal with the two-player case. Throughout this paper we will assume that each player has a (quadratic) cost function (1) he wants to minimize given by:

$$J_i(u_1,u_2)=\int_0^T \{x^T(t)Q_ix(t)+u_i^T(t)R_{ii}u_i(t)+u_j^T(t)R_{ij}u_j(t)\}dt + x^T(T)Q_{T,i}x(T),\ i=1,2,\ j\neq i. \quad (1)$$

Here the matrices Q_i, R_{ii}, and $Q_{T,i}$ are assumed to be symmetric and R_{ii} positive definite (denoted as $R_{ii} > 0$). Sometimes some additional positive definiteness assumptions are made w.r.t. the matrices Q_i and $Q_{T,i}$. In the minimization, a state variable $x(t)$ occurs. This is a dynamic variable that can be influenced by both players. Its dynamics are described by:

$$\dot{x}(t)=Ax(t)+B_1u_1+B_2u_2,\ x(0)=x_0, \quad (2)$$

where A and $B_i,\ i=1,2$, are constant matrices, and u_i is a vector of variables which can be manipulated by player i. The objectives are possibly conflicting. That is, a set of policies u_1 which is optimal for player one, may have rather negative effects on the evolution of the state variable x from another player's point of view.

2 The Cooperative Game

In this section we assume that players can communicate and can enter into binding agreements. Furthermore it is assumed that they cooperate in order to achieve their objectives. However, no side-payments take place. Moreover, it is assumed that every player has all information on the state dynamics and cost functions of his opponents and all players are able to implement their decisions. Concerning the strategies used by the players we assume that there are no restrictions. That is, every $u_i(.)$ may be chosen arbitrarily from a set $\mathcal{U}$ which is chosen such that we get a well-posed problem (in particular, it is chosen such that the differential Eq. (2) has a unique solution for every initial state).

By cooperation, in general, the cost one specific player incurs is not uniquely determined anymore. If all players decide, e.g., to use their control variables to reduce the cost of player 1 as much as possible, a different minimum is attained for player 1 then in the case when all players agree collectively to help a different player in minimizing his cost. So, depending on how the players choose to "divide" their contol efforts, a player incurs different "minima". Consequently, in general, each player is confronted with a whole set of possible outcomes from which somehow one outcome (which in general does not coincide with a player's overall lowest cost) is cooperatively selected. Now, if there are two strategies, γ_1 and γ_2, such that every player has a lower cost if strategy γ_1 is played, then it seems reasonable to assume that all players will prefer this strategy. We say that the solution induced by strategy γ_1 *dominates* in that case the solution induced by the strategy γ_2. So, dominance means that the outcome is better for all players. Proceeding in this line of thinking, it seems reasonable to consider only those cooperative outcomes which have the property that if a different strategy than the one corresponding with this cooperative outcome is chosen, then at least one of the players has higher costs. Or, stated differently, to consider only solutions that are such that they can not be improved upon by all players simultaneously. This motivates the concept of Pareto efficiency.

Definition 2.1. A set of actions $(\hat{u_1}, \hat{u_2})$ is called *Pareto efficient* if the set of inequalities $J_i(u_1, u_2) \leq J_i(\hat{u_1}, \hat{u}_2),\ i = 1, 2$, where at least one of the inequalities is strict, does not allow for any solution $(u_1, u_2) \in \mathcal{U}$. The corresponding point $(J_1(\hat{u_1}, \hat{u}_2), J_2(\hat{u_1}, \hat{u}_2)) \in \mathbb{R}^2$ is called a *Pareto solution*. The set of all Pareto solutions is called the *Pareto frontier*. □

A Pareto solution is therefore never dominated, and for that reason called an *undominated* solution. Typically, there is always more than one Pareto solution, because dominance is a property which generally does not provide a total ordering.

It turns out that if we assume $Q_i \geq 0,\ i = 1, 2$, in our cost functions (1) that there is a simple characterization for all Pareto solutions in our cooperative linear quadratic game. Below we will adopt the notation

$$u := (u_1, u_2) \text{ and } \mathcal{A} := \{\alpha = (\alpha_1, \alpha_2) | \alpha_i \geq 0 \text{ and } \sum_{i=1}^{2} \alpha_i = 1\}.$$

Theorem 2.2. Assume $Q_i \geq 0$. Let $\alpha_i > 0,\ i = 1, 2$, satisfy $\sum_{i=1}^{2} \alpha_i = 1$. If $\hat{u} \in \arg\min_{u \in \mathcal{U}} \{\sum_{i=1}^{2} \alpha_i J_i(u)\}$, then $\hat{u}$ is Pareto efficient.
Moreover, if $\mathcal{U}$ is convex, then for all Pareto efficient $\hat{u}$ there exist $\alpha \in \mathcal{A}$, such that $\hat{u} \in \arg\min_{u \in \mathcal{U}} \{\sum_{i=1}^{2} \alpha_i J_i(\gamma)\}$. □

Theorem 2.2 shows that to find all cooperative solutions for the linear quadratic game one has to solve a regular linear quadratic optimal control problem which depends on a parameter α. The existence of a solution for this problem is related to the existence of solutions of Riccati equations. In Lancaster et al. [22, Section 11.3], it is shown that if the parameters appearing in an algebraic Riccati equation are, e.g., differentiable functions of some parameter α (or, more general, depend analytically on a parameter α), and the maximal solution exists for all α in some open set V, then this maximal solution of the Riccati equation will be a differentiable function of this parameter α too on V (c.q., depend analytically on this parameter α too). Since in the linear quadratic case the parameters depend linearly on α, this implies that in the infinite horizon case the corresponding Pareto frontier will be a smooth function of α (provided the maximal solution exists for all α). A similar statement holds for the finite planning horizon case. In case for all $\alpha \in V$ the cooperative linear quadratic differential game has a solution or, equivalently, the corresponding Riccati differential equations have a solution, then (see e.g., Perko [26, Theorem 2.3.2] for a precise statement and proof) the solution of the Riccati differential equation is a differentiable function of α, since all parameters in this Riccati differential equation are differentiable functions of α.

The Pareto frontier does not always have to be a one-dimensional surface in $\mathbb{R}^2$, like in the above corollary. This is, e.g., already illustrated in the two-player case when both players have the same cost function. In that case the Pareto frontier reduces to a single point in $\mathbb{R}^2$.

Example 2.3. Consider the following differential game on government debt stabilization (see van Aarle et al. [1]). Assume that government debt accumulation, $\dot{d}(t)$, is the sum of interest payments on government debt, $rd(t)$, and primary fiscal deficits, $f(t)$, minus the seignorage i.e., the issue of base money) $m(t)$. So:

$$\dot{d}(t) = rd(t) + f(t) - m(t), d(0) = d_0.$$

Here, $d(t)$, $f(t)$, and $m(t)$ are expressed as fractions of GDP and r represents the rate of interest on outstanding government debt minus the growth rate of output. The interest rate $r > 0$ is assumed to be exogenous. Assume that fiscal and monetary policies are controlled by different institutions, the fiscal authority, and the monetary authority, respectively, which have different objectives.
The objective of the fiscal authority is to minimize a sum of time profiles of the primary fiscal deficit, base-money growth, and government debt: $J_1 = \int_0^\infty e^{-\delta t}\{f^2(t) + \eta m^2(t) + \lambda d^2(t)\}dt$. The parameters, η and λ express the relative priority attached to base-money growth and government debt by the fiscal authority.
The monetary authorities are assumed to choose the growth of base money such that a sum of time profiles of base-money growth and government debt is minimized. That is, $J_2 = \int_0^\infty e^{-\delta t}\{m^2(t) + \kappa d^2(t)\}dt$. Here, $1/\kappa$ can be interpreted as a measure for the conservatism of the central bank w.r.t. the money growth. Furthermore, all variables are normalized such that their targets are zero, and all parameters are positive.

Introducing $\tilde{d}(t) := e^{-\frac{1}{2}\delta t}d(t)$, $\tilde{m} := e^{-\frac{1}{2}\delta t}m(t)$ and $\tilde{f} := e^{-\frac{1}{2}\delta t}f(t)$ the above model can be rewritten as:

$$\dot{\tilde{d}}(t) = (r - \frac{1}{2}\delta)\tilde{d}(t) + \tilde{f}(t) - \tilde{m}(t),\ \tilde{d}(0) = d_0.$$

Where the cost functions of both players are:

$$J_1 = \int_0^\infty \{\tilde{f}^2(t) + \eta\tilde{m}^2(t) + \lambda\tilde{d}^2(t)\}dt \text{ and } J_2 = \int_0^\infty \{\tilde{m}^2(t) + \kappa\tilde{d}^2(t)\}dt.$$

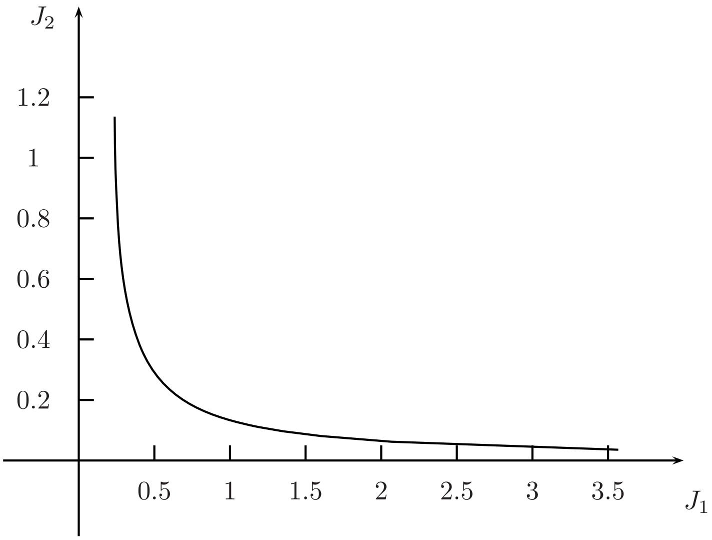

Figure 1: Pareto frontier Example 2.3 if $\eta = 0.1$; $\lambda = 0.6$; $\kappa = 0.5$; $r = 0.06$; and $\delta = 0.04$.

If both the monetary and fiscal authority agree to cooperate in order to reach their goals, by Theorem 2.2 the set of all Pareto solutions is obtained by considering the simultaneous minimization of

$$J_c(\alpha) := \alpha J_1 + (1-\alpha)J_2 = \int_0^\infty \{\alpha \tilde{f}^2(t) + \beta_1 \tilde{m}^2(t) + \beta_2 \tilde{d}^2(t)\}dt,$$

where $\beta_1 = 1 + \alpha(-1+\eta)$ and $\beta_2 = \kappa + \alpha(\lambda - \kappa)$. This cooperative game problem can be reformulated as the minimization of

$$J_c(\alpha) = \int_0^\infty \{\beta_2 \tilde{d}^2(t) + [\tilde{f}\ \tilde{m}] \begin{bmatrix} \alpha & 0 \\ 0 & \beta_1 \end{bmatrix} \begin{bmatrix} \tilde{f} \\ \tilde{m} \end{bmatrix}\}dt,$$

subject to $\dot{\tilde{d}}(t) = (r - \frac{1}{2}\delta)\tilde{d}(t) + [1\ -1] \begin{bmatrix} \tilde{f} \\ \tilde{m} \end{bmatrix}$, $\tilde{d}(0) = d_0$.
In Fig. 1 we plotted the set of Pareto solutions in case $\eta = 0.1$; $\lambda = 0.6$; $\kappa = 0.5$; $r = 0.06$; and $\delta = 0.04$. □

As Theorem 2.2 already indicates, in general, there are a lot of Pareto solutions. This raises the question which one is the "best". By considering this question we enter the arena of what is called bargaining theory.

This theory has its origin in two papers by Nash [24] and [25]. In these papers a bargaining problem is defined as a situation in which two (or more) individuals or organizations have to agree on the choice of one specific alternative from a set

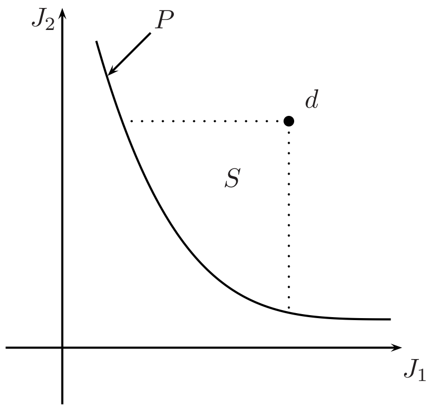

Figure 2: The bargaining game.

of alternatives available to them, while having conflicting interests over this set of alternatives. Nash proposes in [25] two different approaches to the bargaining problem, namely the *axiomatic* and *strategic* approach. The axiomatic approach lists a number of desirable properties the solution must have, called the *axioms*. The strategic approach on the other hand, sets out a particular bargaining procedure and asks what outcomes would result from rational behavior by the individual players.

So, bargaining theory deals with the situation in which players can realize—through cooperation—other (and better) outcomes than the one which becomes effective when they do not cooperate. This non-cooperative outcome is called the *threatpoint*. The question is to which outcome the players may possibly agree.

In Fig. 2 a typical bargaining game is sketched (see also Fig. 1). The inner part of the ellipsoid marks out the set of possible outcomes, the *feasible set* S, of the game. The point d is the threatpoint. The edge P is the set of individually rational Pareto-optimal outcomes.

We assume that if the agents unanimously agree on a point $x = (J_1, J_2) \in S$, they obtain x. Otherwise, they obtain d. This presupposes that each player can enforce the threatpoint, when he does not agree with a proposal. The outcome x the players will finally agree on is called the solution of the bargaining problem. Since the solution depends on the feasible set S as well as the threatpoint d, it will be written as $F(S, d)$. Notice that the difference for player i between the solution and the threatpoint, $J_i - d_i$, is the reduction in cost player i incurs by accepting the solution. In the sequel, we will call this difference the utility gain for player i. We will use the notation $J := (J_1, J_2)$ to denote a point in S and $x \succ y (x \prec y)$ to denote the *vector inequality*, i.e., $x_i > y_i (x_i < y_i),\ i = 1, 2$. In axiomatic bargaining theory a number of solutions have been proposed. In Thomson [28], a survey is given on this theory. We will present here the three most commonly used solutions: the Nash bargainig solution, the Kalai-Smorodinsky solution, and the

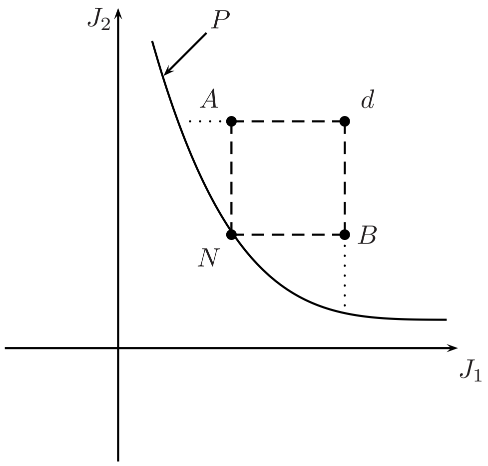

Figure 3: The Nash Bargaining solution $N(S,d)$.

Egalitarian solution.

The *Nash bargaining solution*, $N(S,d)$, selects the point of S at which the product of utility gains from d is maximal. That is:

$$N(S,d) = \arg\max_{J \in S} \prod_{i=1}^{N} (J_i - d_i), \text{ for } J \in S \text{ with } J \preceq d.$$

In Fig. 3 we sketched the N solution. Geometrically, the Nash Bargaining solution is the point on the edge of S (that is a part of the Pareto frontier) which yields the largest rectangle (N, A, B, d).

The Kalai-Smorodinsky solution, $K(S,d)$, sets utility gains from the threatpoint proportional to the player's most optimistic expectations. For each agent, the most optimistic expectation is defined as the lowest cost he can attain in the feasible set subject to the constraint that no agent incurs a cost higher than his coordinate of the threatpoint. Defining the *ideal point* as

$$I(S,d) := \max\{J_i \mid J \in S,\ J \preceq d\},$$

the *Kalai-Smorodinsky solution* is then

$$K(S,d) := \text{ maximal point of } S \text{ on the segment connecting } d \text{ to } I(S,d).$$

In Fig. 4 the Kalai-Smorodinsky solution is sketched for the two-player case. Geometrically, it is the intersection of the Pareto frontier P with the line which connects the threatpoint and the ideal point. The components of the ideal point are the minima each player can reach when the other player is fully altruistic under

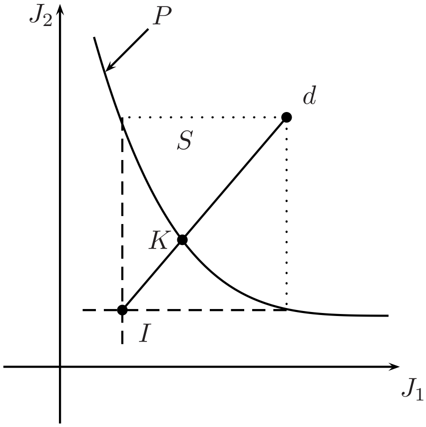

Figure 4: The Kalai-Smorodinsky solution $K(S, d)$.

cooperation.

Finally, the *Egalitarian solution*, $E(S, d)$, represents the idea that gains should be equal divided between the players. Formal, $E(S, d) :=$ maximal point in S for which

$$E_i(S, d) - d_i = E_j(S, d) - d_j,\ i, j = 1, \cdots, N.$$

Again, we sketched this solution for the two-player case. In Fig. 5 we observe that geometrically this Egalitarian solution is obtained as the intersection point of the $45°$-line through the threatpoint d with the Pareto frontier P.

Notice that, in particular, in contexts where interpersonal comparisons of utility is inappropriate or impossible, the first two bargaining solutions still make sense.

As already mentioned above, these bargaining solutions can be motivated using an "axiomatic approach". In this case, some people prefer to speak of an arbitration scheme instead of a bargaining game. An arbiter draws up the reasonable axioms and depending on these axioms, a solution results.

Algorithms to calculate the first two solutions numerically are outlined in [9]. The calculation of the Egalitarian solution requires the solution of one nonlinear constrained equations problem. The involved computer time to calculate this E-solution approximately equals that of calculating the N-solution.

3 The Open-Loop Game

In the rest of this paper we consider the case when players do not cooperate in order to realize their goals. In this section we will be dealing with the *open-loop*

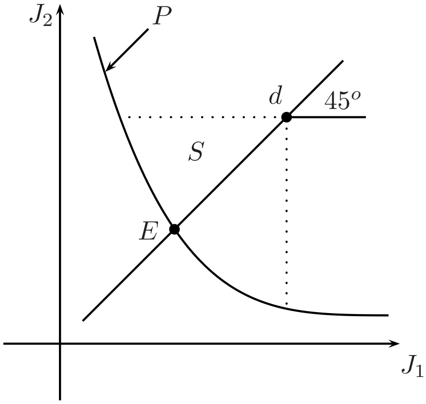

Figure 5: The Egalitarian solution $E(S, d)$.

information structure. That is, the case where every player knows at time $t \in [0, T]$ just the initial state x_0 and the model structure. This scenario can be interpreted as when the players simultaneously determine their actions and submit their actions to some authority who then enforces these plans as binding commitments. For the game (1),(2) we will study the set of Nash equilibria. A formal Nash equilibrium is defined as follows.

Definition 3.1. An admissible set of actions (u_1^*, u_2^*) is a *Nash equilibrium* for the game (1),(2), if for all admissible (u_1, u_2) the following inequalities hold: $J_1(u_1^*, u_2^*) \leq J_1(u_1, u_2^*)$ and $J_2(u_1^*, u_2^*) \leq J_2(u_1^*, u_2)$. □

Here admissibility is meant in the sense that $u_i(.)$ belongs to some restricted set, where this set depends on the information players have on the game, the set of strategies the players like to use to control the system, and the system (2) must have a unique solution.

So, the Nash equilibrium is defined such that it has the property that there is no incentive for any unilateral deviation by any one of the players. Notice that in general one cannot expect to have a unique Nash equilibrium. Moreover, it is easily verified that whenever a set of actions (u_1^*, u_2^*) is a Nash equilibrium for a game with cost functions $J_i,\ i = 1, 2$, these actions also constitute a Nash equilibrium for the game with cost functions $\alpha_i J_i,\ i = 1, 2$, for every choice of $\alpha_i > 0$.

Using the shorthand notation $S_i := B_i R_{ii}^{-1} B_i^T$, we have the following theorem in case the planning horizon T is finite.

Theorem 3.2. Consider matrix

$$M := \begin{bmatrix} A & -S_1 & -S_2 \\ -Q_1 & -A^T & 0 \\ -Q_2 & 0 & -A^T \end{bmatrix}. \tag{3}$$

Assume that the two Riccati differential equations,

$$\dot{K}_i(t) = -A^T K_i(t) - K_i(t)A + K_i(t)S_iK_i(t) - Q_i,\ K_i(T) = Q_{iT},\ i = 1,2, \tag{4}$$

have a symmetric solution $K_i(.)$ on $[0,T]$.
Then, the two-player linear quadratic differential game (1)–(2) has an open-loop Nash equilibrium for every initial state x_0 if and only if matrix

$$H(T) := [I\ 0\ 0]e^{-MT}\begin{bmatrix} I \\ Q_{1T} \\ Q_{2T} \end{bmatrix} \tag{5}$$

is invertible. Moreover, if for every x_0 there exists an open-loop Nash equilibrium, then the solution is unique. The unique equilibrium actions as well as the associated state trajectory can be calculated from the linear two-point boundary value problem

$$\dot{y}(t) = My(t),\ \text{with}\ Py(0) + Qy(T) = [x_0^T\ 0\ 0]^T. \tag{6}$$

Here $P = \begin{bmatrix} I\,0\,0 \\ 0\,0\,0 \\ 0\,0\,0 \end{bmatrix}$ and $Q = \begin{bmatrix} 0 & 0 & 0 \\ -Q_{1T} & I & 0 \\ -Q_{2T} & 0 & I \end{bmatrix}$.
Denoting $[y_0^T(t),\ y_1^T(t),\ y_2^T(t)]^T := y(t)$, with $y_0 \in \mathbb{R}^n$, and $y_i \in \mathbb{R}^{m_i},\ i = 1,2$, the state and equilibrium actions are:

$$x(t) = y_0(t)\ \text{and}\ u_i(t) = -R_{ii}^{-1}B_i^T y_i(t),\ i = 1,2,\ \text{respectively.} \qquad \square$$

Assumption (4) is equivalent to the statement that for both players a with this game problem associated linear quadratic control problem should be solvable on $[0,T]$. That is, the optimal control problem that arises in case the action of his opponent(s) would be known must be solvable for each player. Generically, one may expect that, if there exists an open-loop Nash equilibrium, the set of Riccati differential Eqs. (4) will have a solution on the closed interval $[0,T]$ (see [9, p.269]).

Next consider the set of coupled asymmetric Riccati-type differential equations:

$$\dot{P}_1 = -A^T P_1 - P_1A - Q_1 + P_1S_1P_1 + P_1S_2P_2;\ P_1(T) = Q_{1T} \tag{7}$$

$$\dot{P}_2 = -A^T P_2 - P_2A - Q_2 + P_2S_2P_2 + P_2S_1P_1;\ P_2(T) = Q_{2T}. \tag{8}$$

Existence of a Nash equilibrium is closely related to the existenc of a solution of the above set of coupled Riccati differential equations. The next result holds.

Theorem 3.3. The following statements are equivalent:

a) For all $T \in [0,t_1)$ there exists for all x_0 a unique open-loop Nash equilibrium for the two-player linear quadratic differential game (1)–(2).

b) The next two conditions hold on $[0,t_1)$.

1. $H(t)$ is invertible for all $t \in [0,t_1)$.

2. The two Riccati differential Eqs. (4) have a solution $K_i(0,T)$ for all $T \in [0,t_1)$.

c) The next two conditions hold on $[0,t_1)$.

1. The set of coupled Riccati differential Eqs. (7),(8) has a solution $(P_1(0,T), P_2(0,T))$ for all $T \in [0,t_1)$.
2. The two Riccati differential Eqs. (4) have a solution $K_i(0,T)$ for all $T \in [0,t_1)$.

Moreover, if either one of the above conditions is satisfied the equilibrium is unique. The set of equilibrium actions is in that case given by:

$$u_i^*(t) = -R_{ii}^{-1} B_i^T P_i(t)\Phi(t,0)x_0,\ i = 1,2.$$

Here, $\Phi(t,0)$ satisfies the transition equation:

$$\dot{\Phi}(t,0) = (A - S_1P_1 - S_2P_2)\Phi(t,0);\ \Phi(t,t) = I. \qquad \square$$

Note that there are situations where the set of Riccati differential Eqs. (7),(8) does not have a solution, while there does exist an open-loop Nash equilibrium for the game. With

$$P := \begin{bmatrix} P_1 \\ P_2 \end{bmatrix};\ D := \begin{bmatrix} A^T & 0 \\ 0 & A^T \end{bmatrix};\ S := [S_1\ S_2];\ \text{and } Q := \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix},$$

the set of coupled Riccati Eqs. (7),(8) can be rewritten as the non-symmetric matrix Riccati differential equation:

$$\dot{P} = -DP - PA + PSP - Q;\ P^T(T) = [Q_{1T},\ Q_{2T}].$$

The solution of such a Riccati differential equation can be obtained by solving a set of linear differential equations. In particular, if this linear system of differential Eqs. (9), below, can be analytically solved we also obtain an analytic solution for (P_1, P_2) (see e.g. [2]). Due to this relationship it is possible to compute solutions of (7),(8) in an efficient reliable way using standard computer software packages like, e.g., MATLAB. We have the following result.

Proposition 3.4. The set of coupled Riccati differential Eqs. (7),(8) has a solution on $[0,T]$ if and only if the set of linear differential equations

$$\begin{bmatrix} \dot{U}(t) \\ \dot{V}_1(t) \\ \dot{V}_2(t) \end{bmatrix} = M \begin{bmatrix} U(t) \\ V_1(t) \\ V_2(t) \end{bmatrix};\ \begin{bmatrix} U(T) \\ V_1(T) \\ V_2(T) \end{bmatrix} = \begin{bmatrix} I \\ Q_{1T} \\ Q_{2T} \end{bmatrix} \qquad (9)$$

has a solution on $[0,T]$, with $U(.)$ non-singular.

Moreover, if (9) has an appropriate solution $(U(.), V_1(.), V_2(.))$, the solution of (7),(8) is obtained as $P_i(t) := V_i(t)U^{-1}(t),\ i = 1,2.$ $\square$

Next we consider the infinite planning horizon case. That is, the case that the performance criterion player $i = 1, 2$, likes to minimize is:

$$\lim_{T\to\infty} J_i(x_0, u_1, u_2, T) \tag{10}$$

where

$$J_i = \int_0^T \{x^T(t)Q_i x(t) + u_i^T(t)R_{ii}u_i(t) + u_j^T(t)R_{ij}u_j(t)\}dt,$$

subject to the familiar dynamic state Eq. (2). We assume that the matrix pairs $(A, B_i),\ i = 1, 2$, are stabilizable. So, in principle, each player is capable to stabilize the system on his own.

Since we only like to consider those outcomes of the game that yield a finite cost to both players and the players are assumed to have a common interest in stabilizing the system, we restrict ourselves to functions belonging to the set:

$$\mathcal{U}_s(x_0) = \left\{u \in L_2 \mid J_i(x_0, u) \text{ exists in } \mathbb{R} \cup \{-\infty, \infty\}, \lim_{t\to\infty} x(t) = 0\right\}.$$

A similar remark as for the finite planning horizon case applies here. That is, the restriction to this set of control functions requires some form of communication between the players.
To find conditions under which this game has a unique equilibrium the next algebraic Riccati equations play a fundamental role:

$$0 = A^T P_1 + P_1 A + Q_1 - P_1 S_1 P_1 - P_1 S_2 P_2, \tag{11}$$
$$0 = A^T P_2 + P_2 A + Q_2 - P_2 S_2 P_2 - P_2 S_1 P_1; \tag{12}$$

and

$$0 = A^T K_i + K_i A - K_i S_i K_i + Q_i,\ i = 1, 2. \tag{13}$$

Notice that the Eqs. (13) are just the ordinary Riccati equations that appear in the regulator control problem.

Definition 3.5. A solution (P_1, P_2) of the set of algebraic Riccati Eqs. (11),(12) is called:

a. *stabilizing*, if $\sigma(A - S_1P_1 - S_2P_2) \subset \mathbb{C}^-$;

b. *strongly stabilizing* if

i. it is a stabilizing solution, and

ii.
$$\sigma\left(\begin{bmatrix} -A^T + P_1S_1 & P_1S_2 \\ P_2S_1 & -A^T + P_2S_2 \end{bmatrix}\right) \subset \mathbb{C}_0^+. \tag{14}$$

Here $\mathbb{C}^-$ denotes the left open half of the complex plane and $\mathbb{C}_0^+$ its complement. □

The next theorem gives an answer under which conditions the set of algebraic Riccati equations has a unique strongly stabilizing solution. One of these conditions is that a certain subspace should be a graph subspace. A subspace $V = \mathrm{Im}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$, with $X_i \in \mathbb{R}^{n\times n}$, which has the additional property that X_1 is invertible is called a *graph subspace* (since it can be "visualized" as the graph of the map: $x \to X_2X_1^{-1}x$).

Theorem 3.6.

1. The set of algebraic Riccati Eqs. (11),(12) has a strongly stabilizing solution (P_1, P_2) if and only if matrix M has an n-dimensional stable graph subspace and M has $2n$ eigenvalues (counting algebraic multiplicities) in $\mathbb{C}_0^+$.
2. If the set of algebraic Riccati Eqs. (11),(12) has a strongly stabilizing solution, then it is unique.

Then we have the following main result (see also [11]).

Theorem 3.7. The linear quadratic differential game (2),(10) has a unique open-loop Nash equilibrium for every initial state if and only if:

1. The set of coupled algebraic Riccati Eqs. (11),(12) has a strongly stabilizing solution, and
2. the two algebraic Riccati Eqs. (13) have a stabilizing solution.

Moreover, the unique equilibrium actions are given by:

$$u_i^*(t) = -R_{ii}^{-1}B_i^TP_i\Phi(t,0)x_0,\ i = 1,2. \tag{15}$$

Here, $\Phi(t,0)$ satisfies the transition equation:

$$\dot{\Phi}(t,0) = (A - S_1P_1 - S_2P_2)\Phi(t,0);\ \Phi(t,t) = I. \qquad \square$$

To calculate the unique equilibrium numerically one can use the next algorithm.

Algorithm 3.8.

1: Calculate the eigenstructure of $H_i := \begin{bmatrix} A & -S_i \\ -Q_i & -A^T \end{bmatrix}$.
If $H_i,\ i = 1,2$, has an n-dimensional stable graph subspace, then proceed. Otherwise go to *5*.

2: Calculate matrix $M := \begin{bmatrix} A & -S_1 & -S_2 \\ -Q_1 & -A^T & 0 \\ -Q_2 & 0 & -A^T \end{bmatrix}$. Next calculate the spectrum of M. If the number of eigenvalues having a strict negative real part (counted with algebraic multiplicities) differs from n, go to *5*.

3: Calculate the n-dimensional M-invariant subspace for which Re $\lambda < 0$ for all $\lambda \in \sigma(M)$. That is, calculate the subspace consisting of the union of all (generalized) eigenspaces associated with each of these eigenvalues. Let $\mathcal{P}$ denote this invariant subspace. Calculate $3n \times n$ matrices $X,\ Y$, and Z such

that $\text{Im}\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathcal{P}$.
Verify whether matrix X is invertible. If X is not invertible go to *5*.

4: Denote $P_1 := YX^{-1}$ and $P_2 := ZX^{-1}$. Then
$u_i^*(t) := -R_{ii}^{-1}B_i^TP_ie^{A_{cl}t}x_0$
is the unique open-loop Nash equilibrium strategy for every initial state of the game. Here, $A_{cl} := A - S_1P_1 - S_2P_2$. The spectrum of the corresponding closed-loop matrix A_{cl} equals $\sigma(M|_{\mathcal{P}})$. The involved cost for player i is $x_0^TM_ix_0$, where M_i is the unique solution of the Lyapunov equation:
$A_{cl}^TM_i + M_iA_{cl} + Q_i + P_i^TS_iP_i + P_jB_j^TR_{jj}^{-1}R_{ij}R_{jj}^{-1}B_jP_j = 0.$

5: End of algorithm. □

Step 1 in the algorithm verifies whether the two algebraic Riccati Eqs. (13) have a stabilizing solution. Steps 2 and 3 verify whether matrix M has an n-dimensional stable graph subspace. Finally, Step 4 determines the unique equilibrium.

The next example illustrates the algorithm.

Example 3.9.
1. Consider the system $\dot{x}(t) = -2x(t) + u_1(t) + u_2(t),\ x(0) = x_0;$
and cost functions

$$J_1 = \int_0^\infty \{x^2(t) + u_1^2(t)\}dt \text{ and } J_2 = \int_0^\infty \{4x^2(t) + u_2^2(t)\}dt.$$

Then, $M = \begin{bmatrix} -2 & -1 & -1 \\ -1 & 2 & 0 \\ -4 & 0 & 2 \end{bmatrix}$.
The eigenvalues of M are $\{-3, 2, 3\}$. An eigenvector corresponding with the eigenvalue -3 is $[5,\ 1,\ 4]^T$.
So, according to Theorem 3.6 item 1, the with this game corresponding set of algebraic Riccati Eqs. (11),(12) has a strongly stabilizing solution. Furthermore, since $q_i > 0,\ i = 1, 2$, the two algebraic Riccati Eqs. (13) have a stabilizing solution. Consequently, this game has a unique open-loop Nash equilibrium for every initial state x_0.
The equilibrium actions are:

$$u_1^*(t) = \frac{-1}{5}x^*(t),\ u_2^*(t) = \frac{-4}{5}x^*(t),\ \text{where } \dot{x}^*(t) = -3x^*(t).$$

The corresponding cost are $J_1 = \frac{13}{75}x_0^2,\ J_2 = \frac{58}{75}x_0^2$.
2. Reconsider the game from item 1, but with the system dynamics replaced by $\dot{x}(t) = 2x(t) + u_1(t) + u_2(t),\ x(0) = x_0$.
The with this game corresponding matrix M has the eigenvalues $\{-3, -2, 3\}$.

Since M has two stable eigenvalues, it follows from Theorem 3.6 item 1 that the with this game corresponding set of algebraic Riccati Eqs. (11),(12) does not have a strongly stabilizing solution. So, see Theorem 3.7, the game does not have for every initial state a unique open-loop Nash equilibrium.
It can be shown (see [9]) that in this example for every initial state there are actually an infinite number of open-loop Nash equilibria. □

4 The Feedback Game

It is often argued that weak time consistency is a minimal requirement for the credibility of an equilibrium solution. That is, if the advertised equilibrium action of, say, player one is not weakly time consistent, player one would have an incentive to deviate from this action during the course of the game. For the other players, knowing this, it is therefore rational to incorporate this defection of player one into their own actions, which would lead to a different equilibrium solution. On the other hand, the property that the equilibrium solution does not have to be adapted by the players during the course of the game, although the system evolutes not completely as expected beforehand, is generally experienced as a very nice property. Since the open-loop Nash equilibria in general do not have this property, the question arises whether there exist strategy spaces Γ_i such that if we look for Nash equilibria within these spaces, the equilibrium solutions do satisfy this strong time consistency property.

Since the system we consider is linear, it is often argued that the equilibrium actions should be a linear function of the state too. This argument implies that we should consider either a refinement of the feedback Nash equilibrium concept, or, strategy spaces that only contain functions of the above mentioned type. The first option amounts to consider only those feedback Nash equilibria which permit a linear feedback synthesis as being relevant. For the second option, one has to consider the strategy spaces defined, for $i = 1, 2$, by $\Gamma_i^{lfb} :=$

$$\{u_i(0,T) | u_i(t) = F_i(t)x(t) \text{ where } F_i(.) \text{ is a piecewise continuous function}\},$$

and consider Nash equilibrium actions (u_1^*, u_2^*) within the strategy space $\Gamma_1^{lfb} \times \Gamma_2^{lfb}$.

It turns out that both equilibrium concepts yield the same characterization of these equilibria for the linear quadratic differential game, which will be presented below in Theorem 4.5. Therefore, we will define just one equilibrium concept here.

Definition 4.1. The set of control actions $u_i^*(t) = F_i^*(t)x(t)$ constitute a *linear feedback Nash equilibrium* solution if both

$$J_1(u_1^*, u_2^*) \leq J_1(u_1, u_2^*) \text{ and } J_2(u_1^*, u_2^*) \leq J_1(u_1^*, u_2), \text{ for all } u_i \in \Gamma_i^{lfb}. \quad \square$$

Remark 4.2. In the sequel, with some abuse of notation, sometimes the pair $(F_1^*(t), F_2^*(t))$ will be called a (linear) feedback Nash equilibrium. □

Similar as with open-loop Nash equilibria, it turns out that linear feedback Nash equilibria can be explicitly determined by solving a set of coupled Riccati equations.

Theorem 4.3. The two-player linear quadratic differential game (1),(2) has for every initial state a linear feedback Nash equilibrium if and only if the next set of coupled Riccati differential equations has a set of symmetric solutions K_1, K_2 on $[0, T]$:

$$\begin{aligned} \dot{K}_1(t) &= -(A - S_2K_2(t))^T K_1(t) - K_1(t)(A - S_2K_2(t)) + K_1(t)S_1K_1(t) \\ &\quad -Q_1 - K_2(t)S_{21}K_2(t),\ K_1(T) = Q_{1T} \qquad (16) \\ \dot{K}_2(t) &= -(A - S_1K_1(t))^T K_2(t) - K_2(t)(A - S_1K_1(t)) + K_2(t)S_2K_2(t) \\ &\quad -Q_2 - K_1(t)S_{12}K_1(t),\ K_2(T) = Q_{2T}. \qquad (17) \end{aligned}$$

Moreover, in that case there is a unique equilibrium. The equilibrium actions are:

$$u_i^*(t) = -R_{ii}^{-1}B_i^T K_i(t)x(t),\ i = 1, 2.$$

The cost incurred by player i is $x_0^T K_i(0)x_0,\ i = 1, 2.$ □

Next we consider the infinite planning horizon case. Like in the open-loop case, we consider the minimization of the performance criterion (10). In line with our motivation for the finite-planning horizon, it seems reasonable to study Nash equilibria within the class of linear time-invariant state feedback policy rules. Therefore, we shall restrain our set of permitted controls to the constant linear feedback strategies. That is, to $u_i = F_i x$, with $F_i \in \mathbb{R}^{m_i \times n},\ i = 1, 2$, and where (F_1, F_2) belongs to the set

$$\mathcal{F} := \{F = (F_1, F_2) \mid A + B_1F_1 + B_2F_2 \text{ is stable}\}.$$

The stabilization constraint is imposed to ensure the finiteness of the infinite-horizon cost integrals that we will consider. This assumption can also be justified from the supposition that one is studying a perturbed system which is temporarily out of equilibrium. In that case, it is reasonable to expect that the state of the system remains close to the origin. Obviously, the stabilization constraint is a bit unwieldy since it introduces dependence between the strategy spaces of the players. So, it presupposes that there is at least the possibility of some coordination between both players. This coordination assumption seems to be more stringent in this case than for the equilibrium concepts we introduced before. However, the stabilization constraint can be motivated from the supposition that both players have a first priority in stabilizing the system. Whether this coordination actually takes place depends on the outcome of the game. Only in case when the players have the impression that their actions are such that the system becomes unstable, will they coordinate their actions in order to realize this meta-objective and adapt their actions accordingly. Probably for most games the equilibria without this stabilization constraint coincide with the equilibria of the game if one does consider this additional stabilization constraint. That is, the stabilization constraint will be in most cases not active. But there are games where it does play a role.

To make sure that our problem setting makes sense, we assume that the set $\mathcal{F}$ is non-empty. A necessary and sufficient condition for this to hold is that the matrix pair $(A, \begin{bmatrix} B_1, & B_2 \end{bmatrix})$ is stabilizable.

Summarizing, we define the concept of a linear feedback Nash equilibrium on an infinite-planning horizon as follows.

Definition 4.4. $(F_1^*, F_2^*) \in \mathcal{F}$ is called a *stationary linear feedback Nash equilibrium* if the following inequalities hold:

$$J_1(x_0, F_1^*, F_2^*) \le J_1(x_0, F_1, F_2^*) \text{ and } J_2(x_0, F_1^*, F_2^*) \le J_2(x_0, F_1^*, F_2)$$

for each x_0 and for each state feedback matrix $F_i, \ i = 1, 2$ such that (F_1^*, F_2) and $(F_1, F_2^*) \in \mathcal{F}$. □

Unless stated differently, the phrases “stationary" and “linear" in the notion of stationary linear feedback Nash equilibrium are dropped. It is clear from the context here which equilibrium concept we are dealing with.

Next, consider the set of coupled algebraic Riccati equations:

$$\begin{aligned} 0 &= -(A - S_2K_2)^T K_1 - K_1(A - S_2K_2) + K_1S_1K_1 \\ &\qquad -Q_1 - K_2S_{21}K_2, \qquad (18) \\ 0 &= -(A - S_1K_1)^T K_2 - K_2(A - S_1K_1) + K_2S_2K_2 \\ &\qquad -Q_2 - K_1S_{12}K_1. \qquad (19) \end{aligned}$$

Theorem 4.5 below states that feedback Nash equilibria are completely characterized by *stabilizing solutions* of (18),(19) that is, by solutions (K_1, K_2) for which the closed-loop system matrix $A - S_1K_1 - S_2K_2$ is stable.

Theorem 4.5. Let (K_1, K_2) be a stabilizing solution of (18),(19) and define $F_i^* := -R_{ii}^{-1}B_i^T K_i$ for $i = 1, 2$. Then (F_1^*, F_2^*) is a feedback Nash equilibrium. Moreover, the cost incurred by player i by playing this equilibrium action is $x_0^T K_i x_0, \ i = 1, 2$.
Conversely, if (F_1^*, F_2^*) is a feedback Nash equilibrium, there exists a stabilizing solution (K_1, K_2) of (18),(19) such that $F_i^* = -R_{ii}^{-1}B_i^T K_i$. □

Theorem 4.5 shows that all infinite-planning horizon feedback Nash equilibria can be found by solving a set of coupled algebraic Riccati equations. Solving the system (18),(19) is, in general, a difficult problem. To get some intuition for the solution set we next consider the scalar two-player game, where players are not interested in the control actions pursued by the other player. In that case, it is possible to derive some analytic results. In particular, it can be shown that in this game never more than three equilibria occur. Furthermore, a complete characterization of parameters which give rise to either 0, 1, 2, or 3 equilibria is possible.

So, consider the game:

$$J_i(x_0, u_1, u_2) = \int_0^{\infty} \{q_i x^2(t) + r_i u_i^2\} dt, \ i = 1, 2, \tag{20}$$

subject to the dynamical system

$$\dot{x}(t) = ax(t) + b_1 u_1(t) + b_2 u_2(t), \ x(0) = x_0. \tag{21}$$

The associated relevant algebraic Riccati equations are obtained from (18),(19) by substitution of $R_{21} = R_{12} = 0$, $A = a$, $B_i = b_i$, $Q_i = q_i$, $R_{ii} = r_i$, and $s_i := b_i^2/r_i$, $i = 1, 2$, into these equations. By Theorem 4.5, then a pair of control actions $f_i^* := -\frac{b_i}{r_i} k_i$, constitute a feedback Nash equilibrium if and only if the next equations have a solution $x_i = k_i, \ i = 1, 2$:

$$s_1 x_1^2 + 2 s_2 x_1 x_2 - 2 a x_1 - q_1 = 0 \tag{22}$$
$$s_2 x_2^2 + 2 s_1 x_1 x_2 - 2 a x_2 - q_2 = 0 \tag{23}$$
$$a - s_1 x_1 - s_2 x_2 < 0. \tag{24}$$

Geometrically, Eqs. (22) and (23) represent two hyperbolas in the (x_1, x_2) plane, whereas the inequality (24) divides this plane into a “stable” and an “anti-stable” region. So, all feedback Nash equilibria are obtained as the intersection points of both hyperbolas in the “stable” region. Example 4.6 illustrates the situation.

Example 4.6. Consider $a = b_i = r_i = 1, \ i = 1, 2$, $q_1 = \frac{1}{4}$, and $q_2 = \frac{1}{5}$. Then the hyperbola describing (22),(23) are:

$$x_2 = 1 - \frac{1}{2} x_1 + \frac{1}{8 x_1}, \text{ and } x_1 = 1 - \frac{1}{2} x_2 + \frac{1}{10 x_2}, \text{ respectively.}$$

Both hyperbola, as well as the "stability-separating" line $x_2 = 1 - x_1$, are plotted in Fig. 6. From the plot we see that both hyperbola have three intersection points in the stable region. So, the game has three feedback Nash equilibria. □

Next, introduce $\sigma_i := s_i q_i$ and for all $x > 0$, satisfying $x^2 \geq \sigma_1$, the functions:

$$f_1(x) = x - \sqrt{x^2 - \sigma_1} - \sqrt{x^2 - \sigma_2} \tag{25}$$
$$f_2(x) = x + \sqrt{x^2 - \sigma_1} - \sqrt{x^2 - \sigma_2} \tag{26}$$
$$f_3(x) = x - \sqrt{x^2 - \sigma_1} + \sqrt{x^2 - \sigma_2} \tag{27}$$
$$f_4(x) = x + \sqrt{x^2 - \sigma_1} + \sqrt{x^2 - \sigma_2}. \tag{28}$$

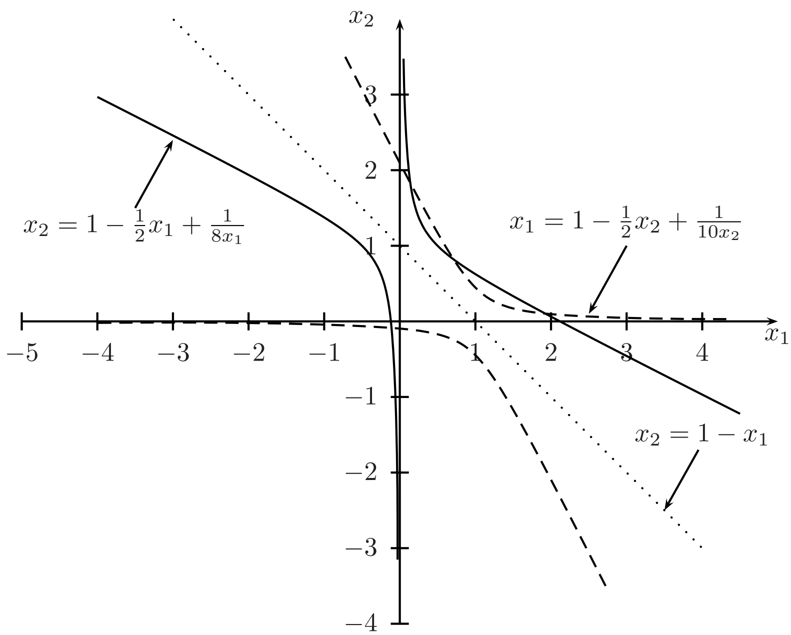

Figure 6: A game with three feedback Nash equilibria: $a = b_i = r_i = 1, q_1 = \frac{1}{4}, q_2 = \frac{1}{5}$.

The number of equilibria coincides with the total number of intersection points all of the functions $f_i(x),\ i = 1, \cdots, 4$, have with the horizontal line a (see Fig. 7). Elaboration of this property yields the following result.

Theorem 4.7. Consider the differential game (20)–(21), with $\sigma_i = \frac{b_i^2 q_i}{r_i},\ i = 1, 2$. Assume, without loss of generality, that $\sigma_1 \geq \sigma_2$. Moreover, let $f_i(x),\ i = 1, \cdots, 4$, be defined as in (25)–(28). Then, if:

1a. $\sigma_1 > 0$ and $\sigma_1 > \sigma_2$, the game has
- one equilibrium if $-\infty < a < \min f_3(x)$
- two equilibria if $a = \min f_3(x)$
- three equilibria if $a > \min f_3(x)$.

1b. $\sigma_1 = \sigma_2 > 0$ the game has
- one equilibrium if $a \leq \sqrt{\sigma_1}$
- three equilibria if $a > \sqrt{\sigma_1}$.

2a. $\sigma_1 < 0$ and $\sigma_1 > \sigma_2$, the game has
- no equilibrium if $\max f_1(x) < a \leq \sqrt{-\sigma_1} - \sqrt{-\sigma_2}$
- one equilibrium if either
  - i) $a = \max f_1(x)$;
  - ii) $a \leq -\sqrt{-\sigma_1} - \sqrt{-\sigma_2}$, or
  - iii) $-\sqrt{-\sigma_1} - \sqrt{-\sigma_2} < a \leq \sqrt{-\sigma_1} + \sqrt{-\sigma_2}$
- two equilibria if either

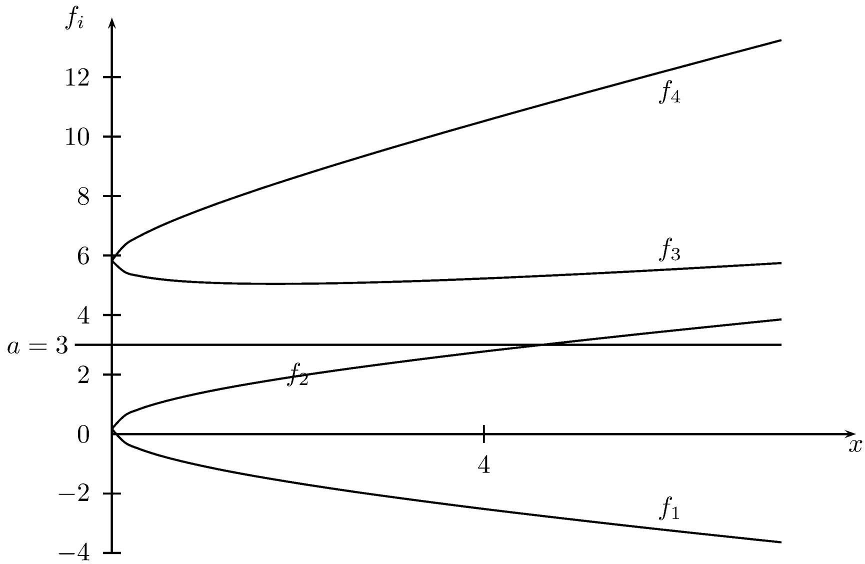

Figure 7: The curves f_i for $\sigma_1 = 9;\ \sigma_2 = 1$.

i) $-\sqrt{-\sigma_1} - \sqrt{-\sigma_2} < a < \max f_1(x)$, or
ii) $-\sqrt{-\sigma_1} + \sqrt{-\sigma_2} < a \leq \sqrt{-\sigma_1} + \sqrt{-\sigma_2}$

- three equilibria if $a > \sqrt{-\sigma_1} + \sqrt{-\sigma_2}$.

2b. $\sigma_1 = \sigma_2 < 0$ the game has

- no equilibrium if $-\sqrt{-3\sigma} < a \leq 0$,
- one equilibrium if $a \leq -2\sqrt{-\sigma}$,
- two equilibria if $-2\sqrt{-\sigma} < a \leq -\sqrt{-3\sigma}$ and $0 < a \leq 2\sqrt{-\sigma}$,
- three equilibria if $a > 2\sqrt{-\sigma}$. □

Example 4.8. Take $a = 3$, $b_i = r_i = 1,\ i = 1, 2$, $q_1 = 9$ and $q_2 = 1$. Then, $\sigma_1 = 9 > 1 = \sigma_2$. Furthermore, $f_3(x) - 3 = x - 3 - \sqrt{x^2 - 9} + \sqrt{x^2 - 1} > 0,$ if $x \geq 3$. So, $\min_{x \geq 3} f_3(x) > 3$, and therefore $-\infty < a < \min f_3(x)$. According Theorem 4.7, item 1a, the game has one feedback Nash equilibrium. In Fig. 7 one can also see this graphically. □

Below we present a numerical algorithm to calculate all feedback Nash equilibria of the two-player scalar game (20)–(21). This algorithm can be generalized for the corresponding N-player scalar game (see [9]). The algorithm follows from the the next result.

Theorem 4.9.

1. Assume that (k_1, k_2) is a feedback Nash equilibrium strategy. Then the negative of the corresponding closed-loop system parameter $\lambda := -a + \sum_{i=1}^{2} s_i k_i > 0$ is an eigenvalue of the matrix

$$M := \begin{bmatrix} -a & s_1 & s_2 & 0 \\ q_1 & a & 0 & -s_2 \\ q_2 & 0 & a & -s_1 \\ 0 & \frac{1}{3}q_2 & \frac{1}{3}q_1 & \frac{1}{3}a \end{bmatrix}. \tag{29}$$

$[1, k_1, k_2, k_1k_2]^T$ is a corresponding eigenvector and $\lambda^2 \geq \sigma_{max}$.

2. Assume that $[1, k_1, k_2, k_3]^T$ is an eigenvector corresponding to a positive eigenvalue λ of M, satisfying $\lambda^2 \geq \sigma_{max}$, and that the eigenspace corresponding with λ has dimension one. Then, (k_1, k_2) is a feedback Nash equilibrium. □

Algorithm 4.10. The following algorithm calculates all feedback Nash equilibria of the linear quadratic differential game (20),(21).

1: Calculate matrix M in (29) and $\sigma := \max_i \frac{b_i^2 q_i}{r_i}$.

2: Calculate the eigenstructure (λ_i, m_i), $i = 1, \cdots, k$, of M, where λ_i are the eigenvalues and m_i the corresponding algebraic multiplicities.

3: For $i = 1, \cdots, k$ repeat the following steps:

3.1) If (i) $\lambda_i \in \mathbb{R}$; ii) $\lambda_i > 0$ and iii) $\lambda_i^2 \geq \sigma$ then proceed with Step 3.2 of the algorithm. Otherwise, return to **3**.

3.2) If $m_i = 1$ then

3.2.1) calculate an eigenvector v corresponding with λ_i of M. Denote the entries of v by $[v_0, v_1, v_2]^T$. Calculate $k_j := \frac{v_j}{v_0}$ and $f_j := -\frac{b_j k_j}{r_j}$. Then, (f_1, f_2) is a feedback Nash equilibrium and $J_j = k_j x_0^2$, $j = 1, 2$. Return to **3**.

If $m_i > 1$ then

3.2.2) Calculate $\sigma_i := \frac{b_i^2 q_i}{r_i}$.

3.2.3) For all 4 sequences (t_1, t_2), $t_k \in \{-1, 1\}$,

(i) calculate $y_j := \lambda_i + t_j \sqrt{\lambda_i^2 - \sigma_j}$, $j = 1, 2$.

ii) If $\lambda_i = -a + \sum_{j=1}^{2} y_j$ then calculate $k_j := \frac{y_j r_j}{b_j^2}$ and $f_j := -\frac{b_j k_j}{r_j}$. Then, (f_1, f_2) is a feedback Nash equilibrium and $J_j = k_j x_0^2$, $j = 1, 2$.

4: End of the algoritm. □

Example 4.11. Reconsider Example 3.9 where, for $a = -2; b_i = r_{ii} = q1 = 1$, and $q_2 = 4$, $i = 1, 2$, we calculated the open-loop Nash equilibrium for an infinite planning horizon. To calculate the feedback Nash equilibria for this game, according Algorithm 4.10, we first have to determine the eigenstructure of matrix:

$$M := \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & -2 & 0 & -1 \\ 4 & 0 & -2 & -1 \\ 0 & \frac{4}{3} & \frac{1}{3} & \frac{-2}{3} \end{bmatrix}.$$

Using Matlab, we find the eigenvalues $\{-2.6959, -1.4626 \pm 0.7625i, 2.9543\}$. Since the square of 2.9543 is larger than $\sigma := 4$, we have to process Step 3 of the algorithm for this eigenvalue. Since its algebraic multiplicity is 1, we calculate an eigenvector v corresponding with this eigenvalue 2.9543. Choosing $v^T = [v_0,\ v_1,\ v_2\ v_3] := [0.7768,\ 0.1355,\ 0.6059,\ 0.1057]$ yields

$$k_1 := \frac{v_1}{v_0} = 0.1744 \text{ and } k_2 := \frac{v_2}{v_0} = 0.7799.$$

This gives the feedback Nash equilibrium actions;

$$u_1(t) = -\frac{b_1 k_1}{r_1} x(t) = -0.1744x(t) \text{ and } u_2(t) = -\frac{b_2 k_2}{r_2} x(t) = -0.7799x(t).$$

The corresponding closed-loop system and cost are:

$$\dot{x}(t) = -2.9543x(t),\ x(0) = x_0; \text{ and } J_1 = 0.1744x_0^2,\ J_2 = 0.7799x_0^2.$$

Note that these cost almost coincide with the open-loop case. □

5 The Uncertain Non-cooperative Game

Dynamic game theory brings together three features that are key to many situations in economy, ecology, and elsewhere: optimizing behavior, presence of multiple agents, and enduring consequences of decisions. In this section, we add a fourth aspect, namely robustness with respect to variability in the environment. In our formulation of dynamic games, so far, we specified a set of differential equations including input functions that are controlled by the players, and players are assumed to optimize a criterion over time. The dynamic model is supposed to be an exact representation of the environment in which the players act; optimization takes place with no regard of possible deviations. It can safely be assumed, however, that agents in reality follow a different strategy. If an accurate model can be formed at all, it will in general be complicated and difficult to handle. Moreover, it may be unwise to optimize on the basis of a too detailed model, in view of possible changes in dynamics that may take place in the course of time and that may be hard to predict. It makes more sense for agents to work on the basis of a relatively simple model and to look for strategies that are robust with respect to deviations between the model and reality. In an economic context, the importance of incorporating aversion to specification uncertainty has been stressed, for instance, by Hansen *et al.* [18].

In control theory, an extensive theory of robust design is already in place; see, e.g., Başar [6]. We use this background to arrive at suitable ways of describing aversion to model risk in a dynamic game context. We assume linear dynamics and quadratic cost functions. These assumptions are reasonable for situations of dynamic quasi-equilibrium, where no large excursions of the state vector are to be expected.

Following a pattern that has become standard in control theory, two approaches can be considered. The first one is based on a stochastic approach. This approach assumes that the dynamics of the system are corrupted by a standard Wiener process (white-noise). Basic assumptions are that the players have access to the current value of the state of the system and that the positive definite covariance matrix does not depend on the state of the system. Basically, it turns out that under these assumptions the feedback Nash equilibria also constitute an equilibrium in such an uncertain environment. For that reason we will not elaborate that case here any further. In the second approach, a malevolent disturbance input is introduced which is used in the modeling of aversion to specification uncertainty. That is, it is assumed that the dynamics of the system are corrupted by a deterministic noise component, and that each player has his own expectation about this noise. This is modeled by adapting for each player his cost function accordingly. The players cope with this uncertainty by considering a worst-case scenario. Consequently, in this approach the equilibria of the game, in general, depend on the worst-case scenario expectations about the noise of the players.

In this section we restrict the analysis to the infinite-planning horizon case. We will first consider the open-loop one-player case.

Consider the problem to find

$$\inf_{u\in\mathcal{U}_s} \sup_{w\in L_2^q(0,\infty)} \int_0^\infty \{x^T(t)Qx(t) + u^T(t)Ru(t) - w^T(t)Vw(t)\}dt, \tag{30}$$

subject to

$$\dot{x}(t) = Ax(t) + Bu(t) + Ew(t),\ x(0) = x_0, \tag{31}$$

where $R > 0$, $V > 0$ and A is stable.

This problem formulation originates from the H_∞ disturbance attenuation control problem and is in literature known (see [4]) as the soft-constrained open-loop differential game. Since no definiteness assumptions were made in Sec. 2 w.r.t. the matrices Q_i, one can use the in that section obtained results directly to derive the next result.

Corollary 5.1. Consider problem (30),(31). Let $S := BR^{-1}B^T$ and $M := EV^{-1}E^T$. This problem has a solution for every initial state x_0 if:

1. The coupled algebraic Riccati equations
$$\begin{aligned} A^T P_1 + P_1 A + Q - P_1 S P_1 - P_1 M P_2 &= 0, \\ A^T P_2 + P_2 A - Q - P_2 M P_2 - P_2 S P_1 &= 0, \end{aligned}$$
have a strongly stabilizing solution; and
2. The two algebraic Riccati equations
$$\begin{aligned} A^T K_1 + K_1 A - K_1 S K_1 + Q &= 0, \\ A^T K_2 + K_2 A - K_2 M K_2 - Q &= 0, \end{aligned}$$
have a symmetric solution K_i, such that $A - SK_1$ and $A - MK_2$ are stable.

Moreover, a worst-case control for the player is $u^*(t) = -R^{-1}B^T P_1 x(t)$.
The corresponding worst-case disturbance is $w^*(t) = -V^{-1}E^T P_2 x(t)$.
Here, $x(t)$ satisfies the differential equation;

$$\dot{x}(t) = (A - SP_1 - MP_2)x(t);\ x(0) = x_0. \qquad \square$$

In case one additionally assumes that $Q \geq 0$, the above result simplifies even further and one gets the next result. More results (and, in particular, converse statements) on the soft-constrained open-loop differential game can be found in [4, Section 4.2.1 and Theorem 9.6] and in [21] some preliminary results for the multi-player case.

Corollary 5.2. Consider problem (30),(31). This problem has a solution for every initial state x_0 if the algebraic Riccati equations

$$A^T P + PA + Q - P(S - M)P = 0 \text{ and } A^T K + KA + KMK + Q = 0$$

have a solution $\bar{P}$ and $\bar{K}$, respectively, such that $A - (S - M)\bar{P}$ and $A + M\bar{K}$ are stable.
Furthermore, a worst-case control for the player is $u^*(t) = -R^{-1}B^T \bar{P} x(t)$.
The corresponding worst-case disturbance is $w^*(t) = V^{-1}E^T \bar{P} x(t)$.
Here, $x(t)$ satisfies the differential equation

$$\dot{x}(t) = (A - (S - M)\bar{P})x(t);\ x(0) = x_0.$$

Moreover, $\bar{J} = x_0^T \bar{P} x_0$. $\square$

From the above corollary we infer, in particular, that if $x_0 = 0$, the best open-loop worst-case controller is $u = 0$, whereas the worst-case signal in that case is $w = 0$. This independent of the choice of V, under the supposition that the Riccati equations have an appropriate solution. So if a stable system is in equilibrium (i.e., $x_0 = 0$), in this open-loop framework the best reaction to potential unknown disturbances is not to react.

Next we consider the corresponding problem within a feedback information framework. That is, consider

$$\dot{x} = (A + BF)x + Ew, \ x(0) = x_0, \tag{32}$$

with (A, B) stabilizable, $F \in \mathcal{F}$ and

$$J(F, w, x_0) = \int_0^\infty \{x^T(Q + F^T RF)x - w^T V w\}dt. \tag{33}$$

The matrices Q, R, and V are symmetric, $R > 0$ and $V > 0$. The problem is to determine for each $x_0 \in \mathbb{R}^n$ the value

$$\inf_{F\in\mathcal{F}} \sup_{w\in L_2^q(0,\infty)} J(F, w, x_0). \tag{34}$$

Furthermore, if this infimum is finite, we like to know whether there is a feedback matrix $\bar{F} \in \mathcal{F}$ that achieves the infimum, and to determine all matrices that have this property. This soft-constrained differential game can also be interpreted as a model for a situation where the controller designer is minimizing the criterion (33) by choosing an appropriate $F \in \mathcal{F}$, while the uncertainty is maximizing the same criterion by choosing an appropriate $w \in L_2^q(0, \infty)$.

A necessary condition for the expression in (34) to be finite is that the supremum

$$\sup_{w\in L_2^q(0,\infty)} J(F, w, x_0)$$

is finite for at least one $F \in \mathcal{F}$. However, this condition is not sufficient. It may happen that the infimum in (34) becomes arbitrarily small. Below we present a sufficient condition under which the soft-constrained differential game has a saddle point.

Theorem 5.3. Consider (32)–(33) and let S and M be as defined before. Assume that the algebraic Riccati equation

$$Q + A^T X + XA - XSX + XMX = 0 \tag{35}$$

has a stabilizing solution X i.e., $A - SX + MX$ stable) and that, additionally, $A - SX$ is stable. Furthermore, assume that there exists a real symmetric $n \times n$ symmetric matrix Y that satisfies the matrix inequality:

$$Q + A^T Y + Y A - Y SY \geq 0. \tag{36}$$

Define $\bar{F} := -R^{-1}B^T X$ and $\bar{w}(t) := V^{-1}E^T Xe^{(A-SX+MX)t}x_0$. Then the matrix $\bar{F}$ belongs to $\bar{\mathcal{F}}$, the function $\bar{w}$ is in $L_2^q(0, \infty)$, and for all $F \in \mathcal{F}$ and $w \in L_2^q(0, \infty)$:

$$J(\bar{F}, w, x_0) \leq J(\bar{F}, \bar{w}, x_0) \leq J(F, \bar{w}, x_0).$$

Moreover, $J(\bar{F}, \bar{w}, x_0) = x_0^T X x_0$. □

If $Q \geq 0$, condition (36) is trivially satisfied by choosing $Y = 0$. Corollary 5.4, below, summarizes the consequences of Theorem 5.3 for problem (34).

Corollary 5.4. Let the assumptions of Theorem 5.3 hold and let X, $\bar{F}$, and $\bar{w}$ be as in that theorem. Then,

$$\min_{F\in\mathcal{F}} \sup_{w\in L_2^q(0,\infty)} J(F,w,x_0) = \max_{w\in L_2^q(0,\infty)} J(\bar{F},w,x_0) = x_0^T X x_0$$

and $\max_{w\in L_2^q(0,\infty)} \inf_{F\in\mathcal{F}} J(F,w,x_0) = \min_{F\in\mathcal{F}} J(F,\bar{w},x_0) = x_0^T X x_0.$ □

Next we consider the multi-player soft-constrained differential game. That is, we consider

$$\dot{x}(t) = (A + B_1F_1 + B_2F_2)x(t) + Ew(t),\ x(0) = x_0, \tag{37}$$

with $(A, [B_1,\ B_2])$ stabilizable, $(F_1, F_2) \in \mathcal{F}$ and $J_i(F_1, F_2, w, x_0) =$

$$\int_0^\infty \{x^T(t)(Q_i + F_1^T R_{i1} F_1 + F_2^T R_{i2} F_2)x(t) - w^T(t)V_i w(t)\}dt. \tag{38}$$

Here the matrices Q_i, R_{ij}, and V_i are symmetric, $R_{ii} > 0$, $V_i > 0$, and

$$\mathcal{F} := \{(F_1, F_2) | A + B_1F_1 + B_2F_2 \text{ is stable}\}.$$

For this game we want to determine all soft-constrained Nash equilibria. That is, to find all $(\bar{F}_1, \bar{F}_2) \in \mathcal{F}$ such that

$$\sup_{w\in L_2^q(0,\infty)} J_1(\bar{F}_1,\bar{F}_2,w,x_0) \leq \sup_{w\in L_2^q(0,\infty)} J_1(F_1,\bar{F}_2,w,x_0),\ \forall (F_1,\bar{F}_2)\in\mathcal{F} \tag{39}$$

and

$$\sup_{w\in L_2^q(0,\infty)} J_2(\bar{F}_1,\bar{F}_2,w,x_0) \leq \sup_{w\in L_2^q(0,\infty)} J_2(\bar{F}_1,F_2,w,x_0),\ \forall (\bar{F}_1,F_2)\in\mathcal{F}, \tag{40}$$

for all $x_0 \in \mathbb{R}^m$.

Because the weighting matrix V_i occurs with a minus sign in (38), this matrix constrains the disturbance vector w in an indirect way so that it can be used to describe the aversion to model risk of player i. Specifically, if the quantity $w^T V_i w$ is large for a vector $w \in \mathbb{R}^q$, this means that player i does not expect large deviations of the nominal dynamics in the direction of Ew. Furthermore, the larger he chooses V_i, the closer the worst-case signal he can be confronted with in this model will approach the zero input signal (that is: $w(.) = 0$).

From Corollary 5.4, a sufficient condition for the existence of a soft-constrained feedback Nash equilibrium follows in a straightforward way. Using the shorthand notation

$$S_i := B_i R_{ii}^{-1} B_i^T, S_{ij} := B_i R_{ii}^{-1} R_{ji} R_{ii}^{-1} B_i^T,\ i \neq j, \text{ and } M_i := E V_i^{-1} E^T,$$

we have the next result.

Theorem 5.5. Consider the differential game defined by (37)–(40). Assume there exist real symmetric $n \times n$ matrices $X_i,\ i = 1, 2$, and real symmetric $n \times n$ matrices $Y_i,\ i = 1, 2$, such that:

$$\begin{aligned} &-(A - S_2X_2)^T X_1 - X_1(A - S_2X_2) + X_1S_1X_1 - Q_1 - \\ &\qquad X_2S_{21}X_2 - X_1M_1X_1 = 0, \end{aligned} \tag{41}$$

$$\begin{aligned} &-(A - S_1X_1)^T X_2 - X_2(A - S_1X_1) + X_2S_2X_2 - Q_2 - \\ &\qquad X_1S_{12}X_1 - X_2M_2X_2 = 0, \end{aligned} \tag{42}$$

$$A - S_1X_1 - S_2X_2 + M_1X_1 \text{ and } A - S_1X_1 - S_2X_2 + M_2X_2 \text{ are stable,} \tag{43}$$

$$A - S_1X_1 - S_2X_2 \text{ is stable} \tag{44}$$

$$-(A - S_2X_2)^T Y_1 - Y_1(A - S_2X_2) + Y_1S_1Y_1 - Q_1 - X_2S_{21}X_2 \leq 0, \tag{45}$$

$$-(A - S_1X_1)^T Y_2 - Y_2(A - S_1X_1) + Y_2S_2Y_2 - Q_2 - X_1S_{12}X_1 \leq 0. \tag{46}$$

Define $\overline{F} = (\overline{F}_1, \overline{F}_2)$ by $\overline{F}_i := -R_{ii}^{-1} B_i^T X_i,\ i = 1, 2.$
Then $\overline{F} \in \mathcal{F}$, and $\overline{F}$ is a soft-constrained Nash equilibrium. Furthermore, the worst-case signal $\bar{w}_i$ from player $i's$ perspective is:

$$\bar{w}(t) = V_i^{-1} E^T X_i e^{(A - S_1X_1 - S_2X_2 + M_iX_i)t} x_0.$$

Moreover, the cost for player i under his worst-case expectations are:

$$\overline{J}_i^{SC}(\overline{F}_1, \overline{F}_2, x_0) = x_0^T X_i x_0,\ i = 1, 2.$$

Conversely, if $(\bar{F}_1, \bar{F}_2)$ is a soft-constrained Nash equilibrium, the Eqs. (41)–(44) have a set of real symmetric solutions (X_1, X_2). □

Again, notice that if $Q_i \geq 0,\ i = 1, 2$, and $S_{ij} \geq 0,\ i, j = 1, 2$, the matrix inequalities (45)–(46) are trivially satisfied with $Y_i = 0,\ i = 1, 2$. So, under these conditions the differential game defined by (37)–(40) has a soft-constrained Nash equilibrium if and only if the Eqs. (41)–(44) have a set of real symmetric $n \times n$ matrices $X_i,\ i = 1, 2$.

Theorem 5.5 shows that the Eqs. (41)–(44) play a crucial role in the question whether the game (37)–(38) will have a soft-constrained Nash equilibrium. Every soft-constrained Nash equilibrium has to satisfy these equations. So, the question arises under which conditions (41)–(44) will have one or more solutions and, if possible, to calculate this (these) solution(s). This is a difficult open question. Similar remarks apply here as were made in Sec. 4 for solving the corresponding set of algebraic Riccati equations to determine the feedback Nash equilibria. But, again for the scalar case, one can devise an algorithm to calculate all soft-constrained Nash equilibria. This algorithm is in the spirit of Algorithm 4.10 and will be discussed now.

As in in Sec. 4, we will consider here just the two-player case under the simplifying assumptions that $b_i \neq 0$ and players have no direct interest in eachothers control actions i.e., $S_{ij} = 0,\ i \neq j$). For more details and the general N-player case, we refer again to the literature ([9], [12]). Again, lower-case notation will be used to stress the fact that we are dealing with the scalar case. For notational convenience let Ω denote either the set $\{1\},\ \{2\}$, or $\{1,2\}$. Furthermore, let:

$$\tau_i := (s_i + m_i)q_i,\ \tau_{max} := \max_i \tau_i,\ \rho_i := \frac{s_i}{s_i + m_i}, \tag{47}$$
$$\gamma_i := -1 + 2\rho_i = \frac{s_i - m_i}{s_i + m_i} \text{ and } \gamma_\Omega := -1 + 2\sum_{i \in \Omega} \rho_i.$$

With some small abuse of notation for a fixed index set Ω, γ_Ω will also be denoted without brackets and comma's. That is, if e.g., $\Omega = \{1,2\}$, γ_Ω is also written as γ_{12}.

An analogous reasoning as in Theorem 4.9 gives the next Theorem.

Theorem 5.6.

1. Assume that (x_1, x_2) solves (41),(42),(44) and $\gamma_i \neq 0,\ i = 1,2,12$. Then $\lambda := -a + \sum_{i=1}^2 s_i x_i > 0$ is an eigenvalue of the matrix

$$M := \begin{bmatrix} -a & s_1 & s_2 & 0 \\ \frac{\rho_1 q_1}{\gamma_1} & \frac{a}{\gamma_1} & 0 & -\frac{s_2}{\gamma_1} \\ \frac{\rho_2 q_2}{\gamma_2} & 0 & \frac{a}{\gamma_2} & -\frac{s_1}{\gamma_2} \\ 0 & \frac{\rho_2 q_2}{\gamma_{12}} & \frac{\rho_1 q_1}{\gamma_{12}} & \frac{a}{\gamma_{12}} \end{bmatrix}; \tag{48}$$

$[1, x_1, x_2, x_1 x_2]^T$ is a corresponding eigenvector and $\lambda^2 \geq \tau_{max}$.

2. Assume that $[1, x_1, x_2, x_3]^T$ is an eigenvector corresponding to a positive eigenvalue λ of M, satisfying $\lambda^2 \geq \tau_{max}$, and that the eigenspace corresponding with λ has dimension one. Then, (x_1, x_2) solves (41),(42),(44). □

Using this result we can calculate soft-constrained feedback Nash equilibria by implementing the next numerical algorithm.

Algorithm 5.7. Let $s_i := \frac{b_i^2}{r_i}$ and $m_i := \frac{e^2}{v_i}$. Assume that for every index set Ω, $\gamma_\Omega \neq 0$. Then, the following algorithm calculates all solutions of (41),(42),(44).

1: Calculate matrix M in (48) and $\tau := \max_i (s_i + m_i) q_i$.

2: Calculate the eigenstructure (λ_i, n_i), $i = 1, \cdots, k$, of M, where λ_i are the eigenvalues and n_i the corresponding algebraic multiplicities.

3: For $i = 1, \cdots, k$ repeat the following steps:

3.1) If (i) $\lambda_i \in \mathbb{R}$; ii) $\lambda_i > 0$ and iii) $\lambda_i^2 \geq \tau$ then proceed with Step 3.2 of the algorithm. Otherwise, return to **3**.

3.2) If $n_i = 1$ then

3.2.1) calculate an eigenvector z corresponding with λ_i of M. Denote the entries of z by $[z_0, z_1, z_2]^T$. Calculate $x_j := \frac{z_j}{z_0}$. Then, (x_1, x_2) solve (41),(42),(44). Return to **3**.

If $n_i > 1$ then

3.2.2) Calculate $\tau_i := s_i q_i$.

3.2.3) For all 4 sequences (t_1, t_2), $t_k \in \{-1, 1\}$,

(i) calculate $y_j := \lambda_i + t_j \frac{s_i}{s_i + m_i} \sqrt{\lambda_i^2 - \sigma_j}$, $j = 1, 2$.

ii) If $\lambda_i = -a + \sum_{j=1}^{2} y_j$ then calculate $x_j := \frac{y_j}{s_j + m_j}$. Then, (x_1, x_2) solves (41),(42),(44).

4: End of the algorithm. □

Example 5.8. Reconsider Example 3.9. That is, consider the two-player scalar game with $a = -2$, $b_i = e = 1$, $r_{ii} = 1$, $v_i = \frac{1}{9}$, $i = 1, 2$, $q_1 = 1$, and $q_2 = 4$. To calculate the soft-constrained Nash equilibria of this game, we first determine all solutions of (41),(42),(44). According Algorithm 5.7, we first have to determine the eigenstructure of the next matrix:

$$M := \begin{bmatrix} 2 & 1 & 1 & 0 \\ -1/8 & 5/2 & 0 & 5/4 \\ -1/2 & 0 & 5/2 & 5/4 \\ 0 & -4/42 & -1/42 & 20/42 \end{bmatrix}.$$

Using Matlab, we find that M has the eigenvalues $\{2.4741, .5435, 2.2293 \pm 0.7671i\}$. Since none of the positive eigenvalues squared is larger than $\tau = 40$, the game has no equilibrium.

Next, consider the case $v_i = 2$. In that case, numerical computations show that M has the eigenvalues $\{-6.8522, -5.67222, -1.7290, 3.0535\}$. The only positive eigenvalue which is squared larger than $\tau = 6$ is $\lambda = 3.0535$. So we have to process Step 3 of the algorithm for this eigenvalue. Since this eigenvalue has a geometric multiplicity of one, there is one solution satisfying (41),(42),(44). From the corresponding eigenspace one obtains the solution tabulated below (with $a_{cl} = a - s_1 x1 - s_2 x_2 = -$eigenvalue):

eigenvalue	(x_1, x_2)	$a_{cl} + m_1 x_1$	$a_{cl} + m_2 x_2$
3.0535	(0.1954,0.8581)	-2.9558	-2.6244

From the last two columns of this table we see that the solution satisfies the additional conditions (43). Since $q_i > 0$, and thus (45),(46) are satisfied with $y_i = 0$, it follows that this game has one soft-constrained Nash equilibrium. The with this equilibrium corresponding equilibrium actions are:

$$u_1^*(t) = -0.1954x(t) \text{ and } u_2^*(t) = -0.8581x(t).$$

Assuming that the initial state of the system is x_0, the worst-case expected cost by the players are $J_1^* = 0.1954x_0^2$ and $J_2^* = 0.8581x_0^2$, respectively.

Compared to the noise-free case we see that player one's cost increases by 12.75% and player two incurs a cost increase of 10.97%. □

6 Concluding Remarks

In this paper we reviewed some main results in the area of linear quadratic differential games. For didactical reasons, the results were presented for the two-player case. The exposition is based on [9] where one can find also additional results for the N-player case, proofs, references, and a historical perspective for further reading. In particular, this book also contains results on convergence properties of the finite planning horizon equilibria in case the planning horizon is extended to infinite. An issue that has not been addressed here.

For the cooperative game, one can find an extension of results from Sec. 2 to more general cost functions (like, e.g., indefinite Q_i matrices) in [14].

For the non-cooperative game, results generalizing on the cost functions considered here can be found for the open-loop information case in more detail in [10]. For the feedback information case in [13]: the existence of equilibria was considered if players can not observe the state of the system directly and use static output feedback to control the system.

A review on computational aspects involved with the calculation of the various equilibria can be found in [16], whereas [15] describes a numerical toolbox that is available on the web to calculate the unique open-loop Nash equilibrium for an infinite planning horizon.

For the infinite planning horizon case both in the open-loop and the feedback information case, the number of equilibria the game may have can vary between zero and infinity. Theorem 3.7 presents both necessary and sufficient conditions under which the game has a unique equilibrium for the open-loop information case. An open problem is whether one can also find for the feedback information case conditions that are both necessary and sufficient for the existence of a unique equilibrium. Under such conditions the numerical calculation of this equilibrium is then possible using one of the algorithms proposed in the literature as described, e.g., in [16].

Several special cases of LQ differential games have been considered in literature. We like to mention here two cases where also recently new numerical results were reported.
First, the set of weakly coupled large-scale systems has been studied extensively by, e.g., Mukaidani in a number of papers (see, e.g. [23]). This are systems where each player controls a set of states which are only marginally affected by other players. So, the corresponding LQ game almost equals an ordinary optimal LQ control problem. It can be shown that under the assumption that the coupling between the various “subsystems" is marginal the LQ game will have a unique equilibrium.
In [3], the set of positive systems has recently been considered. That is, the case that both the state and used controls should be positive at any point in time. In this paper conditions are stated under which such a system has an equilibrium and some algorithms are devised to calculate an equilibrium.

We hope this survey convinced the reader that the last decennia progress has been made in the theory about linear quadratic differential games and that there are many challenges left in this area for research. We would not be complete, if we did not mention the following important references that contributed to this research during the last decennium: Feucht [17], Weeren [29], Başar and Bernhard [4], Başar and Olsder [5], Kun [21], van den Broek [7], and Kremer [20].

References

[1] Aarle B. van, Bovenberg L. and Raith M., *Monetary and fiscal policy interaction and government debt stabilization*, Journal of Economics **62** (1995), 111–140.

[2] Abou-Kandil H. and Bertrand P., *Analytic solution for a class of linear quadratic open-loop Nash games*, International Journal of Control **43** (1986), 997–1002.

[3] Azevedo-Perdicoúlis T.P. and Jank G., *Linear quadratic Nash games on positive linear systems*, European Journal of Control **11** (2005), 525–535.

[4] Başar T. and Bernhard P., *H_∞-Optimal Control and Related Minimax Design Problems*, Birkhäuser, Boston (1995).

[5] Başar T. and Olsder G.J., *Dynamic Noncooperative Game Theory*, SIAM, Philadelphia (1999).

[6] Başar T., *Paradigms for robustness in controller and filter design*, In: Neck (Ed.) *Modeling and Control of Economic Systems 2002*, Elsevier, (2003), 1–13.

[7] Broek W.A. van den, *Uncertainty in Differential Games*, PhD. Thesis, CentER Dissertation Series, Tilburg University, The Netherlands (2001).

[8] Dockner E., Jørgensen S., Long N. van and Sorger G., *Differential Games in Economics and Management Science*, Cambridge University Press, Cambridge (2000).

[9] Engwerda J.C., *LQ Dynamic Optimization and Differential Games*, Wiley Chichester, (2005).

[10] Engwerda J.C., *Uniqueness conditions for the infinte-planning horizon linear quadratic differential game*, Automatica **44**(2) (2008), 504–511, online available as: http://dx.doi.org/10.1016/j.automatica.2007.06.006.

[11] Engwerda J.C., *The open-loop linear Quadratic Differential game revisited*, CentER DP 2005-34, Tilburg University, The Netherlands (2005).

[12] Engwerda J.C., 2006, *A numerical algorithm to find soft-constrained Nash equilibria in scalar LQ-games*, International Journal of Control **79** (2006), 592–603.

[13] Engwerda J.C. and Weeren A.J.T.M., *Output feedback Nash equilibria in LQ games*, Automatica **44**(1) (2008), 265–271, online available as: http://dx.doi.org/10.1016/j.automatica.2007.04.025.

[14] Engwerda J.C., *A note on cooperative linear quadratic control*, CentER DP 2007-15, online available as: http: //greywww.kub.nl:2080/greyfiles/center/ctr_py_2007.html (2007).

[15] Engwerda J.C., Michalak T. and Plasmans J.E.J., *A numerical toolbox to solve the N-player affine LQ open-loop differential game*, under construction, Dept. of Econometrics and O.R., Tilburg University, The Netherlands (2007).

[16] Engwerda J.C., *Algorithms for computing Nash equilibria in LQ games*, Computational Management Science **4** (2007), 113–140.

[17] Feucht M., *Linear-quadratische Differentialspiele und gekoppelte Riccatische Matrixdifferentialgleichungen*, Ph.D. Thesis, Universität Ulm, Germany (1994).

[18] Hansen L.P., Sargent T.J. and Tallarini Jr. T.D., *Robust permanent income and pricing*, Review of Economic Studies **66** (1999), 873–907.

[19] Jørgensen S. and Zaccour G., *Differential Games in Marketing*, Kluwer, Deventer (2004).

[20] Kremer D., *Non-Symmetric Riccati Theory and Noncooperative Games*, Ph.D. Thesis, RWTH-Aachen, Germany (2002).

[21] Kun G., *Stabilizability, Controllability, and Optimal Strategies of Linear and Nonlinear Dynamical Games*, PhD. Thesis, RWTH-Aachen, Germany (2001).

[22] Lancaster, P. and Rodman, L., *Algebraic Riccati Equations*, Oxford, Oxford University Press (1995).

[23] Mukaidani H., *A numerical analysis of the Nash strategy for weakly coupled large-scale systems*, IEEE Trans. Aut. Control **51** (2006), 1371–1377.

[24] Nash J., *The bargaining problem*, Econometrica, **18** (1950), 155–162.

[25] Nash J., *Two-person cooperative games*, Econometrica, **21** (1953), 128–140.

[26] Perko L.,*Differential Equations and Dynamical Systems*, Springer-Verlag, Berlin (2001).

[27] Plasmans J., Engwerda J., Aarle B. van, Bartolomeo G. di, Michalak T., *Dynamic Modeling of Monetary and Fiscal Cooperation Among Nations*, Springer-Verlag, Berlin (2006).

[28] Thomson W., 1994, *Cooperative models of bargaining*, in Aumann and Hart (Eds.), *Handbook of Game Theory* **2** (1994), Elsevier Science, 1238–1277.

[29] Weeren A.J.T.M., *Coordination in Hierarchical Control*, Ph.D. Thesis, Tilburg University, The Netherlands, (1995).

A Neumann Boundary Control for Multidimensional Parabolic "Minmax" Control Problems

Andrzej Nowakowski
Faculty of Mathematics
University of Lodz
Lodz, Poland
annowako@math.uni.lodz.pl

Abstract

In this paper, the optimal control problems of "minmax" type, governed by parabolic equations with the second control on the boundary (Neumann condition), are considered. We apply a new dual dynamic programming approach to derive sufficient optimality conditions for such problems. The idea is to move all the notions from a state space to a dual space and to obtain a new verification theorem providing the conditions which should be satisfied by a solution of the dual partial differential equation of dynamic programming. We also give sufficient optimality conditions for the existence of an optimal dual feedback control.

Key words. Dual dynamic programming, dual feedback control, parabolic equation, dynamic game, Neumann boundary control, verification theorem.

AMS Subject Classifications. Primary 49K20, 49K35, 49N35; Secondary 49J45.

1 Introduction

Consider the following optimal control problem (P):

$$\begin{aligned}\min_u \max_v J(x,u,v) &= \int_{[0,T]\times\Omega} L(t,z,x(t,z),u(t,z))dtdz \\ &+ \int_{\Omega} l\left(x\left(T,z\right)\right)dz + \int_{\Sigma} h(t,z,v(t,z))dtdz\end{aligned}$$

subject to

$$x_t\left(t,z\right) + \Delta_z x(t,z) = f(t,z,x(t,z),u(t,z)) \text{ a. e. on } [0,T]\times\Omega, \tag{1}$$

$$x\left(0,z\right) = \varphi\left(0,z\right) \text{ on } \Omega, \tag{2}$$

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_4,

$$\partial_\nu x(t,z) = v\,(t,z) \quad \text{on} \quad (0,T)\times\Gamma, \tag{3}$$

$$u(t,z) \in U \quad \text{a. e. on} \quad (0,T)\times\Omega, \tag{4}$$

$$v(t,z) \in \mathbf{V} \quad \text{on} \quad (0,T)\times\Gamma, \tag{5}$$

where $\Omega \subset R^n$, is bounded with C^2 boundary $\Gamma = \partial\Omega$, ∂_ν is the normal derivative to Γ , $U \subset R^m$, and $\mathbf{V} \subset R$ are given nonempty sets; $L, f : [0,T]\times\Omega\times R\times R^m \to R$, $l : R \to R$, $h : [0,T]\times\bar{\Omega}\times R \to R$, and $\varphi : R^{n+1} \to R$ are given functions; $x : [0,T]\times\Omega \to R$, $x \in W^{2,2}([0,T]\times\Omega)\cap C([0,T];L^2(\Omega))$ and $u : [0,T]\times\Omega \to R^m$ $v : [0,T]\times\Gamma \to R$ are Lebesgue measurable functions in suitable sets. We assume that for each s in R, the functions $(t,z,u) \to L(t,z,s,u)$, $(t,z,u) \to f(t,z,s,u)$ are $(L\times B)$-measurable, where $L\times B$ is the σ -algebra of subsets of $[0,T]\times\Omega\times R^m$ generated by products of Lebesgue measurable subsets of $[0,T]\times\Omega$ and Borel subsets of R^m, and for each $(t,z,u) \in [0,T]\times\Omega\times R^m$, the functions $s \to L(t,z,s,u)$, $s \to f(t,z,s,u)$ are continuous. Moreover, we assume that $(t,z,v) \to h(t,z,v)$ is Borel measurable in $[0,T]\times\Omega\times\mathbf{V}$. We call a trio $x(t,z), u(t,z), v(t,z)$ admissible if it satisfies (1)–(5) and $L(t,z,x(t,z),u(t,z))$, $h(t,z,v(t,z))$ and $l\,(x\,(T,z))$ are summable; then the corresponding trajectory $x(t,z)$ is said to be admissible. In this paper, we assume that system (1)–(5) admits at least one solution belonging to $W^{2,2}((0,T)\times\Omega)\cap C([0,T];L^2(\Omega))$.

The aim of this paper is to present "minmax" sufficient optimality conditions for problem (P) in terms of dynamic programming conditions directly. In the literature, there is no work in which problem (P) is studied directly by the dynamic programming method. The only results known to the author (see, e.g., [1]–[12], [14], [16], and references therein) treat problem (P) as an abstract problem with an abstract evolution equation (1) with one control and later derive from abstract Hamilton-Jacobi equations the suitable sufficient optimality conditions for problem (P). We propose an almost direct method to study (P) by a dual dynamic programming approach following the method described in [17] for the one-dimensional case and in [10] for the multidimensional case. We move all notions of dynamic programming to a dual space (the space of multipliers) and then develop a dual dynamic approach together with a dual Bellman-Isaacs equation and, as a consequence, sufficient optimality conditions for (P). We also define an optimal dual feedback control in the terms of which we formulate sufficient conditions for optimality. Such an approach allows us to weaken significantly the assumptions on the data. This paper was motivated first by the paper [15] where minimax control of parabolic systems with Dirichlet boundary conditions is investigated from the point of view of existence of solutions. The second motivation is the following. It is well known that each Hamilton-Jacobi equation can be considered as the fundamental equation (Isaac equation) for the corresponding differential game. This is why many results for Hamilton-Jacobi equation can be interpreted in the framework of differential games and, conversely, many results and methods of differential games have their applications in Hamilton-Jacobi equation. We briefly sketch how the control u appears in (1) in a game theory framework. Let an initial position $(t_0, \cdot, x(t_0, \cdot))$ be given and a function $\mathbf{U} :[t_0, T] \times \Omega \times R \to U$ and a division $\Theta = \{t_0 = \tau_0 < \tau_1 < ... < \tau_{k+1} = T\}$. The control u we define as:

$$u(t, \cdot) = \mathbf{U}(\tau_i, \cdot, x(\tau_i, \cdot)),\ \tau_i \leq t \leq \tau_{i+1}, i = 0, 1, ..., k, \tag{6}$$

together with arbitrarily control $v(t, z) \in \mathbf{V}$ on $(0, T) \times \Gamma$. Both controls define, by (1)–(5), a trajectory with initial condition $x(t_0, \cdot) = x_0(\cdot)$. Denote this solution by $x(t, \cdot, t_0, x_0, \mathbf{U}, v, \Theta)$, $t \in [t_0, T]$. Usually it is called a movement position and the set of all movement positions is extended to a set $\mathbf{X}(t_0, x_0, \mathbf{U}, \Theta)$ of all solutions of (1)–(5) ($x \in W^{2,2}([t_0, T] \times \Omega) \cap C([t_0, T]; L^2(\Omega))$) with $v \in co\mathbf{V}$ and u defined by (6). Of course,

$$\{x(t, \cdot, t_0, x_0, \mathbf{U}, v, \Theta),\ t \in [t_0, T] :\ v \in L^1((0, T) \times \Gamma)\} \subset \mathbf{X}(t_0, x_0, \mathbf{U}, \Theta).$$

Now, let us put

$$\mathbf{X}(t_0, x_0, \mathbf{U}) = \overline{\mathbf{LIM}}_{\text{diam}(\Theta) \to 0} \mathbf{X}(t_0, x_0, \mathbf{U}, \Theta),$$

where

$$\text{diam}(\Theta) = \max\{(\tau_{i+1} - \tau_i) :\ i = 0, 1, ..., k\}.$$

The set $\mathbf{X}(t_0, x_0, \mathbf{U})$ contains elements which are limits $x(t, \cdot)$, $t \in [t_0, T]$ of sequences $x_l(\cdot, \cdot) \in \mathbf{X}(t_0, x_0, \mathbf{U}, \Theta_l)$, $l = 1, 2, ...,$ where $\text{diam}(\Theta_l) \to 0$ if $l \to \infty$. Having that, the function $\mathbf{U} :[t_0, T] \times \Omega \times R \to U$ is called a positioning strategy and elements of $\mathbf{X}(t_0, x_0, \mathbf{U})$ limit movements. Following this way, we come to the theory of differential positioning game and value function (for $\min_u \max_v\ J(x, u, v)$) satisfying suitable Hamilton-Jacobi Equation (see, e.g., [18], [19], [13] compare [4]).

2 A Dual Dynamic Programming

First we describe the idea of a dual dynamic approach to optimal control problems governed by parabolic equations with the second control on the boundary. Let us recall what dynamic programming means. We have an initial condition $(t_0, x_0(t_0, z))$, $z \in \Omega$ for which we assume that we have an optimal solution $(\overline{x}, \overline{u}, \overline{v})$. Then by necessary optimality conditions there exists a conjugate function $\overline{p}(t, z) = (\overline{y}^0, \overline{y}(t, z))$ on $[0, T] \times \Omega$ being a solution to the corresponding adjoint system (see, e.g., [6], [12]). The element $p = (y^0, y)$ plays a role of multipliers from the classical Lagrange problem with constraints (with the multiplier y^0 corresponding to the functional, the y to the constraints). If we perturb (t_0, x_0), then, assuming that the optimal solution exists for each perturbed problem, we also have a conjugate function corresponding to it. Therefore, making perturbations of our initial conditions we obtain two sets of functions: optimal trajectories $\overline{x}$ and corresponding to them conjugate functions $\overline{p}$. The graphs of optimal trajectories cover some set in a state space (t, z, x), say a set X (in the classical calculus of variation it is called a field of extremals), and the graphs of the conjugate functions cover some set in a conjugate space (t,z,p), say a set P (in classical mechanics it is called the space of momentums). In the classical dynamic programming approach we explore the state space (t,z,x), i.e., the set X (see, e.g., [1]), but in the dual dynamic programming approach we explore the conjugate space (the dual space) (t,z,p), i.e., the set P (see [17] for the one-dimensional case and [10] for the multidimensional case). It is worth noting that although in elliptic control optimization problems we have no possibilities to perturb that problem, it is still possible to apply dual dynamic programming (see [11]). It is natural that if we want to explore the dual space (t,z,p) then we need a mapping between the set P and the set X: $P \ni (t,z,p) \to (t,z,\tilde{x}(t,z,p)) \in X$ to have a possibility to formulate, at the end of some consideration in P, any conditions for optimality in our original problem as well as on an optimal solution $\overline{x}$. Of course, such a mapping should have the property that for each admissible trajectory $x(t,z)$ lying in X we must have a function $p(t,z)$ lying in P such that $x(t,z) = \tilde{x}(t,z,p(t,z))$. Hence, we conduct all our investigations in a dual space (t,z,p), i.e., most of our notions concerning the dynamic programming are defined in the dual space including a dynamic programming equation which becomes now a dual dynamic programming equation.

Therefore, let $P \subset R^{n+3}$ be a set of the variables $(t,z,p) = (t,z,y^0,y)$, $(t,z) \in [0,T] \times \Omega$, $y^0 \le 0$, $y \in R$. Let $\tilde{x} : P \to R$ be a function such that for each admissible trajectory $x(t,z)$ there exists a function $p(t,z) = (y^0, y(t,z))$, $p \in W^{2,2}([0,T] \times \Omega)$, $(t,z,p(t,z)) \in P$ satisfying

$$x(t,z) = \tilde{x}(t,z,p(t,z)) \ \text{ for } \ (t,z) \in [0,T] \times \Omega. \tag{7}$$

Now, let us introduce an auxiliary C^2 function $V(t,z,p) : P \to R$ such that for $(t,z,p) \in P$ the following condition is satisfied:

$$V(t,z,p) = y^0 V_{y^0}(t,z,p) + y V_y(t,z,p) = p V_p(t,z,p), \tag{8}$$

$$\begin{aligned} \nabla_z V(t,z,p)\nu\,(z) &= y^0 \nabla_z V_{y^0}(t,z,p)\nu(z), \\ \text{for } (t,z) &\in [0,T] \times \partial\Omega, (t,z,p) \in P, \end{aligned} \tag{9}$$

where $\nu(\cdot)$ is the exterior unit normal vector to $\partial\Omega$ and $\nabla V(t,z,p)$ denotes the gradient of the function $z \to V(t,z,p)$. Condition (8) is a generalization of transversality condition known in classical mechanics as the orthogonality of the momentum to the wave front. Condition (9) has the same meaning but taken on the boundary. Similarly, as in classical dynamic programming, define at $(t,\tilde{p}(\cdot))$, where $\tilde{p}(\cdot) = (\tilde{y}^0, \tilde{y}(\cdot))$ is any function $\tilde{p} \in W^{2,2}(\Omega)$, $(t,z,\tilde{p}(z)) \in P$, a dual value function S_D by the formula

$$\begin{aligned} S_D\left(t,\tilde{p}\left(\cdot\right)\right) := \inf\nolimits_u \sup\nolimits_v \Big\{ & -\tilde{y}^0 \textstyle\int_{[t,T]\times\Omega} L(\tau,z,x(\tau,z),u(\tau,z))d\tau dz \\ & -\tilde{y}^0 \textstyle\int_\Omega l\left(x\left(T,z\right)\right)dz - \tilde{y}^0 \int_{[t,T]\times\partial\Omega} h(\tau,z,v(\tau,z))d\tau dz \Big\}, \end{aligned} \tag{10}$$

where the supremum (infimum) is taken over all admissible pairs $x(\tau,\cdot)$, $v(\tau,\cdot)$, $\tau \in [t,T]$, $(x(\tau,\cdot),\ u(\tau,\cdot),\ \tau \in [t,T])$ satisfying:

$$x(t,z) = \tilde{x}(t,z,\tilde{p}(z)) \ \text{ for } \ z \in \Omega, \tag{11}$$

$$\partial_\nu \tilde{x}(t,z,\tilde{p}(z)) = v\,(t,z) \ \text{ for } \ z \in \partial\Omega, \tag{12}$$

i.e., whose trajectories start at $(t, \tilde{x}(t,\cdot,\tilde{p}(\cdot))$ and for which there exists such a function $p(\tau,z) = (\tilde{y}^0, y(\tau,z))$, $p \in W^{2,2}([t,T]\times\bar{\Omega}) \cap C([t,T];L^2(\Omega))$, $(\tau,z,p(\tau,z)) \in P$, that $x(\tau,z) = \tilde{x}(\tau,z,p(\tau,z))$ for $(\tau,z) \in (t,T)\times\bar{\Omega}$ and

$$y(t,z) = \tilde{y}(z) \ \text{ for } \ z \in \bar{\Omega}. \tag{13}$$

Then, integrating (8) over Ω, for any function $p(z) = (y^0, y(z))$, $p \in W^{2,2}(\Omega)$, $(t,z,p(z)) \in P$, such that $x(\cdot,\cdot)$ satisfying $x(t,z) = \tilde{x}(t,z,p(z))$ for $z \in \overline{\Omega}$, is an admissible trajectory, we also have the equality

$$\begin{aligned} &\textstyle\int_\Omega V(t,z,p(z))dz + \int_{\partial\Omega} \nabla_z V(t,z,p(z))\nu(z)dz \\ &= -\textstyle\int_{\partial\Omega} y(z)\nabla\tilde{x}(t,z,p(z))\nu(z)dz - S_D\,(t,p\,(\cdot)) \end{aligned} \tag{14}$$

with

$$\int_\Omega \tilde{y}^0 V_{y^0}(t,z,p(z))dz + \tilde{y}^0 \int_{\partial\Omega} \nabla_z V_{y^0}(t,z,p(z))\nu(z)dz = -S_D\,(t,p\,(\cdot))\,, \tag{15}$$

and assuming that $\tilde{x}(t,z,p(z)) = -V_y(t,z,p(z))$ for $(t,z) \in [0,T]\times\Omega$, $(t,z,p(z)) \in P$. Denote by the symbol $\Delta_z h$ the sum of the second partial derivatives of the function $h : P \longrightarrow R$ with respect to the variable z_i, $i = 1,...,n$, i.e.,

$$\Delta_z h(t,z,p) := \textstyle\sum_{i=1}^n \left(\partial^2/\partial z_i^2\right) h\,(t,z,p)\,. \tag{16}$$

It turns out that the function $V(t,z,p)$ being defined by (14), (15) satisfies the second-order partial differential equation

$$V_t(t,z,p) + \Delta_z V(t,z,p) + H(t,z,-V_y(t,z,p),p) = 0, \tag{17}$$

where

$$\begin{aligned} H(t,z,x,p) &= y^0 L(t,z,x,u(t,z,p)) + yf(t,z,x,u(t,z,p)), \\ H_\Sigma(t,z,p) &= y^0 h(t,z,v(t,z,p)) \end{aligned} \tag{18}$$

and $u(t,z,p)$, $v(t,z,p)$ are optimal dual feedback controls, and the dual second-order partial differential equation of multidimensional dynamic programming

(DSPDEMDP)

$$\begin{aligned} \sup\nolimits_{u\in U}\{V_t(t,z,p)+\Delta_z V(t,z,p)\\ +y^0L(t,z,-V_y(t,z,p),u)+yf(t,z,-V_y(t,z,p),u)\}=0,\\ \inf\nolimits_{v\in \mathbf{V}}\{-\nabla_z V(t,z,p)\nu(z)+y^0h(t,z,v)\}=0,\\ (t,z)\in(0,T)\times\partial\Omega,(t,z,p)\in P. \end{aligned} \tag{19}$$

Let us note that the function $\tilde{x}(t,z,p)$, introduced a little bit artificially at the beginning of this section, is defined in fact by $-V_y(t,z,p)$, where V is a solution to (19), i.e., knowing the set P and V_y we are able to describe the set X in which our original problem we need to consider. The assumption that the auxiliary function $V(t,z,p)$ is of C^2 is important in this paper and we cannot weaken it. However, we would like to stress that it is only an auxiliary assumption and it is not put on a value function which in our case need not to be even continuous.

Remark. We would like to stress that the duality which is sketched in this section is not a duality in the sense of convex optimization. It is a new nonconvex duality, first described in [17] and next developed in [10], for which we do not have the relation $\sup(D)\le\inf(P)$ (D means a dual problem, P a primal one). But instead of it we have other relations, namely (8) and (14), which are generalizations of transversality conditions from classical mechanics. If we find a solution to (17) then checking the relation (8) for concrete problems is not very difficult.

3 A Verification Theorem

The most important conclusion of dynamic programming is a verification theorem. We present it in a dual form according to our dual dynamic programming approach described in the previous section.

Theorem 3.1. *Let $\overline{x}(t,z),\overline{u}(t,z)$, $(t,z)\in[0,T]\times\bar{\Omega}$, $\bar{v}(t,z)$, $(t,z)\in(0,T)\times\partial\Omega$, be an admissible trio. Assume that there exists a C^2 solution $V(t,z,p)$ of DSPDEMDP* (19) *on P such that* (8), (9) *hold. Let further $\overline{p}(t,z)=(\overline{y}^0,\overline{y}(t,z))$, $\overline{p}\in W^{2,2}([0,T]\times\Omega)\cap C([0,T];L^2(\Omega))$, $(t,z,\overline{p}(t,z))\in P$ be a function such that $\overline{x}(t,z)=-V_y(t,z,\overline{p}(t,z))$ for $(t,z)\in[0,T]\times\Omega$, $\bar{v}(t,z)=-\partial_\nu V_y(t,z,\bar{p}(t,z))$ for $(t,z)\in(0,T)\times\partial\Omega$. Suppose that $V(t,z,p)$ satisfies the boundary condition for $(T,z,p)\in P$,*

$$\overline{y}^0\int_\Omega V_{y^0}(T,z,p)dz=\overline{y}^0\int_\Omega l\left(-V_y\left(T,z,p\right)\right)dz. \tag{20}$$

Moreover, assume that:

$$\begin{aligned} &V_t(t,z,\overline{p}(t,z))+\Delta_z V(t,z,\overline{p}(t,z))+\overline{y}^0L(t,z,-V_y(t,z,\overline{p}(t,z)),\overline{u}(t,z))\\ &+\overline{y}(t,z)f(t,z,-V_y(t,z,\overline{p}(t,z)),\overline{u}(t,z))=0,\ \textit{for } (t,z)\in(0,T)\times\Omega,\\ &-(\nabla_z)\,V(t,z,\overline{p}(t,z))\nu(z)+y^0h(t,z,\bar{v}(t,z))=0,\ \textit{for } (t,z)\in(0,T)\times\partial\Omega. \end{aligned} \tag{21}$$

Then, $\overline{x}(t,z), \overline{u}(t,z)$, $(t,z) \in (0,T) \times \Omega$, $\bar{v}(t,z)$, $(t,z) \in (0,T) \times \partial\Omega$, *is an optimal trio relative to all admissible trios* $x(t,z), u(t,z)$, $(t,z) \in [0,T] \times \Omega$, $v(t,z)$, $(t,z) \in (0,T) \times \partial\Omega$, *for which there exists a function* $p(t,z) = (\overline{y}^0, y(t,z))$, $p \in W^{2,2}([0,T] \times \Omega) \cap C([0,T]; L^2(\Omega))$, $p \in L^2([0,T] \times \partial\Omega)$, $(t,z,p(t,z)) \in P$, *such that* $x(t,z) = -V_y(t,z,p(t,z))$ *for* $(t,z) \in [0,T] \times \Omega$, $v(t,z) = -\partial_\nu V_y(t,z,p(t,z))$ *for* $(t,z) \in (0,T) \times \partial\Omega$ *and*

$$y(0,z) = \overline{y}(0,z) \ \textit{for} \ z \in \Omega, \tag{22}$$

$$y(t,z) = \overline{y}(t,z) \ \textit{for} \ (t,z) \in [0,T] \times \partial\Omega. \tag{23}$$

Proof. Let $x(t,z)$, $u(t,z)$, $(t,z) \in [0,T] \times \Omega, \overline{v}(t,z)$, $(t,z) \in (0,T) \times \partial\Omega$, be an admissible trio for which there exists a function $p(t,z) = (\overline{y}^0, y(t,z))$, $p \in W^{2,2}([0,T] \times \Omega) \cap C([0,T]; L^2(\Omega))$, $p \in L^2([0,T] \times \partial\Omega)$, $(t,z,p(t,z)) \in P$, such that $x(t,z) = -V_y(t,z,p(t,z))$ for $(t,z) \in [0,T] \times \Omega$, $\bar{v}(t,z) = -\partial_\nu V_y(t,z,p(t,z))$ for $(t,z) \in (0,T) \times \partial\Omega$ and (22), (23) are satisfied. From the transversality conditions (8), (9), we obtain that for $(t,z) \in [0,T] \times \Omega$,

$$\begin{aligned} &V_t\,(t,z,p(t,z)) + \Delta_z V(t,z,p(t,z)) \\ &\quad = \overline{y}^0 \left[(d/dt)\, V_{y^0}(t,z,p(t,z)) + \Delta_z V_{y^0}(t,z,p(t,z))\right] \\ &+ y(t,z)\left[(d/dt)\, V_y(t,z,p(t,z)) + \Delta_z V_y(t,z,p(t,z))\right], \end{aligned} \tag{24}$$

and for $(t,z) \in (0,T) \times \partial\Omega$,

$$\begin{aligned} &(\nabla_z)\, V(t,z,p(t,z))\nu\,(z) \\ = \;& \overline{y}^0\,(\nabla_z)\, V_{y^0}(t,z,p(t,z))\nu(z). \end{aligned} \tag{25}$$

Since $x(t,z) = -V_y(t,z,p(t,z))$, for $(t,z) \in [0,T] \times \Omega$, (1) shows that for $(t,z) \in [0,T] \times \Omega$,

$$\begin{aligned} &(d/dt)\, V_y(t,z,p(t,z)) + \Delta_z V_y(t,z,p(t,z)) \\ &\quad = -f(t,z,-V_y(t,z,p(t,z)), u(t,z)) \end{aligned} \tag{26}$$

and boundary control (3) shows that, for $(t,z) \in (0,T) \times \partial\Omega$,

$$-\partial_\nu V_y(t,z,p(t,z)) = \overline{v}(t,z).$$

Now define a function $W(t,z,p(t,z))$ on P by the following requirement for $(t,z) \in [0,T] \times \Omega$,

$$\begin{aligned} W(t,z,p(t,z)) &= \overline{y}^0 \left[(d/dt)\, V_{y^0}(t,z,p(t,z))\right. \\ &\left. + \Delta_z V_{y^0}(t,z,p(t,z)) + L(t,z,-V_y(t,z,p(t,z)), u(t,z))\right] \end{aligned} \tag{27}$$

and for $(t,z) \in (0,T) \times \partial\Omega$,

$$W(t,z,p(t,z)) = -\overline{y}^0\,(\nabla_z)\, V_{y^0}(t,z,p(t,z))\nu\,(z)$$

$$+\overline{y}^0 h(t,z,\overline{v}(t,z)). \tag{28}$$

We conclude from (24)–(27) that for $(t,z)\in[0,T]\times\Omega$,

$$\begin{aligned} W(t,z,p(t,z)) &= V_t\,(t,z,p(t,z)) + \Delta_z V(t,z,p(t,z)) \\ &+\overline{y}^0 L(t,z,-V_y(t,z,p(t,z)),u(t,z)) \\ &+y(t,z)f(t,z,-V_y(t,z,p(t,z)),u(t,z)) \end{aligned} \tag{29}$$

and for $(t,z)\in(0,T)\times\partial\Omega$,

$$\begin{aligned} W(t,z,p(t,z)) &= -\left(\nabla_z\right)V(t,z,p(t,z))\nu\,(z) \\ &+\overline{y}^0 h(t,z,\overline{v}(t,z)); \end{aligned} \tag{30}$$

hence, by (19) and (29), that

$$W(t,z,p(t,z)\le 0 \;\text{ for } (t,z)\in[0,T]\times\Omega \tag{31}$$

and by (23),

$$W(t,z,p(t,z)=0 \;\text{ for } (t,z)\in(0,T)\times\partial\Omega; \tag{32}$$

and finally, after integrating (31) and applying (27), that

$$\begin{aligned} \overline{y}^0\int_{[0,T]\times\Omega}\left[(d/dt)\,V_{y^0}(t,z,p(t,z)) + div\nabla_z V_{y^0}(t,z,p(t,z))\right]dtdz \\ \le -\overline{y}^0\int_{[0,T]\times\Omega}L(t,z,x(t,z),u(t,z))dtdz. \end{aligned} \tag{33}$$

Similarly, in the set $(0,T)\times\partial\Omega$ we have:

$$\begin{aligned} -\overline{y}^0\int_{[0,T]\times\partial\Omega}\left(\nabla_z\right)V_{y^0}(t,z,p(t,z))\nu(z)dtdz \\ = -\overline{y}^0\int_{[0,T]\times\partial\Omega}h(t,z,\overline{v}(t,z))dtdz. \end{aligned} \tag{34}$$

Thus from (33), (20), (22), (23) and by the Green formula it follows that:

$$\begin{aligned} \overline{y}^0\int_{\Omega}\left[l\left(-V_y(T,z,p(T,z))\right) - V_{y^0}(0,z,\overline{y}^0,\overline{y}(0,z))\right]dz \\ +\overline{y}^0\int_{[0,T]}\left(\int_{\partial\Omega}\nabla_z V_{y^0}(t,z,\overline{y}^0,\overline{y}(t,z))\nu(z)dz\right)dt \\ \le -\overline{y}^0\int_{[0,T]\times\Omega}L(t,z,x(t,z),u(t,z))dtdz. \end{aligned} \tag{35}$$

So by (35) and (34) we get:

$$\begin{aligned} -\overline{y}^0\int_{\Omega}V_{y^0}(0,z,\overline{y}^0,\overline{y}(0,z))dz \le -\overline{y}^0\int_{[0,T]\times\Omega}L(t,z,x(t,z),u(t,z))dtdz \\ -\overline{y}^0\int_{\Omega}l(x(T,z))dz - \overline{y}^0\int_{[0,T]\times\partial\Omega}h(t,z,\overline{v}(t,z))dtdz. \end{aligned} \tag{36}$$

In the same manner, applying (21) and (29) we have:

$$W(t,z,\overline{p}(t,z)=0 \ \text{ for } (t,z)\in[0,T]\times\Omega \tag{37}$$

and

$$W(t,z,\overline{p}(t,z)=0 \ \text{ for } (t,z)\in(0,T)\times\partial\Omega.$$

Now from (37), (27), (20), and the Green formula we have:

$$\begin{aligned}
&\overline{y}^0\int_{[0,T]}\left(\int_{\partial\Omega}\nabla_z V_{y^0}(t,z,\overline{y}^0,\overline{y}(t,z))\nu(z)dz\right)dt\\
&-\overline{y}^0\int_\Omega V_{y^0}(0,z,\overline{y}^0,\overline{y}(0,z))dz\\
&=-\overline{y}^0\int_{[0,T]\times\Omega}L(t,z,\overline{x}(t,z),\overline{u}(t,z))dtdz-\overline{y}^0\int_\Omega l(\overline{x}(T,z))dz.
\end{aligned}\tag{38}$$

Combining (36) with (38) gives

$$\begin{aligned}
&-\overline{y}^0\int_{[0,T]\times\Omega}L(t,z,\overline{x}(t,z),\overline{u}(t,z))dtdz-\overline{y}^0\int_\Omega l(\overline{x}(T,z))dz\\
&-\overline{y}^0\int_{[0,T]\times\partial\Omega}h(t,z,\overline{v}(t,z))dtdz\le-\overline{y}^0\int_{[0,T]\times\Omega}L(t,z,x(t,z),u(t,z))dtdz\\
&-\overline{y}^0\int_\Omega l(x(T,z))dz-\overline{y}^0\int_{[0,T]\times\partial\Omega}h(t,z,\overline{v}(t,z))dtdz,
\end{aligned}\tag{39}$$

which gives the right-hand side of the inequality for saddle point.

Starting at the beginning of the proof with an admissible trio $x(t,z),\overline{u}(t,z)$, $(t,z)\in[0,T]\times\Omega$, $v(t,z)$, $(t,z)\in[0,T]\times\partial\Omega$, instead of $x(t,z),u(t,z)$, $(t,z)\in[0,T]\times\Omega$, $\overline{v}(t,z)$, $(t,z)\in[0,T]\times\partial\Omega$, and following the same way as above with the inequality for v in (19), we come to the second inequality we need for saddle point:

$$\begin{aligned}
&-\overline{y}^0\int_{[0,T]\times\Omega}L(t,z,x(t,z),\overline{u}(t,z))dtdz-\overline{y}^0\int_\Omega l(x(T,z))dz\\
&-\overline{y}^0\int_{[0,T]\times\partial\Omega}h(t,z,v(t,z))dtdz\le-\overline{y}^0\int_{[0,T]\times\Omega}L(t,z,\overline{x}(t,z),\overline{u}(t,z))dtdz\\
&-\overline{y}^0\int_\Omega l(\overline{x}(T,z))dz-\overline{y}^0\int_{[0,T]\times\partial\Omega}h(t,z,\overline{v}(t,z))dtdz.
\end{aligned}\tag{40}$$

□

4 An Optimal Dual Feedback Control

In practice, a feedback control is more important than a value function. It turns out that the dual dynamic programming approach allows one to also investigate a kind of a feedback control which we call a dual feedback control. Surprisingly, it can have better properties than the classical one— now our state equation depends only on the parameter and not additionally on the state in a feedback function, which makes the state equation difficult to solve.

Definition 4.1. A pair of functions $\tilde{u} = \tilde{u}(t,z,p)$, from P of the points $(t,z,p) = (t,z,y^0,y)$, $(t,z) \in (0,T) \times \Omega$, $y^0 \leq 0$, $y \in R$, into U and $\tilde{v}(t,z,p)$ from a subset P of those points $(t,z,p) = (t,z,y^0,y)$, $(t,z) \in (0,T) \times \partial\Omega$, $(t,z,p) \in P$, into $\mathbf{V}$ is called a dual feedback control, if there is any solution $\tilde{x}(t,z,p)$, $(t,z,p) \in P$, of the partial differential equation

$$x_t(t,z,p) + \Delta_z x(t,z,p) = f(t,z,x(t,z,p),\tilde{u}(t,z,p)) \tag{41}$$

satisfying Neumann boundary condition

$$\partial_\nu \tilde{x}(t,z,p) = \tilde{v}\,(t,z,p) \;\text{ on }\; (0,T) \times \Gamma,\, (t,z,p) \in P,$$

such that for each admissible trajectory $x(t,z)$, $(t,z) \in [0,T] \times \Omega$, there exists a function $p(t,z) = (y^0, y(t,z))$, $p \in W^{2,2}([0,T] \times \Omega) \cap C([0,T]; L^2(\Omega))$, $p \in L^2([0,T] \times \partial\Omega)$, $(t,z,p(t,z)) \in P$, such that (7) holds.

Definition 4.2. A dual feedback control $\overline{u}(t,z,p)$, $\overline{v}(t,z,p)$ is called an optimal dual feedback control, if there exist a function $\overline{x}(t,z,p)$, $(t,z,p) \in P$, corresponding to $\overline{u}(t,z,p)$, $\overline{v}(t,z,p)$ as in Definition 4.1, and a function $\overline{p}(t,z) = (\overline{y}^0, \overline{y}(t,z))$, $\overline{p} \in W^{2,2}([0,T] \times \Omega) \cap C([0,T]; L^2(\Omega))$, $\overline{p} \in L^2([0,T] \times \partial\Omega)$, $(t,z,\overline{p}(t,z)) \in P$, $(t,z,\overline{p}(t,z)) \in P$, such that dual value function S_D (see (10)) is defined at $(t,\overline{p}\,(t,\cdot))$ by $\overline{u}(\tau,z,p)$, $\overline{v}(\tau,z,p)$ and corresponding to them $\overline{x}(\tau,z,p)$, $(\tau,z,p) \in P$, $\tau \in [t,T]$, i.e.,

$$\begin{aligned}
&S_D(t,\overline{p}\,(t,\cdot)) = \\
&-\overline{y}^0 \textstyle\int_{[t,T]\times\Omega} L(\tau,z,\overline{x}(\tau,z,\overline{p}(\tau,z)),\overline{u}(\tau,z,\overline{p}(\tau,z))))d\tau dz \\
&-\overline{y}^0 \textstyle\int_{\Omega} l\left(\overline{x}(T,z,\overline{p}(T,z))\right) dz - y^0 \int_{[t,T]\times\partial\Omega} h(\tau,z,\overline{v}(\tau,z,\overline{p}(\tau,z)))d\tau dz
\end{aligned} \tag{42}$$

and, moreover, there is $V(t,z,p)$ satisfying (8) and (9) for which V_{y^0} satisfies the equality

$$\begin{aligned}
&\overline{y}^0 \int_\Omega V_{y^0}(t,z,\overline{p}(t,z))dz + \overline{y}^0 \int_{\partial\Omega} (\nabla_z)\, V_{y^0}(t,z,\overline{p}(t,z))\nu(z)dz \\
= \;& -S_D\,(t,\overline{p}\,(t,\cdot))
\end{aligned} \tag{43}$$

and V_y satisfies

$$\begin{aligned}
V_y(t,z,p) &= -\overline{x}(t,z,p) \;\text{ for }\; (t,z) \in (0,T) \times \Omega, (t,z,p) \in P, \\
V_y(t,z,p) &= -\partial_\nu \overline{x}(t,z,p) \;\text{ for }\; (t,z) \in (0,T) \times \partial\Omega, (t,z,p) \in P.
\end{aligned} \tag{44}$$

The next theorem is nothing more that the verification theorem formulated in terms of a dual feedback control.

Theorem 4.1. *Let $\overline{u}(t,z,p)$, $\overline{v}(t,z,p)$ be a dual feedback control in P. Suppose that there exists a C^2 solution $V(t,z,p)$ of DSPDEMDP* (19) *on P such* *that* (8) *and* (20) *hold. Let* $\overline{p}(t,z) = (\overline{y}^0, \overline{y}(t,z))$, $\overline{p} \in W^{2,2}([0,T]\times\Omega) \cap C([0,T];L^2(\Omega))$, $\overline{p} \in L^2([0,T]\times\partial\Omega)$, $(t,z,\overline{p}(t,z)) \in P$ *be a function such that* $(\overline{x}(t,z), \overline{u}(t,z))$, *where* $\overline{x}(t,z) = \overline{x}(t,z,\overline{p}(t,z))$ *and* $\overline{u}(t,z) = \overline{u}(t,z,\overline{p}(t,z))$, $(t,z) \in [0,T]\times\Omega$, $\bar{v}(t,z) = \overline{v}(t,z,\overline{p}(t,z))$, $(t,z) \in (0,T)\times\partial\Omega$, *is an admissible trio with* $\overline{x}(t,z,p)$, $(t,z,p) \in P$, *corresponding to* $\overline{u}(t,z,p)$, $\overline{v}(t,z,p)$ *as in Definition 4.1. Assume further that* V_y *and* V_{y^0} *satisfy:*

$$\begin{aligned} V_y(t,z,p) &= -\overline{x}(t,z,p) \;\; \textit{for} \;\; (t,z,p) \in P, \\ V_y(t,z,p) &= -\partial_\nu \overline{x}(t,z,p) \;\; \textit{for} \;\; (t,z) \in (0,T)\times\partial\Omega, (t,z,p) \in P, \end{aligned} \tag{45}$$

$$\begin{aligned} &\overline{y}^0 \int_\Omega V_{y^0}(t,z,\overline{p}(t,z))dz \\ &+\overline{y}^0 \int_{[0,T]} \left(\int_{\partial\Omega} (\nabla_z)\, V_{y^0}(t,z,\overline{y}^0,\overline{y}(t,z))\nu(z)dz \right) dt \\ &= -\overline{y}^0 \int_{[0,T]\times\Omega} L(t,z,\overline{x}(t,z,\overline{p}(t,z)),\overline{u}(t,z,\overline{p}(t,z)))dtdz \\ &\qquad -\overline{y}^0 \int_\Omega l\left(\overline{x}(T,z,\overline{p}(T,z))\right) dz \\ &-\overline{y}^0 \int_{[0,T]\times\partial\Omega} h(t,z,\overline{v}(t,z,\overline{p}(t,z)))dtdz. \end{aligned} \tag{46}$$

Then $\overline{u}(t,z,p)$, $\overline{v}(t,z,p)$ *is an optimal dual feedback control.*

Proof. Take any function

$$p(t,z) = (\overline{y}^0, y(t,z)),\; p \in W^{2,2}([0,T]\times\Omega) \cap C([0,T];L^2(\Omega)),$$

$$p \in L^2([0,T]\times\partial\Omega), (t,z,p(t,z)) \in P,$$

such that $x(t,z) = \overline{x}(t,z,p(t,z))$, $u(t,z) = \overline{u}(t,z,p(t,z))$, $(t,z) \in (0,T)\times\Omega$, $v(t,z) = \overline{v}(t,z,p(t,z))$, $(t,z) \in [0,T]\times\partial\Omega$, is an admissible trio and (22), (23) hold. By (45), it follows that $x(t,z) = -V_y(t,z,p(t,z))$ for $(t,z) \in [0,T]\times\Omega$. As in the proof of Theorem 3.1, (46) gives:

$$\begin{aligned} &-\overline{y}^0 \int_{[0,T]\times\Omega} L(t,z,\overline{x}(t,z,\overline{p}(t,z)),\overline{u}(t,z,\overline{p}(t,z)))dtdz \\ &-\overline{y}^0 \int_\Omega l\left(\overline{x}(T,z,\overline{p}(T,z))\right) dz - y^0 \int_{[0,T]\times\partial\Omega} h(t,z,\overline{v}(t,z,\overline{p}(t,z)))dtdz \leq \\ &\qquad -\overline{y}^0 \int_{[0,T]\times\Omega} L(t,z,\overline{x}(t,z,p(t,z)),\overline{u}(t,z,p(t,z))dtdz \\ &-\overline{y}^0 \int_\Omega l(\overline{x}(T,z,p(T,z)))dz - y^0 \int_{[0,T]\times\partial\Omega} h(t,z,\overline{v}(t,z,\overline{p}(t,z)))dtdz, \end{aligned} \tag{47}$$

and similarly,

$$\begin{aligned} &-\overline{y}^0 \int_{[0,T]\times\Omega} L(t,z,\overline{x}(t,z,p(t,z)),\overline{u}(t,z,\overline{p}(t,z)))dtdz \\ &-\overline{y}^0 \int_\Omega l\left(\overline{x}(T,z,p(T,z))\right) dz - y^0 \int_{[0,T]\times\partial\Omega} h(t,z,\overline{v}(t,z,p(t,z)))dtdz \leq \\ &\qquad -\overline{y}^0 \int_{[0,T]\times\Omega} L(t,z,\overline{x}(t,z,\overline{p}(t,z)),\overline{u}(t,z,\overline{p}(t,z))dtdz \\ &-\overline{y}^0 \int_\Omega l(\overline{x}(T,z,\overline{p}(T,z)))dz - y^0 \int_{[0,T]\times\partial\Omega} h(t,z,\overline{v}(t,z,\overline{p}(t,z)))dtdz. \end{aligned}$$

We conclude from (47) that:

$$\begin{aligned} &S_D(0,\overline{p}\,(0.\cdot)) = \\ &-\overline{y}^0 \int_{[0,T]\times\Omega} L(t,z,\overline{x}(t,z,\overline{p}(t,z)),\overline{u}(t,z,\overline{p}(t,z)))dtdz \\ &-\overline{y}^0 \int_\Omega l\left(\overline{x}(T,z,\overline{p}(T,z))\right)dz - y^0 \int_{[0,T]\times\partial\Omega} h(t,z,\overline{v}(t,z,\overline{p}(t,z)))dtdz\ , \end{aligned} \tag{48}$$

and it is sufficient to show that $\overline{u}(t,z,p)$, $\overline{v}(t,z,p)$ is an optimal dual feedback control, by Theorem 3.1 and Definition 4.2. □

References

[1] Barbu V. *Analysis and control of nonlinear infinite dimensional systems*. Academic Press, Boston 1993.

[2] Barbu V. *The dynamic programming equation for the time-optimal control problem in infinite dimensions*, SIAM J. Control Optim. **29** (1991), 445-456.

[3] Barbu V. and Da Prato G. *Hamilton-Jacobi equations in Hilbert spaces*. Pitman Advanced Publishing Program, Boston 1983.

[4] Berkovitz L.D. *Differential games of generalized pursuit and evasion*, SIAM J. Control Optim. **24** (1986), 24:361–373.

[5] Cannarsa P. and Carja O. *On the Bellman equation for the minimum time problem in infinite dimensions*, SIAM J. Control Optim. **43** (2004), 532-548.

[6] Casas E. *Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations*, SIAM J. Control Optim. **35** (1997), 1297–1327.

[7] Fattorini H.O. *Infinite-Dimensional Optimization and Control Theory*. Cambridge University Press, Cambridge (1999).

[8] Fattorini H.O. and Murphy T. *Optimal control for nonlinear parabolic boundary control systems: the Dirichlet boundary conditions*, Diff. Integ. Equ. **7** (1994), 1367–1388.

[9] Fursikov A.V. *Optimal control of distributed systems. Theory and applications*. American Mathematical Society. Providence, RI (2000).

[10] Galewska E. and Nowakowski A. *Multidimensional dual dynamic programming*, J. Optim. Theor. Applic. **124** (2005), 175-186.

[11] Galewska E. and Nowakowski A. *A dual dynamic programming for multidimensional elliptic optimal control problems*, Num. Func. Anal. Optim. **27** (2006), 279-280.

[12] Gozzi F. and Tessitore ME. *Optimality conditions for Dirichlet boundary control problems of parabolic type*, J. Math. Syst. Estim. Contr. **8** (1998), 143–146.

[13] Krasovskii N.N. and Subbotin A.I. *Game Theoretical Control Problems*, Springer-Verlag, New York (1987).

[14] Li X. and Yong J. *Optimal control theory for infinite dimensional systems*. Birkhauser, Boston (1994).

[15] Mordukhovich B.S. and Zhang K. *Minimax control of parabolic systems with Dirichlet boundary conditions and state constraints*, Appl. Math. Optim. **36** (1997), 323–360.

[16] Neittaanmaki P. and Tiba D. *Optimal control of nonlinear parabolic systems*. Marcel Dekker, New York (1994).

[17] Nowakowski A. *The dual dynamic programming*, Proc. Amer. Math. Soc. **116** (1992), 1089–1096.

[18] Subbotin A.I. *Generalized solutions of first-order PDEs. The dynamical optimization perspective* Translated from the Russian. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, (1995).

[19] Subbotin A.I. *Constructive theory of positional differential games and generalized solutions to Hamilton-Jacobi equations. Stochastic and differential games*, Ann. Internat. Soc. Dynam. Games, **4**, Birkhäuser Boston, Boston, (1999), 3–67.

Non-Cooperative and Semi-Cooperative Differential Games

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

Abstract

In this paper we review some recent results on non-cooperative and semi-cooperative differential games. For the n-person non-cooperative games in one-space dimension, we consider the Nash equilibrium solutions. When the system of Hamilton-Jacobi equations for the value functions is strictly hyperbolic, we show that the weak solution of a corresponding system of hyperbolic conservation laws determines an n-tuple of feedback strategies. These yield a Nash equilibrium solution to the non-cooperative differential game.

However, in the multi-dimensional cases, the system of Hamilton-Jacobi equations is generically elliptic, and therefore ill posed. In an effort to obtain meaningful stable solutions, we propose an alternative "semi-cooperative" pair of strategies for the two players, seeking a Pareto optimum instead of a Nash equilibrium. In this case, the corresponding Hamiltonian system for the value functions is always weakly hyperbolic.

Key words. Non-cooperative differential games, Nash equilibrium, system of Hamilton-Jacobi equations, hyperbolic system of conservation laws, BV solutions, optimal control theory, discontinuous ODE, ill-posed Cauchy problem.

AMS Subject Classifications. Primary 91A23, 49N70, 93B52, 35L65; Secondary 91A10, 49N90, 49N35, 49L20, 34A36.

1 Introduction

In this paper we review some recent results on non-cooperative differential games. Non-cooperative games provide a mathematical model for the behavior of two or more individuals, operating in the same environment with different (possibly conflicting) goals. In the case of n players, the evolution of the system is governed by system of differential equations of the form

$$\dot{x}(t) = \sum_{i=1}^{n} f_i(x, u_i), \qquad x(\tau) = y. \tag{1}$$

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_5,

Here the real-valued map $t \mapsto u_i(t)$ is the control implemented by the i-th player. Together with (1) we consider the payoff functionals:

$$J_i = J_i(\tau, y, u_1, \ldots, u_n) \doteq g_i(x(T)) - \int_\tau^T h_i(x(t), u_i(t))\, dt, \qquad (2)$$

Notice that (2) is the sum of a terminal payoff g_i, depending on the state of the system at the final time T, and of a running cost h_i, incurred while implementing the control u_i. The goal of the i-th player is to maximize J_i.

To see an example in economy, one can consider n companies which are planning to sell the same type of product. Let $x_i(t)$ be the market share of the i-th company at time t. This can change in time, depending on the level of advertising $(u_1, \ldots, u_n)$ chosen by the various companies. At the final time T when the products are being sold, the payoff J_i of the i-th company will depend on its market share $x_i(T)$ and on its total advertising cost, as in (2).

A major step toward the understanding of non-cooperative games with several players was provided by the concept of Nash non-cooperative equilibrium, introduced by J. Nash [26]. Roughly speaking, a set of strategies $(U_1^*, \ldots, U_n^*)$ constitutes a *Nash equilibrium* if, whenever one single player modifies his strategy (and the other players do not change theirs), his own payoff will not increase. This concept was first formulated in the context of static games, where no time-evolution is involved. It is natural to explore the relevance of Nash equilibria also in connection with differential games.

Results on the existence of Nash equilibrium solutions for open-loop strategy can be found in [17,32]. In this case, each player has knowledge only of the initial state of the system. His strategy is thus a function of time only, say $u_i = U_i(t)$.

In this paper, we analyze the existence and stability of Nash equilibrium strategies in feedback (closed-loop) form. Here, the players can directly observe the state $x(t)$ of the system at every time $t \in [0, T]$, therefore their strategies depend on x. More precisely, an n-tuple of feedback strategies

$$u_i = U_i^*(t, x), \qquad i = 1, \ldots, n,$$

is called a *Nash equilibrium solution* if the following holds. For each i, if the i-th player chooses an alternative strategy U_i, while every other player $j \neq i$ sticks to his previous strategy U_j^*, then the i-th payoff does not increase:

$$\begin{aligned} &J_i(\tau, y, U_1^*, \ldots, U_{i-1}^*, U_i, U_{i+1}^*, \ldots, U_n^*) \\ &\qquad \leq\ J_i(\tau, y, U_1^*, \ldots, U_{i-1}^*, U_i^*, U_{i+1}^*, \ldots, U_n^*). \end{aligned}$$

Therefore, for the i-th player, the feedback strategy $u_i = U_i^*(t, x)$ provides the solution to the optimal control problem

$$\max_{u_i(\cdot)} \left\{ g_i(x(T)) - \int_\tau^T h_i(x(t), u_i(t))\, dt \right\}, \qquad (3)$$

in connection with the system

$$\dot{x} = f_i(x, u_i) + \sum_{j \neq i} f_j(x, U_j^*(t, x)) \,. \tag{4}$$

Assuming that the dynamics of the system and the payoff functions are sufficiently regular, the problem can be attacked using tools from P.D.E. theory. As in the theory of optimal control, the basic objects of our study are the *value functions* V_i. Roughly speaking, $V_i(\tau, y)$ denotes the payoff expected by the i-th player, if the game were to start at time τ, with the system in the state $x(\tau) = y$. As shown in [17, p. 292], these value functions satisfy a system of first order partial differential equations with terminal data:

$$\frac{\partial}{\partial t} V_i + H_i(x, \nabla V_1, \ldots, \nabla V_n) = 0, \qquad V_i(T, x) = g_i(x), \quad i = 1, \ldots, n. \tag{5}$$

In the case of a two-person, zero-sum differential game, the value function is obtained from the scalar Bellman-Isaacs equation [17]. The analysis can thus rely on comparison principles and on the well-developed theory of viscosity solutions for Hamilton-Jacobi equations; for example, see [2]. In our case, one has to study a highly nonlinear system of Hamilton-Jacobi equations. Previous results in this direction include only particular examples as in [3,13,14,27].

In the one-dimensional case, differentiating (5) one obtains a system of conservation laws for the gradient functions $p_i \doteq V_{i,x}$, namely:

$$p_{i,t} + H_i(x, p)_x = 0. \tag{6}$$

Under the assumption on strict hyperbolicity (which is discussed in more detail in Sec. 2 the known results on systems of conservation laws can be applied. The theorem of Glimm [19] or its more general versions [4,21,24] provide then the existence of a global solution to the Hamilton-Jacobi equations for terminal data g_i whose gradients have sufficiently small total variation.

The Nash feedback strategies can then be recovered from the gradients of the value functions. Establishing the optimality of this feedback strategy is a non-trivial task due to lack of regularity. In Sec. 2, we prove the optimality by using the special structure of solutions of hyperbolic systems of conservation laws.

However, when the state space is multi-dimensional, the corresponding system of P.D.E's is generically not hyperbolic, and the Cauchy problem is not well posed. In Sec. 3, we study in detail a particular one-dimensional example where hyperbolicity fails, and construct a family of unstable, highly oscillatory solutions. Our conclusion is that the concept of Nash equilibrium is not appropriate for the study of feedback strategies for differential games in continuous time. Indeed, solutions are extremely sensitive to small perturbations of the data, so that the mathematical model has no predictive power.

To readdress the situation, one possibility is to introduce some stochasticity in the system; see [18,25] and the references therein. The presence of random inputs, in the form of white noise, has a well-known stabilizing effect since it transforms the system into a parabolic one. Another possibility, explained in more detail in Sec. 3, is to allow some degree of cooperation among the players. As proved by Smale in connection with the repeated prisoner's dilemma [31], even if the players do not communicate with each other, over a period of time they can devise strategies converging to a Pareto optimum. In the setting of differential games we prove that, if these semi-cooperative strategies are implemented, then the system of P.D.E's for the value functions turns out to be always hyperbolic, at least in a weak sense. Partial cooperation thus removes the most severe instabilities found among Nash non-cooperative equilibrium solutions.

2 Feedback Nash Equilibrium to Non-Cooperative Differential Games

Consider a differential game for n players in one space dimension, with the simple form:

$$\dot{x} = f_0 + \sum_i u_i, \qquad x(\tau) = y\,. \tag{1}$$

Here the controls u_i can be any measurable, real-valued functions, while $f_0 \in \mathbb{R}$ is a fixed constant. The payoff functionals are given by:

$$J_i = J_i(\tau, y, u_1, \ldots, u_n) = g_i(x(T)) - \int_\tau^T h_i(u_i(t))\,dt\,. \tag{2}$$

A key assumption, used throughout the paper, is that the cost functions h_i are smooth and strictly convex, with a positive second derivative

$$\frac{\partial^2}{\partial\omega^2} h_i(\omega) > 0. \tag{3}$$

The Hamiltonian functions H_i are thus defined as follows. By (3), for any $j = 1, \ldots, n$ and any given gradient vector $p_j = \nabla V_j \in \mathbb{R}^m$, there exist a unique optimal control value $u_j^*(p_j)$ such that:

$$p_j \cdot u_j^*(p_j) - h_j(u^*(p_j)) = \max_{\omega\in\mathbb{R}} \{p_j \cdot \omega - h_j(\omega)\} \doteq \phi_j(p_j)\,. \tag{4}$$

Then

$$H_i(p_1, \ldots, p_n) = p_i \cdot \left(f_0 + \sum_j u_j^*(p_j) \right) - h_i(u_i^*(p_i)). \tag{5}$$

The corresponding Hamilton-Jacobi equation for V_i takes the form

$$V_{i,t} + H_i(V_{1,x}, \ldots, V_{n,x}) = 0, \tag{6}$$

with data given at the terminal time $t = T$:

$$V_i(T, x) = g_i(x), \qquad i = 1, \dots, n. \tag{7}$$

In turn, the gradients $p_i \doteq V_{i,x}$ of the value functions satisfy the system of conservation laws with terminal data

$$\frac{\partial}{\partial t} p_i + \frac{\partial}{\partial x} H_i(p_1, \dots, p_n) = 0, \qquad p_i(T, x) = g_i'(x)\,. \tag{8}$$

In recent years, considerable progress has been achieved in the understanding of weak solutions to hyperbolic systems of conservation laws in one-space dimension. In particular, entropy admissible solutions with small total variation are known to be unique and depend continuously on the initial data [7,8]. Moreover, they can be obtained as the unique limits of vanishing viscosity approximations [4]. We apply these new results to prove the existence and stability of Nash equilibrium solutions, in the context of differential games.

The key question is whether this system of conservation laws admits a solution. Moreover, is this solution unique? How is it affected by small perturbations of the data? Classical P.D.E. theory provides conditions under which the Cauchy problem is “well posed”, i.e., it admits a unique solution depending continuously on the initial data. The basic requirement is that the system should be hyperbolic. For a given system of P.D.E’s, hyperbolicity amounts to an algebraic condition on the matrices of coefficients and can be checked in practice.

2.1 Hyperbolicity conditions

Now we describe the hyperbolicity conditions for the system of conservation laws (8). The Jacobian matric $A(p)$ of this system, with entries $A_{ij} = \partial H_i / \partial p_j$, takes the form:

$$A(p) = \begin{pmatrix} \dot{x} & p_1\phi_2'' & p_1\phi_3'' & \cdots & p_1\phi_n'' \\ p_2\phi_1'' & \dot{x} & p_2\phi_3'' & \cdots & p_2\phi_n'' \\ p_3\phi_1'' & p_3\phi_2'' & \dot{x} & \cdots & p_2\phi_n'' \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_n\phi_1'' & p_n\phi_2'' & p_n\phi_3'' & \cdots & \dot{x} \end{pmatrix}, \tag{9}$$

where $\phi_j(p_j)$ is defined in (4) and $\phi_j''(p_j) = -(h_j''(u_j^*(p_j)))^{-1}$ is always negative. The system (8) is strictly hyperbolic at a point $p = (p_1, \dots, p_n)$ if the Jacobian matrix $A(p)$ has n real distinct eigenvalues. Our first result provides a sufficient condition for this to happen.

Lemma 2.1. *Assume that all components p_i, $i = 1, \dots, n$, have the same sign. Moreover, assume that there are no distinct indices $i \neq j \neq k$ such that:*

$$p_i\phi_i'' = p_j\phi_j'' = p_k\phi_k''. \tag{10}$$

Then the system (8) *is strictly hyperbolic at* p*. Moreover, all eigenvalues* $\lambda_i(p)$ *of the matrix* $A(p)$ *satisfy the inequality:*

$$\lambda_i(p) \neq \dot{x} = f_0 + \sum_j u_j^*(p_j) \qquad i = 1, \ldots, n. \tag{11}$$

We note that if all the p_i have the same sign, the eigenvalues will be real. The further condition (10) ensures that they are all distinct. If this condition fails, say, we have $p_{i-1}\phi''_{i-1} = p_i\phi''_i = p_{i+1}\phi''_{i+1}$, then $-p_i\phi''_i$ becomes a multiple zero of $\det(B - \lambda I)$. In this case, the system of conservation laws is only called hyperbolic (not strictly hyperbolic). Furthermore, from (11), the characteristic wave speeds λ_i are all different from the speed $\dot{x}$ at which the state of the system changes. For a proof of the above result, see [10]. We mention that, in the case of two-player games, the condition $p_1 p_2 > 0$ is also necessary for the strict hyperbolicity of the system. However, one can give an example of a three-player game, where the system (8) is strictly hyperbolic even without p_1, p_2, p_3 having all the same sign.

2.2 Solutions of the hyperbolic system

Next, assume that the system of conservation laws (8) is strictly hyperbolic in a neighborhood of a point $p^* = (p_1^*, \ldots, p_n^*)$. In this case, assuming that the terminal conditions have small total variation, one can apply the following theorem (cf. [4,7,8,19]) and obtain the global existence and uniqueness of a weak solution.

Proposition 1. Assume that the flux function $H : I\!R^n \mapsto I\!R^n$ is smooth and that, at some point p^*, the Jacobian matrix $A(p^*) = DF(p^*)$ has n real distinct eigenvalues. Then there exists $\delta > 0$ for which the following holds. If

$$\|\bar{p}(\cdot) - p^*\|_{\mathbf{L}^\infty} < \delta, \qquad \text{Tot.Var.}\{\bar{p}\} < \delta, \tag{12}$$

then the Cauchy problem

$$p_t + H(p)_x = 0, \qquad p(0, x) = \bar{p}(x) \tag{13}$$

admits a unique entropy weak solution $p = p(t, x)$ defined for all $t \geq 0$, obtained as the limit of vanishing viscosity approximations.

2.3 Optimal trajectory; solutions of a discontinuous O.D.E.

In general, a weak solution of the hyperbolic system of conservation laws (8) uniquely determines a family of discontinuous feedback controls $U_i^* = U_i^*(t, x)$. Inserting these feedback controls in (1), we obtain the O.D.E. for the optimal trajectory:

$$\dot{x}(t) = f_0 + \sum_{i=1}^n u_i^*(p_i(t, x)), \qquad x(\tau) = y. \tag{14}$$

Note that the right-hand side of this ODE is discontinuous, due to the discontinuities in the feedback controls $U_i^* = U_i^*(t,x)$. In spite of this, the solution of the Cauchy problem (14) is unique and depends continuously on the initial data, thanks to the special structure of the BV solutions of hyperbolic systems of conservation laws. Indeed, every trajectory of (14) crosses transversally all lines of discontinuity in the functions p_i. Because of the bound on the total variation, the uniqueness result in [6] can thus be applied. We explain the ideas below.

First, we observe that the solution $p = p(t,x)$ of (8) has bounded directional variation along a cone Γ, strictly separated from all characteristic directions. Indeed, by assumption, the matrix $A(p^*)$ has distinct eigenvalues $\lambda_1^* < \lambda_2^* < \cdots < \lambda_n^*$. By continuity, there exists $\varepsilon > 0$ such that, for all p in the ε-neighborhood

$$\Omega_\varepsilon^* \doteq \{p;\ |p - p^*| \leq \varepsilon\},$$

the characteristic speeds range inside disjoint intervals

$$\lambda_j(p) \in [\lambda_j^-, \lambda_j^+]. \tag{15}$$

Moreover, if $p^-, p^+ \in \Omega_\varepsilon^*$ are two states connected by a j-shock, the speed $\lambda_j(p^-, p^+)$ of the shock remains inside the interval $[\lambda_j^-, \lambda_j^+]$.

Now consider an open cone of the form

$$\Gamma \doteq \{(t,x);\ t > 0,\ a < x/t < b\}. \tag{16}$$

Following [6], we define the *directional variation* of the function $(t,x) \mapsto p(t,x)$ along the cone Γ as:

$$\sup \left\{ \sum_{i=1}^{N} |p(t_i, x_i) - p(t_{i-1}, x_{i-1})| \right\}, \tag{17}$$

where the supremum is taken over all finite sequences $(t_0, x_0), (t_1, x_1), \ldots, (t_N, x_N)$ such that:

$$(t_i - t_{i-1}, x_i - x_{i-1}) \in \Gamma \quad \text{for every } i = 1, \ldots, N; \tag{18}$$

see Fig. 1. The next lemma shows that the weak solution $p = p(t,x)$ has bounded directional variation along a suitable cone Γ.

Lemma 2.2. *Let $p = p(t,x)$ be an entropy weak solution of* (13) *taking values inside the domain Ω_ε^*. Assume that $\lambda_{k-1}^+ < a < b < \lambda_k^-$ for some k. Then p has bounded directional variation along the cone Γ in* (16).

See [10] for a detailed proof. Together with Γ we now consider a strictly smaller cone, say,

$$\Gamma' \doteq \{(t,x);\ t > 0,\ a' < x/t < b'\}, \tag{19}$$

with $a < a' < b' < b$. A standard theorem in real analysis states that a BV function of one real variable admits left and right limits at every point. An analogous result for functions with bounded directional variation is proved in [10].

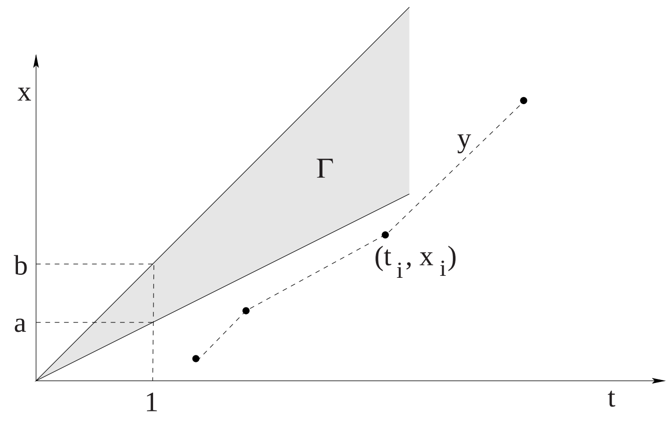

Figure 1: Directional variation along the cone Γ.

Lemma 2.3. *Let $p = p(t,x)$ be a function with bounded directional variation along the cone Γ in* (16)*, and consider the smaller cone $\Gamma' \subset \Gamma$ in* (19)*, with $a < a' < b' < b$. Then at every point $P = (t,x)$ there exist the directional limits:*

$$p^+(P) \doteq \lim_{Q \to P,\ Q-P \in \Gamma'} p(Q), \qquad p^-(P) \doteq \lim_{Q \to P,\ P-Q \in \Gamma'} p(Q). \tag{20}$$

Now, due to the transversality condition and the bound on directional total variation, the result in [6] can be applied, providing the uniqueness and continuous dependence of trajectories of (14). We refer to [10] for more details.

2.4 Optimal feedback strategies

We see that from the gradients of the value functions one can recover the Nash feedback strategies for the various players. To obtain an existence result for solutions of differential games, one has to show that, for each single player, the feedback strategy corresponding to the solution of the Hamilton-Jacobi system actually provides the optimal solution to the control problem (3)–(4). We remark that, if the value functions V_i were smooth, the optimality would be an immediate consequence of the equations. The main technical difficulty stems from the non-differentiability of these value functions.

In the literature on control theory, sufficient conditions for optimality have been obtained along two main directions. On one hand, there is the “regular synthesis” approach developed by Boltianskii [5], Brunovskii [12], and Sussmann and Piccoli [29]. In this case, one typically requires that the value function be piecewise $\mathcal{C}^1$ and satisfy the Hamilton-Jacobi equations outside a finite or countable number of smooth manifolds $\mathcal{M}_i$. On the other hand, one can use the Crandall-Lions theory of viscosity solutions, and show that the value function is the unique solution of the Hamilton-Jacobi equation in the viscosity sense [2].

None of these approaches is applicable in the present situation because of lack of regularity, for both the value functions and the system itself. Indeed, each player now has to solve an optimal control problem for a system whose dynamics (determined by the feedbacks used by all other players) is discontinuous. Our proof of optimality strongly relies on the special structure of BV solutions of hyperbolic systems of conservation laws. In particular, the solution has bounded directional variation along a cone Γ bounded away from all characteristic directions. As a consequence, the value functions V_i always admit a directional derivative in the directions of the cone Γ. For trajectories whose speed remains inside Γ, the optimality can thus be tested directly from the equations. An additional argument, using Clarke's generalized gradients [15], rules out the optimality of trajectories whose speed falls outside the above cone of directions. The following Theorem is proved in [10].

Theorem 2.1. *Consider the differential game* (1)–(2), *where the cost functions* h_i *are smooth and satisfy the convexity assumption* (3). *In connection with the functions* ϕ_j *at* (4), *let* $p^* \doteq (p_1^*, \ldots, p_n^*)$ *be a point where the assumptions of Lemma* 2.1 *are satisfied. Then there exists* $\delta > 0$ *such that the following holds. If*

$$\|g_i' - p_i^*\|_{\mathbf{L}^\infty} < \delta\,, \qquad \textit{Tot.Var.}\{g_i'(\cdot)\} < \delta\,, \qquad\qquad i = 1, \ldots, n\,, \tag{21}$$

then for any $T > 0$ *the terminal value problem* (8) *has a weak solution* $p : [0, T] \times I\!R \mapsto I\!R^n$. *The (possibly discontinuous) feedback controls* $U_j^*(t, x) \doteq u_j^*(p_j(t, x))$ *implicitly defined by* (4) *provide a Nash equilibrium solution to the differential game. The trajectories* $t \mapsto x(t)$ *depend Lipschitz continuously on the initial data* (τ, y).

It is interesting to observe that the entropy admissibility conditions play no role in our analysis. For example, a solution of the system of conservation laws consisting of a single, non-entropic shock still determines a Nash equilibrium solution, provided that the amplitude of the shock is small enough. There is, however, a way to distinguish entropy solutions from all others that is also in the context of differential games. Indeed, entropy solutions are precisely the ones obtained as vanishing viscosity limits [4]. They can thus be derived from a stochastic differential game of the form

$$dx = \sum_{i=1}^{n} f_i(x, u_i)\, dt + \varepsilon\, dw,$$

letting the white noise parameter $\varepsilon \to 0$. Here, dw formally denotes the differential of a Brownian motion. For a discussion of stochastic differential games we refer to [18].

3 Semi-Cooperative Differential Games

3.1 Lack of hyperbolicity in vector cases

Unfortunately, though strict hyperbolicity can usually be found in one-space dimension, it is not the case in higher-space dimensions. When the state of the system is described by a vector $x \in I\!R^m$, $m \geq 2$, the system of Hamilton-Jacobi equations (5) for the value functions is generically not hyperbolic. For the reader's convenience, we recall here some basic definitions.

The linear multidimensional system with constant coefficients

$$\frac{\partial}{\partial t} v + \sum_{\alpha=1}^{m} A_\alpha \frac{\partial}{\partial x_\alpha} v = 0 \tag{1}$$

is said to be *hyperbolic* if, for each vector $\xi = (\xi_1, \ldots, \xi_m) \in I\!R^m$, the matrix

$$A(\xi) \doteq \sum_\alpha \xi_\alpha A_\alpha \tag{2}$$

admits a basis of real eigenvectors [30]. We shall say that (1) is *weakly hyperbolic* if all the eigenvalues of $A(\xi)$ are real, for every $\xi \in I\!R^m$.

Next, given a point $(x,p) = (x, p_1, \ldots, p_n) \in I\!R^{(1+n)m}$, with

$$x \in I\!R^m, \qquad p_i = \nabla_x V_i = (p_{i1}, \ldots, p_{im}) \in I\!R^m\,,$$

consider the linearized system

$$\frac{\partial v_i}{\partial t} + \sum_{j,\alpha} \left[\frac{\partial H_i}{\partial p_{j\alpha}}(x, p_1, \ldots, p_n) \right] \cdot \frac{\partial v_j}{\partial x_\alpha} = 0 \qquad i = 1, \ldots, n\,, \tag{3}$$

where all derivatives are computed at the point (x,p). This is equivalent to (1), with

$$(A_\alpha)_{ij} = \frac{\partial H_i}{\partial p_{j\alpha}}(x, p_1, \ldots, p_n)\,. \tag{4}$$

We now say that the system in (5) is hyperbolic (weakly hyperbolic) on a domain $\Omega \subseteq I\!R^{(1+n)m}$ if, for every $(x,p) \in \Omega$, the linearized system (3) is hyperbolic (weakly hyperbolic, respectively).

To understand why the hyperbolicity condition fails, in a generic multi-dimensional situation, consider, for example, a two-player game on $I\!R^m$. In the scalar case, we have seen that the 2×2 system of Hamiton-Jacobian equations is not hyperbolic if the gradients of the value functions have opposite signs. In the multidimensional case, whenever $\nabla V_1, \nabla V_2 \in I\!R^m$ are not parallel to each other, we can find a vector ξ such that $\nabla V_1 \cdot \xi < 0$ and $\nabla V_2 \cdot \xi > 0$; see Fig. 2. In this case, the eigenvalues of the corresponding matrix $A(\xi)$ in (2) and (4) are complex, and the system is called elliptic.

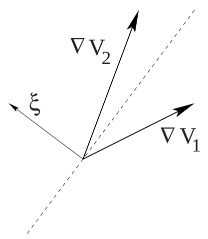

Figure 2: Hyperbolicity fails generically in multi-space dimensions.

3.2 Ill-posedness of the Cauchy problem

When the system is not hyperbolic, the Cauchy problem (5) is ill posed. See [16,23] for recent discussion on this subject. Here we show a brief analysis on how vanishing viscosity approximations can fail to converge to a well-defined solution. For more details, see [11].

Consider a two-persons non-cooperative differential game in one-space dimension, with the simple dynamics

$$\dot{x} = u_1 + u_2 \,, \qquad x(\tau) = y \,, \tag{5}$$

and payoff functionals

$$J_i = J_i(\tau, y, u_1, u_2) = g_i\big(x(T)\big) - \int_\tau^T \frac{u_i^2}{2}\, dt \qquad i = 1, 2.$$

Here, u_i is the control implemented by the i-th player, while g_i is his terminal payoff. Let V_1, V_2 be the corresponding value functions, and call $p_1 \doteq V_{1,x}$ and $p_2 \doteq V_{2,x}$ their spatial derivatives. The corresponding optimal feedback control u_i^* for the i-th player is

$$u_i^*(p_i) = \arg\max_\omega \big\{p_i \cdot \omega - (\omega^2/2)\big\} = p_i \,, \tag{6}$$

and the Hamiltonian functions are

$$H_i(p_1, p_2) = (p_1 + p_2)p_i - p_i^2/2, \qquad i = 1, 2 \,.$$

Therefore, $p = (p_1, p_2)$ satisfies a 2×2 system of conservation laws, solved backward in time

$$p_{i,t} + H_i(p_1, p_2)_x = 0, \qquad p_i(x, T) = g_{i,x}(x) \,.$$

Setting $\tau = T - t$, and still using t as time variable, we obtain a more standard Cauchy problem, to be solved forward in time:

$$\begin{cases} p_{1,t} - (p_1^2/2 + p_1 p_2)_x &= 0, \\ p_{2,t} - (p_2^2/2 + p_1 p_2)_x &= 0, \end{cases} \tag{7}$$

with the initial data

$$p_1(0,x) = g_{1,x}(x), \qquad p_2(0,x) = g_{2,x}(x). \tag{8}$$

The system (7) can be written in quasi-linear form:

$$p_t - A(p)p_x = 0, \qquad A(p) \doteq \begin{pmatrix} p_1 + p_2 & p_1 \\ p_2 & p_1 + p_2 \end{pmatrix}. \tag{9}$$

The eigenvalues of the matrix $A(p)$ are real if $p_1 p_2 \geq 0$, and complex if $p_1 p_2 < 0$.

Throughout the following, we focus our attention on solutions with $p_1 p_2 < 0$, so that hyperbolicity fails. As a first step, we add a small viscosity and consider the parabolic system

$$p_t^\varepsilon - A(p^\varepsilon)p_x^\varepsilon = \varepsilon p_{xx}^\varepsilon\,. \tag{10}$$

This system is related to a stochastic differential game with dynamics

$$dx = (u_1 + u_2)dt + \varepsilon\, d\omega\ ,$$

where ω denotes a standard Brownian motion, as in [18]. Observe that $p^\varepsilon = (p_1^\varepsilon, p_2^\varepsilon)$ provides a solution to (10) if and only if

$$p^\varepsilon(t,x) = p(t/\varepsilon,\ x/\varepsilon)\ ,$$

where $p = (p_1, p_2)$ solves the system with unit viscosity

$$\begin{cases} p_{1,t} - (p_1^2/2 + p_1 p_2)_x \;=\; p_{1,xx}\ , \\ p_{2,t} - (p_2^2/2 + p_1 p_2)_x \;=\; p_{2,xx}\,. \end{cases} \tag{11}$$

To achieve an understanding of solutions of (10), it thus suffices to study the system (11). An interesting class of solutions of (11) are the **traveling waves**, having the form $p(t,x) = P(x - \sigma t)$. The function $P : I\!R \mapsto I\!R^2$ must then satisfy the second-order O.D.E.

$$P'' = -[A(P) + \sigma I]P', \tag{12}$$

where $A = DH$ is the Jacobian matrix in (9) and I denotes the 2×2 identity matrix.

Integrating equation (12) once, we obtain:

$$P' = (H(\overline{P}) + \sigma\overline{P}) - \big(H(P) + \sigma P\big),$$

where $\overline{P} = (\bar p_1, \bar p_2)$ is some constant vector. We are particularly interested in periodic solutions of the O.D.E.:

$$\begin{cases} p_1' \;= (\bar p_1 \bar p_2 + \bar p_1^2/2) - (p_1 p_2 + p_1^2/2) - \sigma(p_1 - \bar p_1)\,, \\ p_2' \;= (\bar p_1 \bar p_2 + \bar p_2^2/2) - (p_1 p_2 + p_2^2/2) - \sigma(p_2 - \bar p_2)\,, \end{cases} \tag{13}$$

taking values inside the elliptic region where $p_1 p_2 < 0$. Linearizing (13) at the equilibrium point $(\bar p_1, \bar p_2)$, one gets:

$$\begin{cases} z_1' &= -(\bar p_1 + \bar p_2 + \sigma) z_1 - \bar p_1 z_2 \,, \\ z_2' &= -\bar p_2 z_1 - (\bar p_1 + \bar p_2 + \sigma) z_2 \,. \end{cases} \tag{14}$$

Notice that, if one chooses $\sigma = \bar\sigma \doteq -\bar p_1 - \bar p_2$, then the two eigenvalues

$$\lambda_1,\ \lambda_2 = -(\bar p_1 + \bar p_2 + \sigma) \pm i\sqrt{-\bar p_1 \bar p_2}$$

are purely imaginary. By the Hopf bifurcation theorem [28], for every $\delta > 0$ sufficiently small there exists a value $\sigma = \sigma(\delta) \approx \bar\sigma$ such that the corresponding system (14) has a periodic orbit passing through the point $(\bar p_1 + \delta\,,\ \bar p_2)$.

In this way, we obtain a family of periodic orbits for the system (13), depending on the parameters $\bar p_1, \bar p_2$, and δ. If $s \mapsto \big(p_1(s),\ p_2(s)\big)$ is any such orbit, then

$$\big(p_1(t,x), p_2(t,x)\big) \doteq \big(p_1(x-\sigma t),\ p_2(x-\sigma t)\big) \tag{15}$$

yields a solution of the parabolic system (11) in the form of a periodic traveling wave. In turn, the functions

$$\Big(p_1^\varepsilon(t,x),\ p_2^\varepsilon(t,x)\Big) = \left(p_1\Big(\frac{x-\sigma t}{\varepsilon}\Big),\ p_2\Big(\frac{x-\sigma t}{\varepsilon}\Big)\right) \tag{16}$$

provide a solution to the system (10) with small viscosity.

We now recall that, by (6), the corresponding dynamic of the system is:

$$\dot x(t) = u_1^* + u_2^* = p_1\Big(\frac{x-\sigma t}{\varepsilon}\Big) + p_2\Big(\frac{x-\sigma t}{\varepsilon}\Big).$$

In our construction,

$$p_1 + p_2 \approx \bar p_1 + \bar p_2 \ \neq\ \sigma \approx -\bar p_1 - \bar p_2\ .$$

As the viscosity parameter $\varepsilon \to 0+$, along each trajectory the controls $(u_1^*,\ u_2^*) = (p_1^\varepsilon,\ p_2^\varepsilon)$ are periodic functions of time with fixed amplitude and with period approaching zero. Because of this oscillatory behavior, there is no strong limit in $\mathbf{L}^1$. Yet, a weak limit exists and can be represented in terms of Young measures [30]. These oscillatory limits can now be interpreted as **chattering feedback controls**. The limit trajectories cover the whole t-x plane. They all have the same constant speed, determined by the weak limit of $p_1^\varepsilon + p_2^\varepsilon$.

A further analysis in [11] shows that these viscous traveling waves have almost the same instability properties as the constant states.

3.3 Transition from Nash equilibrium to a Pareto optimum

The eventual conclusion of our analysis is that, except for the one-dimensional case, the concept of Nash non-cooperative equilibrium is not appropriate to study games with complete information, in continuous time. The highly unstable nature of the solutions makes it impossible to extract useful information from the mathematical model.

In the literature, various approaches have been proposed, to overcome this basic difficulty. Following [3], one can study a special class of multi-dimensional games, with linear dynamics and quadratic cost functionals. In this case, the system of Hamilton-Jacobi equations (5) may be ill posed, but one can always find a unique solution within the set of quadratic polynomial functions. An alternative strategy is to add some random noise to the system. This leads to the analysis of a stochastic differential game [18,25], with dynamics described by

$$dx = f(x, u_1, \ldots, u_n)\, dt + \varepsilon\, dw \ ,$$

where dw is the differential of a Brownian motion. The corresponding system describing the value functions is now parabolic, and admits a unique smooth solution. However, one should be aware that, as $\varepsilon \to 0$, the solutions become more and more unstable and may not approach a well-defined limit.

An entirely different approach was proposed in [11], where the authors explored the possibility of partial cooperation among players. To explain the heart of the matter, we first observe that the Hamiltonian functions at (5)–(4) are derived from the following instantaneous optimization problem. Given $p_1, \ldots, p_n \in I\!R^m$, the i-th player seeks a control value u_i which maximizes his instantaneous payoff:

$$Y_i = p_i \cdot \Big(f_0 + \sum_j u_j \Big) - h_i(u_i)\,. \tag{17}$$

In the case of two players, the set of possible payoffs (Y_1, Y_2) attainable as $(u_1, u_2) \in I\!R^{2m}$ corresponds to the shaded region in Fig. 3. The Nash equilibrium strategy produces the payoffs at **N**, and corresponds to the Hamiltonian in (5)–(4).

In this context, it is interesting to examine alternative strategies for the two players, resulting in different Hamiltonian functions. If full cooperation were possible, then the players would simply choose the strategy that maximizes the sum $Y_1 + Y_2$ of the two payoffs, i.e., the point **C** in Fig. 3. In this case, u_1, u_2 can be regarded as components of one single control function. The optimization problem thus fits entirely within the framework of optimal control theory. The only additional issue arising from the differential game is the possible side payment that one of the players should make to the other, to preserve fairness.

Alternatively, the players may choose strategies u_1, u_2 corresponding to a Pareto optimum $\mathbf{P}$. See [1] for basic definitions. In the case where the two players cannot communicate and are not allowed to make side payments, their behavior can still drift away from a Nash equilibrium and approach a Pareto optimum

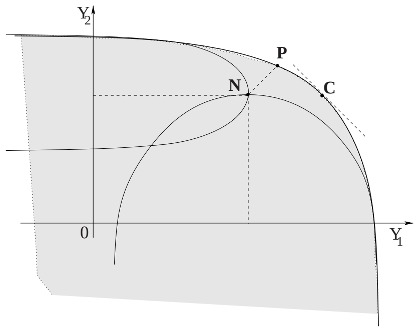

Figure 3: From Nash equibrium to Pareto optimal.

which improves both of their payoffs. For a game modeling an iterated prisoner's dilemma, in [31] Smale introduced a class of "good" strategies, which induce the other player to cooperate. Asymptotically for large times, the outcome of the game thus drifts away from the Nash equilibrium, approaching a Pareto optimum. It is remarkable that these strategies do not require any direct communication among the players. These same ideas are appropriate also in continuous time.

Real life situations where such transition happens can often be observed. For example, you were invited to a party. You went there, and it turned out to be very boring. You wished you could leave but it would look bad if you were the first to leave, so you stayed. Many of the other guests were having the same thoughts as you, and they stayed also. This is a Nash Equilibrium of this non-cooperative game. You were not discussing with other guests about leaving, but everybody had the same intention. Then, if someone eventually got up and approached the host, many other guests would join him, so the first person would not look bad to be the first one to leave. So, suddenly everyone was leaving. This is a Pareto optimum. The key in the transition is the willingness of each player to cooperative, even though without actual cooperation.

A full description of such transition in differential games can be found in [11]. To illustrate the main ideas, let (U_1^N, U_2^N) and (Y_1^N, Y_2^N) be the Nash strategies and the gains for two players, respectively, and let (U_1^P, U_2^P) and (Y_1^P, Y_2^P) be the strategies and gains with some Pareto optimum that both players think is fair, respectively. A strategy is called a **good strategy** for player 1 if the following three conditions hold.

(C1) If the gain of the first player is smaller than what he gets by playing the Nash strategy, then he leans back toward U_1^N.

(C2) If the second player is gaining more than his due profit Y_2^P, then the first player should again lean back toward U_2^N.

(C3) If the second player is cooperating, then the first player should approach the Pareto strategy.

Notice that the first two conditions say that player 1 should play "tough" whenever the other player is taking advantage of him. The last condition implies that he should play “soft” when the game goes in his favor or when the other player is cooperating.

For a given Pareto optimum (U_1^P, U_2^P), the definition of a good strategy for player 2 is completely analogous. One can now envision a situation where each player adopts a partially cooperative strategy, based on the behavior of the other player. In [11], we showed that if both players adopt good strategies, then the outcome of the game will approach the Pareto optimum.

3.4 A smart anti-cheating strategy

Assume that the first player adopts a good strategy, expecting the payoffs (Y_1^P, Y_2^P). It is possible for the second player to “out-smart” him and gain more than his fair share Y_2^P. Indeed, assume $p_1, p_2 > 0$, and two players are semi-cooperating, so the system is moving toward the Pareto optimum P. See Fig. 4.

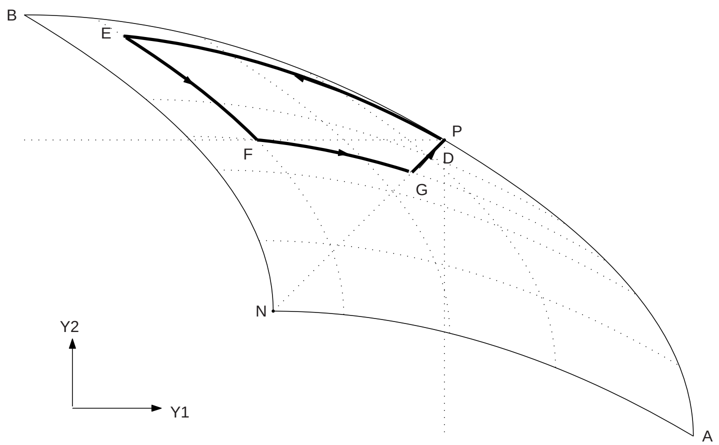

Figure 4: A possible “cheating cycle” by player 2.

At a point D, close to P, player 2 can quickly change his control, moving close to the Nash equilibrium. The payoffs (Y_1, Y_2) will move from Y^D to Y^E, favoring the second player. Of course player 1 doesn’t like it, and will move towards his Nash equilibrium. As a consequence, the payoff of the second player will now decrease. When the payoffs reach a point Y^E such that $Y_2^F = Y_2^P$, player 2 decides to cooperate again, setting his control back to U_2^P, reaching the pointing G. Now the new payoffs (Y_1^G, Y_2^G) are in favor of the first player, so he decides to cooperate, and the system is moving toward the Pareto optimum. The whole cycle can then be repeated.

Notice now, in this cycle, player 2 can make the transition from D to E, and from F to G very quickly. Along the arc from E to F, player 2 gains more than his fair share Y_2^P, while along the arc from G to D, player 2 gains less than Y_2^P. If the time spent along the first arc is longer compared with the time spend on the latter, his strategy will be profitable in a long run.

Based on this analysis, player 1 could design his counter-strategy, making it a smart anti-cheating one. In order to discourage the above behavior, player 1 should quickly go back to his Nash strategy when the other player is not cooperating, and approach the Pareto optimum slowly in the cooperative case. In other words, if player 2 tries to cheat, player 1 should not be too quick in restoring cooperation. By doing so, the gain of player 2 in the long run will not be more than his fair share Y_2^P. Again, see [11] for details.

3.5 Weak hyperbolicity of the semi-cooperative games

If, for any given p_1, p_2, the players choose control values $u_i(p_1, p_2)$ which yield some Pareto optimum, we say that their strategy is *semi-cooperative*. When these strategies are implemented, the value functions will satisfy a different system of Hamilton-Jacobi equations.

Consider a two-person game with the dynamic

$$\dot{x} = u_1 + u_2 \ .$$

The instantaneous gain functionals Y_1, Y_2 become:

$$\begin{aligned} Y_1(p_1, p_2, u_1, u_2) \ &= \ p_1 \cdot (u_1 + u_2) - h_1(u_1), \\ Y_2(p_1, p_2, u_1, u_2) \ &= \ p_2 \cdot (u_1 + u_2) - h_2(u_2). \end{aligned}$$

Let $U_i^P(p_1, p_2, s)$ denote the Pareto optimum that maximizes the combined payoff $sY_1 + Y_2$ for some $s > 0$, and let's write

$$Y_i^P(p_1, p_2, s) \doteq Y_i\left(p_1, p_2, U_1^P(p_1, p_2, s), U_2^P(p_1, p_2, s)\right) .$$

Assume that the players adopt feedback strategies of the form

$$u_1 = u_1^*(p_1, p_2) \,, \qquad\qquad u_2 = u_1^*(p_1, p_2) \ ,$$

then the value functions V_1, V_2 will satisfy the system of Hamilton-Jacobi equations

$$\begin{cases} V_{1,t} + H_1(\nabla_x V_1, \ \nabla_x V_2) \ = 0, \\ V_{2,t} + H_2(\nabla_x V_1, \ \nabla_x V_2) \ = 0, \end{cases}$$

with

$$\begin{aligned} H_1(p_1, p_2) \ &= Y_1(p_1, p_2, u_1^*(p_1, p_2), u_2^*(p_1, p_2)) \,, \\ H_2(p_1, p_2) \ &= Y_2(p_1, p_2, u_1^*(p_1, p_2), u_2^*(p_1, p_2)) \,. \end{aligned}$$

A remarkable fact, proved in [11], is that the above system is always weakly hyperbolic for a very general class of strategies $u_i^*(p_1, p_2)$, under the only assumption that they achieve Pareto optima (Y_1^P, Y_2^P). In particular, this includes the semi-cooperative strategy $(u_1^*, u_2^*) = (U_1^P, U_2^P)$ considered in this paper. This means that the Jacobian matrix of the Hamiltonian functions has real (possibly coinciding) eigenvalues. The result is true for both one-dimensional and multi-dimensional cases. This looks promising in connection with the Cauchy problem (5) for the value functions. Indeed, our semi-cooperative solutions will not experience the severe instabilities of the Nash solutions. It is thus expected that some existence theorem should hold in greater generality.

References

[1] J.P. Aubin. *Mathematical Methods of Game and Economic Theory*, North Holland, Amsterdam 1979.

[2] M. Bardi and I. Capuzzo Dolcetta. *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Birkhäuser, Boston 1997.

[3] T. Basar and G.J. Olsder. *Dynamic Non-cooperative Game Theory*, 2nd edition, Academic Press, London 1995.

[4] S. Bianchini and A. Bressan. Vanishing viscosity solutions of nonlinear hyperbolic systems, *Annals of Mathematics*, to appear.

[5] V.G. Boltianskii. Sufficient conditions for optimality and the justification of the dynamic programming principle, *SIAM J. Control Optim.* **4** (1966), 326-361.

[6] A. Bressan. Unique solutions for a class of discontinuous differential equations, *Proc. Amer. Math. Soc.* **104** (1988), 772-778.

[7] A. Bressan. *Hyperbolic systems of conservation laws. The one-dimensional Cauchy problem*, Oxford University Press, Oxford 2000.

[8] A. Bressan, T.P. Liu and T. Yang. L^1 stability estimates for $n \times n$ conservation laws, *Arch. Rational Mech. Anal.* **149** (1999), 1–22.

[9] A. Bressan and F. Priuli. Infinite horizon noncooperative differential games, *J. Diff. Eq.* **227** (2006), no. 1, 230–257.

[10] A. Bressan and W. Shen. Small BV solution of hyperbolic noncooperative differential games, *SIAM J. Control Optim.* **43** (2004) 104–215.

[11] A. Bressan and W. Shen. Semi-cooperative strategies for differential games, *Int. J. Game Th.* **32** (2004), 561–593.

[12] P. Brunovsky. Existence of regular syntheses for general problems, *J. Diff. Eq.* **38** (1980), 317–343.

[13] P. Cardaliaguet and S. Plaskacz. Existence and uniqueness of a Nash equilibrium feedback for a simple non-zero-sum differential game, *Int. J. Game Th.* **32** (2003), 33–71.

[14] G.Q. Chen and A. Rustichini. The Riemann solution to a system of conservation laws, with application to a nonzero sum game, *Contemp. Math.* **100** (1988), 287–297,

[15] F.H. Clarke. *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

[16] H. Frid and I.S. Liu. Oscillation waves in Riemann problems inside elliptic regions for conservation laws of mixed type, *Z. Angew. Math. Phys.* **46** (1995), 913–931.

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

[18] A. Friedman. Stochastic Differential Games, *J. Diff. Eq.* **11** (1972), 79–108.

[19] J. Glimm. Solutions in the large for nonlinear hyperbolic systems of equations, *Comm. Pure Appl. Math.* **18** (1965), 697–715.

[20] O. Hajek. Discontinuous differential equations I, *J. Diff. Eq.* **32** (1979), 149–170.

[21] T. Iguchi and P.G. LeFloch. Existence theory for hyperbolic systems of conservation laws with general flux-functions, *Arch. Rational Mech. Anal.* **168** (2003), 165–244.

[22] R. Isaacs. *Differential Games*, Wiley, New York 1965.

[23] H.O. Kreiss and J. Yström. Parabolic problems which are ill-posed in the zero dissipation limit, *Math. Comput. Model* **35** (2002), 1271–1295.

[24] T.P. Liu. Admissible solutions of hyperbolic conservation laws, *Amer. Math. Soc. Memoir* **240** (1981).

[25] P. Mannucci. Nonzero sum stochastic differential games with discontinuous feedback, *SIAM J. Control Optim.* **43** (2004), 1222-1233.

[26] J. Nash. Non-cooperative games, *Ann. Math.* **2** (1951), 286-295.

[27] G.J. Olsder. On open- and closed-loop bang-bang control in nonzero-sum differential games, *SIAM J. Control Optim.* **40** (2001), 1087–1106.

[28] L. Perko. *Differential Equations and Dynamical Systems*, Springer-Verlag, 1991.

[29] B. Piccoli and H.J. Sussmann. Regular synthesis and sufficiency conditions for optimality, *SIAM J. Optim. Control* **39** (2000), 359-410.

[30] D. Serre. *Systems of Conservation Laws I, II.* Cambridge University Press, Cambridge 2000.

[31] S. Smale. The prisoner's dilemma and dynamical systems associated to non-cooperative games, in *The Collected Papers of Stephen Smale*, F. Cucker and R. Wong Eds., World Scientific, 2000.

[32] E.M. Vaisbord and V.I. Zhukovskii. *Introduction to Multi-players Differential Games and their Applications,* Gordon and Breach Science Publishers, 1988.

Some Geometrical Properties of the Phase Space Structure in Games on Manifolds

Andrei Akhmetzhanov
Institute for Problems in Mechanics
Russian Academy of Sciences
Moscow, Russia
akhmetzhanov@gmail.com

Abstract

Pursuit-evasion games with simple motion on bounded surfaces in 3D space, like ellipsoids of revolution or two-sided plane ellipses, are considered. In each case, we state the existence conditions for singular trajectories and specify the domain of problem parameters where the optimal phase-portrait does not contain singular trajectories and is similar to the case of Euclidean plane.

Key words. Pursuit-evasion games, calculus of variations.

AMS Subject Classifications. Primary 49N75; Secondary 49Q20.

1 Introduction

In this paper we consider differential games on bounded 2D manifolds in 3D Euclidean space like ellipsoids of revolution or two-sided plane ellipses. Dynamics is simple. This means that the players control their velocity by choosing its value and direction. The constraints on the control parameters are symmetric and spherical. The game payoff function is the time required to the pursuer to capture the evader. Thus, we study here pursuit-evasion games with simple dynamics and zero capture radius. Moreover, both players have complete information on the position of each other at each moment of time.

Differential games on various surfaces (manifolds) have been the subject of investigation in a number of publications. In [1], a pursuit-evasion game on the plane has been formulated in the presence of obstacle (two ships in a sea with a circular island). Its complete solution has been obtained later in [2,3]. The differential games on 2D manifolds in the Euclidean space have been investigated lately: in [4]-[8] (on cone, hyperboloid of revolution), in [9] (on multidimensional spheres), in [10] (on two-sided plane figures). Some review of such games one can find in [11].

Note that the motion of players along the same shortest geodesic line is an optimal one in the greater part of the game space (called primary domain). Generally, a secondary non-empty domain can also exist. Starting at the points of the secondary

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_6,

domain the pursuer can guarantee a better result comparing with the pursuit along the geodesic lines. The secondary domain arises due to the existence of two or more different geodesic lines of the equal length for some positions of the players on the surface. For the case of the bounded surfaces, the secondary domain can be empty for some parameters of the game and thus the primary strategies of players will be optimal in the whole space.

Differential games of simple pursuit on 2D manifolds are of a certain interest because in some cases they can be a good approximation to more complicated models, or can serve as test examples for numerical computations.

The game on two-sided plane ellipse is a limit case for the game on a triaxial ellipsoid, when one of the axes is arbitrary small. Thus, its consideration is important for understanding of the game properties in general case. In its turn, a game on a triaxial ellipsoid (or on a bounded convex surface) can be regarded as a subgame for a pursuit-evasion game in 3D space with obstacle (state constraint).

2 Problem formulation

Let the motion of players P (pursuer) and E (evader) occurs along the bounded surface $\mathcal{M}$ in $\mathbb{R}^3$. Dynamics is described by the following equations:

$$\dot{x} = u, \quad \dot{y} = v, \quad x(0) = x^0, \quad y(0) = y^0, \quad x, y, u, v \in \mathbb{R}^2 \tag{1}$$

$$\sqrt{\langle G(x)u, u\rangle} \le 1, \quad \sqrt{\langle G(y)v, v\rangle} \le \nu < 1 \ ,$$

where $\langle\cdot,\cdot\rangle$ is the scalar product in Euclidean space, $x = (x_1, x_2)$, and $y = (y_1, y_2)$ are the local coordinates of the players P and E on the surface $\mathcal{M}$. The 2x2 matrix $G(\xi)$ is symmetric, positive definite, $\det G(\xi) \neq 0$. It is called the metric tensor of the manifold $\mathcal{M}$ and locally described by the coordinates $\xi \in \mathbb{R}^2$. The matrix G is said to determine the Riemannian metric on the manifold $\mathcal{M}$ with the element of the arc length ds:

$$ds^2 = \langle G(\xi)d\xi, d\xi\rangle, \quad ds = \sqrt{\langle G(\xi)\dot{\xi}, \dot{\xi}\rangle}dt \ .$$

For a differential game on a Euclidean plane, the metric tensor is an identity matrix and the expression for the control constraints has the most simple form

$$u_1^2 + u_2^2 \le 1, \quad v_1^2 + v_2^2 \le \nu^2 \ .$$

The value of ν can be considered here as the ratio of the velocities of the pursuer and the evader.

The end of the game will take place when we have the coordinate coincidence of the players

$$t = T: \quad x(T) = y(T) \ , \tag{2}$$

where the optimal time T is regarded as the game value.

3 Game space

Consider the above-formulated differential game on a Euclidian plane, where coordinates of the players P and E are defined through a couple of Cartesian coordinates. In this case, the optimal behavior of the players is the motion along the line PE and the game value can be written in the form

$$V_1(x,y) = \frac{L(x,y)}{1-\nu} \;, \tag{3}$$

where $L(x,y) = L_{PE}$ is the length of the segment PE. Turning back to our initial game problem on 2D surface $\mathcal{M}$ in $\mathbb{R}^3$, let's do the following partition of the phase space $D = \mathcal{M} \times \mathcal{M}$. Let's say that the point of the phase space belongs to the *primary domain* D_1 if and only if the game value has the same structure as in the game on a Euclidian plane (3). Now the function $L(x,y)$ is the length of the shortest geodesic line connecting players. The value of L can be defined from the corresponding problem of the calculus of variations:

$$L(x,y) = \min_{\xi(\cdot)} \int_{t_0}^{t_1} \sqrt{\langle G(\xi)\dot{\xi}, \dot{\xi}\rangle}\, dt, \quad \xi(t_0) = x, \quad \xi(t_1) = y \;. \tag{4}$$

As mentioned above, the optimal behavior of players in D_1 is the motion along the geodesic line. Therefore, the value of the controls parameters of the players should be taken in the (feedback) form; see [12, p. 139]:

$$u(x,y) = -a(x,y) = -G^{-1}(x)\, \partial L/\partial x$$
$$v(x,y) = \nu b(x,y) = \nu G^{-1}(y)\, \partial L/\partial y \;.$$

Here, a and b are the unit tangent vectors to the geodesic line at the end points P and E correspondingly; see Fig. 1:

$$|a|_x^2 = \langle G(x)a, a\rangle = 1, \quad |b|_y^2 = \langle G(y)b, b\rangle = 1 \;.$$

Note that the game value at any initial position can not be greater than the time of motion along the shortest geodesic line connecting the initial positions of the players [11].

The subdomain $D_2 \subset \mathcal{M} \times \mathcal{M}$ is called the *secondary domain* if the following condition is fulfilled for any position of the players from D_2

$$V(x,y) < V_1(x,y) \;.$$

We can see that the whole phase space is a sum of the primary and secondary domain $D = D_1 + D_2$.

Note the following important property of the games on manifolds. There are some positions of the players on the manifold such that two or more different

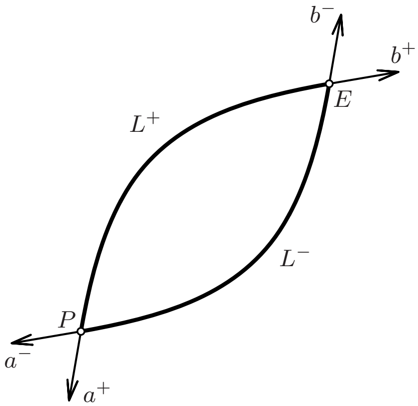

Figure 1: Geometry of geodesic lines on manifolds.

geodesic lines (extremals of the variational problem mentioned above) can connect them. In this case, the shortest geodesic line is the extremal with minimal length

$$L(x,y) = \min_{\alpha \in A} L_\alpha(x,y) = L_{\widetilde{\alpha}}(x,y), \quad \widetilde{\alpha} \in A^* \subset A \,,$$

where A is a set of all extremals of (4), A^* is a set of the minimal extremals with equal length. In general, the set A^* does not consist only of one element. For the wide variety of manifolds we can restrict ourself to the consideration of the set A consisted of two elements $A = \{+,-\}$. In this case,

$$L(x,y) = \min[L^+(x,y), L^-(x,y)] \,. \tag{5}$$

Let's denote Γ the domain consisted of all positions of the players with two different geodesic lines with equal length

$$(x,y) \in \Gamma : \quad L^+(x,y) = L^-(x,y) \,. \tag{6}$$

As was shown earlier [11], the primary strategies are still optimal if the following condition holds:

$$|a^+ + a^-| - \nu|b^+ + b^-| < 2(1-\nu) \,. \tag{7}$$

Therefore, the edge of the secondary domain on the intersection with the domain Γ can be defined through the conditions:

$$(x,y) \in B: \quad |a^+ + a^-| - \nu|b^+ + b^-| = 2(1-\nu), \quad L^+(x,y) = L^-(x,y) \,. \tag{8}$$

The aim of the following investigation is the consideration of differential games on ellipsoids of revolution and two-sided plane ellipses. We will try to find the boundary of the domain Γ, build it in some special coordinates, and construct the edge B of the secondary domain when it is not empty. One of the most interesting questions is to study the value of parameters when the secondary domain arises. In this case, singular characteristics are appeared in the phase space and the study of differential games requires more detailed consideration.

4 Differential games on 2D bounded surfaces

4.1 Game on the ellipsoid of revolution

Construction of the geodesic lines. Consider the ellipsoid of revolution about the major axis (oblate spheroid). Without loss of generality, take the semiminor axis equal to the one and eccentricity ε as a parameter defined the surface $\mathcal{M}$. Then the semimajor axis of ellipsoid equals to $a = 1/\sqrt{1-\varepsilon^2}$.

Introduce the so-called elliptic coordinates of a point on the manifold $\mathcal{M}$ as:

$$x_* = \frac{\cos\varphi\cos\theta}{\sqrt{1-\varepsilon^2}}, \quad y_* = \sin\varphi\cos\theta, \quad z_* = \sin\theta\ .$$

The problem of the construction of geodesics leads to the following problem in the calculus of variations. It has been proved [13] that the extremals of the *curve length functional* (4) are contained in the set of extremals of the *curve action functional* (square root is omitted):

$$E(x,y) = \min_{\xi(\cdot)} \int_{t_0}^{t_1} \langle G(\xi)\dot{\xi}, \dot{\xi}\rangle\, dt, \quad \xi(t_0) = x, \quad \xi(t_1) = y\ , \tag{9}$$

where t is the natural parameter of the curve. We should put the auxiliary condition that the velocity of the motion on a geodesic line equals one.

In elliptic coordinates, the Lagrangian of the functional (9) has the form

$$\mathcal{L}(\varphi,\theta,\dot{\varphi},\dot{\theta}) = \left(\cos^2\theta + \frac{\varepsilon^2\sin^2\varphi\cos^2\theta}{1-\varepsilon^2}\right)\dot{\varphi}^2 + \frac{\varepsilon^2\sin 2\varphi\sin 2\theta}{2(1-\varepsilon^2)}\,\dot{\varphi}\,\dot{\theta} + \left(1 + \frac{\varepsilon^2\cos^2\varphi\sin^2\theta}{1-\varepsilon^2}\right)\dot{\theta}^2\ ,$$

as earlier $\dot{f}$ means the differentiation of the function f with respect to t.

Euler-Lagrange equations are written as follows:

$$\frac{\partial\mathcal{L}}{\partial\varphi} - \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\varphi}} = 0: \quad \frac{\varepsilon^2}{2}\sin 2\varphi\sin 2\theta\,\ddot{\theta} + 2(1-\varepsilon^2\cos^2\varphi)\cos^2\theta\,\ddot{\varphi} + \varepsilon^2\sin 2\varphi\cos^2\theta(\dot{\varphi}^2+\dot{\theta}^2) - 2(1-\varepsilon^2\cos^2\varphi)\sin 2\theta\,\dot{\varphi}\,\dot{\theta} = 0$$

$$\frac{\partial\mathcal{L}}{\partial\theta} - \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\theta}} = 0: \quad 2(1-\varepsilon^2+\varepsilon^2\cos^2\theta\sin^2\theta)\ddot{\theta} + \frac{\varepsilon^2}{2}\sin 2\varphi\sin 2\theta\,\ddot{\varphi} + (1-\varepsilon^2\sin^2\varphi)\sin 2\theta\,\dot{\varphi}^2 + \varepsilon^2\cos^2\varphi\sin 2\theta\,\dot{\theta}^2 - 2\varepsilon^2\sin 2\varphi\sin^2\theta\,\dot{\varphi}\,\dot{\theta} = 0\ .$$

The last equation can be replaced by the condition that the velocity along the curve has to be equal one, $\mathcal{L}(\varphi,\theta,\dot{\varphi},\dot{\theta}) = 1$. Such a system of differential equations can be solved either numerically or analytically using Jacobi special functions.

Construction of the boundary $\partial\Gamma$. The symmetry of the surface $\mathcal{M}$ leads to the following statement: there are two different geodesic lines with equal length connecting players only if this pair of points is on the same ellipse, generating original ellipsoid (generatrix). It is a necessary condition. To obtain the sufficient condition on this statement we have also to require that the distance between points P and E should exceed the distance from one of the points to the corresponding focus, defined using Jacobi equation; see [14]. To find the position of the focus we should fix one of the points, e.g., P, and solve the following problem of the calculus of variations—initial value problem for the Jacobi equation.

Consider the curve functional like (4): $L(\gamma) = \int F(\varphi, \theta, \theta')\, d\varphi$, where the variable φ is playing a role of the independent variable. The Lagrangian is written as follows:

$$F(\varphi,\theta,\theta') = \left[\cos^2\theta + \frac{\varepsilon^2 \sin^2\varphi \cos^2\theta}{1-\varepsilon^2} + \frac{\varepsilon^2 \sin 2\varphi \sin 2\theta}{2(1-\varepsilon^2)}\,\theta' \right.$$
$$\left. + \left(1 + \frac{\varepsilon^2 \cos^2\varphi \sin^2\theta}{1-\varepsilon^2}\right)\theta'^2\right]^{1/2} .$$

In this expression, the derivative f' means the differentiation of the function f with respect to the parameter φ. Then the Jacobi equation takes the form

$$(Ph')' - (R - Q')h = 0, \quad h(\varphi_0) = 0, \quad h'(\varphi_0) = 1\,,$$

where

$$P(\varphi) = F_{\theta'\theta'}\Big|_{\substack{\theta'=0\\ \theta=0}} = \sqrt{\frac{1-\varepsilon^2}{1-\varepsilon^2\cos^2\varphi}}, \quad R(\varphi) = F_{\theta\theta}\Big|_{\substack{\theta'=0\\ \theta=0}} = \sqrt{\frac{1-\varepsilon^2\cos^2\varphi}{1-\varepsilon^2}}$$

$$Q(\varphi) = F_{\theta'\theta}\Big|_{\substack{\theta'=0\\ \theta=0}} = \frac{\varepsilon^2 \sin\varphi\cos\varphi}{\sqrt{(1-\varepsilon^2)(1-\varepsilon^2\cos^2\varphi)}}\,.$$

Substituting these expressions into the equation leads to the following ordinary differential equation:

$$h''(1-\varepsilon^2\cos^2\varphi) - h'\,\varepsilon^2 \sin\varphi\cos\varphi + h = 0, \quad h(\varphi_0) = 0, \quad h'(\varphi_0) = 1\,. \tag{10}$$

The next zero of its solution determines the position of the focus.

For the case of sphere ($\varepsilon = 0$), we have simple equation $h'' + h = 0$ with initial conditions $h(\varphi_0) = 0$, $h'(\varphi_0) = 1$. Its solution $h(\varphi)$ goes to zero only in the diametrically opposite point $\varphi = \varphi_0 + \pi$. The situation for an arbitrary case ($0 < \varepsilon < 1$) is shown in Fig. 2. We can see that two foci of the Jacobi equation P_1 and P_2 exist. Therefore, the position of the players is in the domain Γ when the point E is located between the points P_1 and P_2. The boundary situation takes place when the player E is exactly at one of these foci. Note that the domain Γ does not depend on the parameter ν.

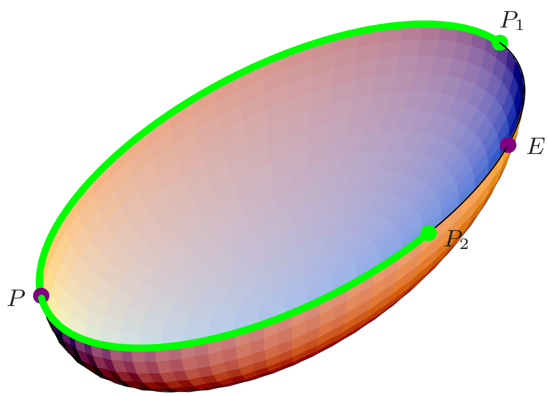

Figure 2: Position of the foci of Jacobi equation.

Let's present results of the construction of the boundary $\partial\Gamma$ in polar coordinates (r, φ). The first coordinate r is the length of the generatrix's segment connecting P and E, the second coordinate φ is an angle between the radius vector of P and semimajor axis of the ellipsoid. For $\varepsilon = 0$, the interior of the domain Γ is empty and the plot of $\partial\Gamma$ consists of two coincident circles. However, when the value of ε increases the boundary splits and the domain Γ becomes non-empty; see Fig. 3.

Construction of the edge B**.** For the beginning let us write the expression for the tangent vector to the geodesic line on ellipsoid. According to the definition it is:

$$\dot{\gamma}(t) = \begin{bmatrix} -\dfrac{1}{\sqrt{1-\varepsilon^2}}(\sin\varphi\cos\theta\,\dot{\varphi} + \cos\varphi\sin\theta\,\dot{\theta}) \\ \cos\varphi\cos\theta\,\dot{\varphi} - \sin\varphi\sin\theta\,\dot{\theta} \\ \cos\theta\,\dot{\theta} \end{bmatrix} .$$

For the case $\theta = 0$ corresponding to the position of players on the generatrix, this expression is simplified and takes the form:

$$\dot{\gamma}(t) = \begin{bmatrix} -\dfrac{\sin\varphi}{\sqrt{1-\varepsilon^2}}\,\dot{\varphi} \\ \cos\varphi\,\dot{\varphi} \\ \dot{\theta} \end{bmatrix} .$$

The position of the players in Γ belongs to the edge B of the secondary domain if the following equation takes place:

$$|a^+ + a^-| - \nu|b^+ + b^-| = 2(1-\nu) .$$

Since there is a mirror symmetry with respect to the plane $\theta = 0$, one has for the

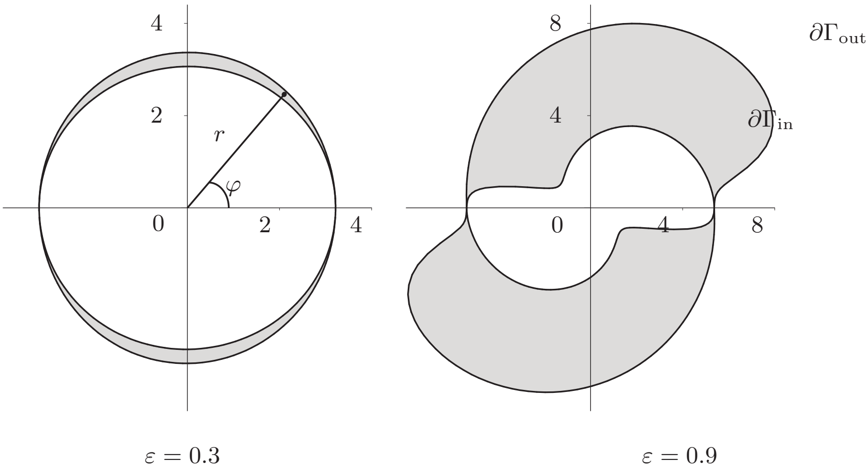

Figure 3: Domain Γ.

point E with a coordinate $\varphi = \varphi_1$ the following equality:

$$b^+ + b^- = 2\begin{bmatrix} -\dfrac{\sin\varphi_1}{\sqrt{1-\varepsilon^2}}\,\dot\varphi_1 \\ \cos\varphi_1\,\dot\varphi_1 \\ 0 \end{bmatrix},$$

and thus

$$|b^+ + b^-| = 2\sqrt{1+\frac{\varepsilon^2\sin^2\varphi_1}{1-\varepsilon^2}}\,|\dot\varphi_1|\,.$$

Analogous expression can be written for another point P with coordinates $\varphi = \varphi_0$, $\theta = 0$:

$$|a^+ + a^-| = 2\sqrt{1+\frac{\varepsilon^2\sin^2\varphi_0}{1-\varepsilon^2}}\,|\dot\varphi_0|\,.$$

Consequently, one gets the following equation for the edge B:

$$\sqrt{1+\frac{\varepsilon^2\sin^2\varphi_0}{1-\varepsilon^2}}\,|\dot\varphi_0| - \nu\sqrt{1+\frac{\varepsilon^2\sin^2\varphi_1}{1-\varepsilon^2}}\,|\dot\varphi_1| = 1-\nu\,. \tag{11}$$

The result of the numerical construction of the edge B is presented in Fig. 4. The edge consists of four arcs, the domain Γ_1 is a part of the primary domain such that $\Gamma/\Gamma_1 \subset D_2$. We see that the secondary domain arises near the boundary $\partial\Gamma$. This gives an idea for the following constructions. Let's build the curve in the space of parameters ε and ν which divides it into two parts. In one of them, A_1 the secondary domain is empty and the primary strategies are optimal in the whole space. In the other one, A_2 is not true. Such a curve is shown in Fig. 5.

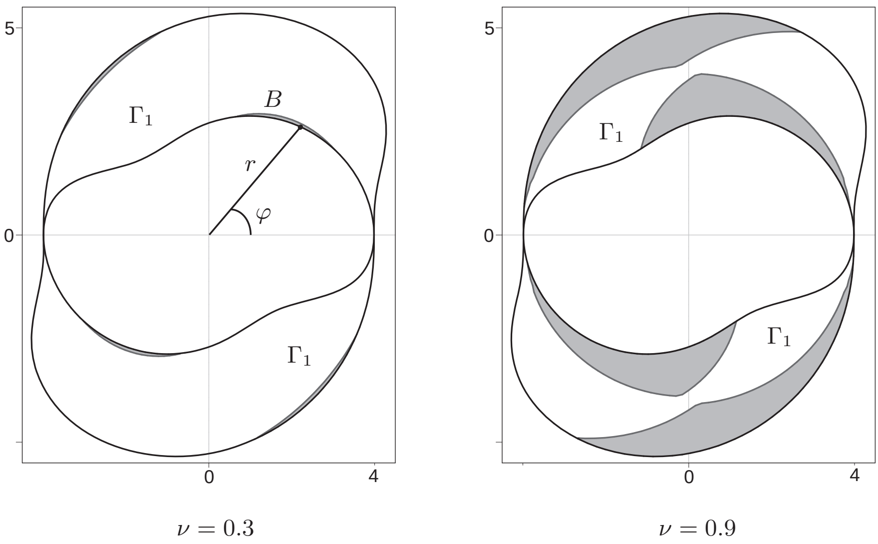

Figure 4: Appearance of the secondary domain for $\varepsilon = 0.75$.

4.2 Game on two-sided plane ellipses

The differential games on two-sided plane ellipses has been partially studied before in [11]. The domain $\partial\Gamma$ has been built in elliptic coordinates and the curve like in Fig. 5 has been also defined. In this paper, we will construct the edge B of the secondary domain and thereby give an answer to a question when singular trajectories should be taken into account.

Construction of the geodesic lines. Let's consider a two-sided plane ellipse $\mathcal{E}_1$. The players can move from one side of the ellipse to the other throughout the edge—a line of ellipse. As it was taken before, the semiminor axis equals one and the parametrization is defined through the eccentricity ε.

When two players P and E are on the same side of the ellipse, the shortest (geodesic) line is the segment PE. In this case, an optimal strategy of the players is to move along this segment. This is proved for the differential games with empty secondary domain. Assume that the players are now on the different sides of the ellipse $\mathcal{E}_1$. Then the geodesic line connecting them is the broken line PME with the break point M on the edge of the ellipse. Note that two different geodesic lines with equal length can exist and connect players only in this situation, when they are on different sides. In particular, when the players P and E are at the different foci of the ellipse, there exist the infinitely many geodesics connecting them. In other cases, there are either two or one geodesic line.

For the construction of the geodesic line PME, one can apply a lot of algebraic and geometrical methods. On the one hand, we can parameterize the boundary of

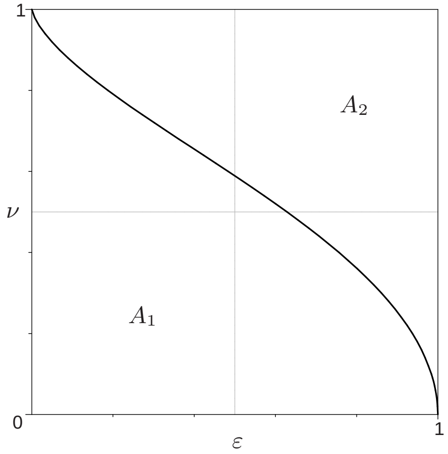

Figure 5: Partition of the parameters' space.

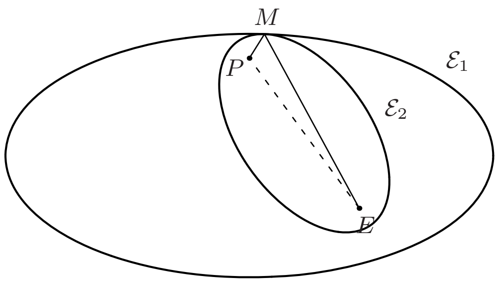

Figure 6: Construction of the geodisic on a two-sided plane ellipse.

the ellipse $\mathcal{E}_1$ by means of the value of α and find the geodesic line as the global minimum of the length of PME, $L_{PME} = \min L_{PM_\alpha E}$. On the other hand, we can use the optical interpretation for which the geodesic line is the shortest way of the light. However, the most convenient method to use is the following one. Fix a position of the players on the ellipse and consider the confocal family of the ellipses with foci in the points P and E. We can see that the geodesic line PME can be defined through the ellipse $\mathcal{E}_2$ which is only tangent to $\mathcal{E}_1$ and does not have any points of intersection. The point of the tangency is the break point of the geodesic line. If there are several points of the tangency there are several number of geodesic lines correspondingly. Their lengths are equal to each other by one of the properties of an ellipse.

Construction of the edge B. In the sequel, we represent the construction results of the manifold B using the following statement. Two points points P and E on the ellipse can be connected by two different geodesics if and only if: $1)$ they are on the different sides of the ellipse; $2)$ they belong to the ellipse $\mathcal{E}_b$ with semiminor axis $0 \leq b < 1$ which is confocal to the $\mathcal{E}_1$; and $3)$ the tangent lines to the ellipse $\mathcal{E}_b$ at the points P and E intersect outside of the ellipse $\mathcal{E}_1$. This statement does not have yet an analytical proof, but was checked by intense numerical simulations.

Consider a couple of points on the ellipse $\mathcal{E}_b$ and emit the tangent lines from these points; see Fig. 7. They will intersect at the point N outside of the ellipse $\mathcal{E}_1$.

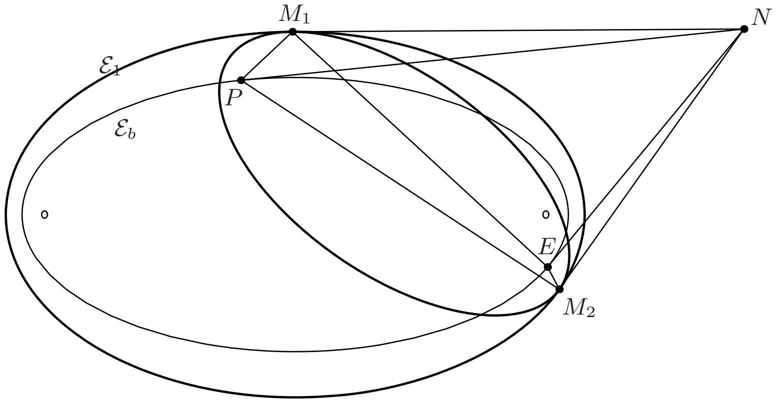

Figure 7: Construction of the pair of geodesics with equal length.

Let us construct the couple of the tangent lines to the ellipse $\mathcal{E}_1$ emitted from the point N, the points of the tangency $M_{1,2}$ will be the break points of the geodesic lines $PM_{1,2}E$ with equal length.

Therefore, the boundary of the domain Γ consists of the points such that the tangent lines emanated from them intersect on the ellipse $\mathcal{E}_1$. To construct the edge B, one can use the above-mentioned geometrical fact. The result of numerical computations for various values of parameters is shown in Fig. 8 in elliptic coordinates. We can see the evolution and expansion of the secondary domain in the game space. Note that it tends to some limit.

5 Conclusions

In this paper, the differential games with simple motion on 2D surfaces (manifolds) are considered. The structure of the solution of such games may be a simple one (with no singular trajectories, possibly with dispersal surfaces) or a complicated one (with singular trajectories and surfaces). The existence of singular trajectories depends upon the specific shape (parameters) of the game space.

We have constructed the full solution of the differential games of the first (simple) kind, for which the so called secondary domain is empty. The problems of the second (complicated) kind have found a partial solution.

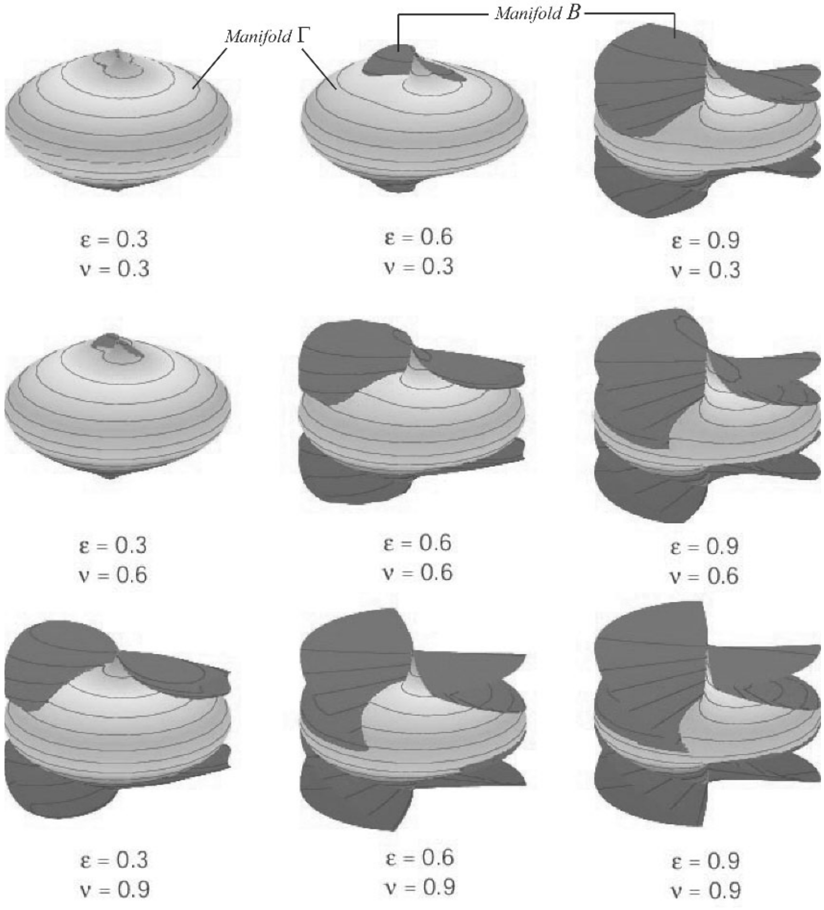

Figure 8: Evolution of the manifold B.

Acknowledgements

The author thanks Prof. Arik Melikyan for helpful discussions. This research has been partially supported by the Russian Foundation for Basic Researches, grant N 04-01-00610.

References

[1] Isaacs R., *Differential Games*, Wiley, New York, (1965).

[2] Vishnevetsky L.S., Melikyan A.A., *Optimal Pursuit on a Plane in the Presence of an Obstacle*, J. Appl. Maths Mechs (PMM), **46**, N 4 (1982), 485–490.

[3] Pozharitsky G.K., *Isaacs' Problem of Island Enveloping*, J. Appl. Maths Mechs (PMM), **46**, N 5 (1982), 707–713.

[4] Melikyan A.A., Hovakimyan N.V., *Game Problem of Simple Pursuit on Two-Dimensional Cone,* J. Appl. Maths Mechs (PMM), **55**, N 5 (1991), 741–751.

[5] Melikyan A.A., Hovakimyan N.V., Harutunian L.L., *Games of simple pursuit and approach on two-dimensional cone,* JOTA, **98**, N 3 (1998), 515–543.

[6] Hovakimyan N.V., Harutunian L.L., *Game problems on rotation surfaces,* in Petrosjan, L.A. (ed.), *Game Theory and Applications*, **IV** (1998), 62–74.

[7] Hovakimyan N., Melikyan A.A., *Geometry of Pursuit-Evasion on Second Order Rotation Surfaces*, Dynamics and Control, **10** (2000), 297–312.

[8] Melikyan A.A., *Some Properties of the Bellman-Isaacs equation for the games on the surfaces of revolution*, International Games Theory Review, **6**, N 1 (2001), 171–179.

[9] Kovshov A. *The simple pursuit by a few objects on the multidimensional sphere*, in L. Petrosjan and V. Mazalov (eds.), *Game Theory & Applications II*, Nova Science Publishers, New York, (1996), 27–36.

[10] Melikyan A.A., *Primary strategies of simple pursuit in differential games on two-sided plane figures*, J. Appl. Maths Mechs (PMM), **68**, N 4 (2004), 545–554.

[11] Melikyan A.A., *Geometry of Pursuit-Evasion Games on Two-Dimensional Manifolds,* S. Jørgensen, M. Quincampoix, T.L. Vincent (eds.), *Annals of the International Society of Dynamic Games*, **9**, Birkhäuser, Boston, (2007).

[12] Melikyan A.A., *Generalized Characteristics of First Order PDEs: Applications in Optimal Control and Differential Games*, Birkhäuser, Boston, (1998).

[13] Mishchenko A., Fomenko A., *A course of differential geometry and topology*, Mir Publishers, Moscow, (2000).

[14] Gelfand I.M., Fomin S.V., *Calculus of Variations*, Prentice-Hall, Englewood Cliffs, New Jersey, (1963).

Strategies for Alternative Pursuit Games

Igor Shevchenko
North Pacific Fisheries Research Center (TINRO-Center)
Vladivostok, Russia
igor@tinro.ru

Abstract

A pursuit game is called alternative if it may be terminated by the pursuer on one of several given terminal manifolds and corresponding payoffs of Boltza type associated with those variants differ only in their terminal parts. For the case of only two alternatives, we describe an approach for construction of the strategies when solutions for all games with fixed targeted terminal alternatives are known. Some alternative dominates the second if it has a less cost for the pursuer. We assume that the pursuer targets a particular alternative only if the dominating condition is stable along the respective optimal trajectory. Thus, it remains to study alternative pursuit games only in the subdomains where the dominating conditions are unstable. When the playing space has a specific structure, the state has to leave these subdomains through a manifold with matched alternatives. To get a solution, three interconnected auxiliary games are to be set up. We present the Hamilton-Jacobi-Isaacs equations for these games. Also, we describe a procedure to construct locally gradient strategies that optimize the local growth of approximations for the minimum of the alternative value functions.

Key words. Targeted alternatives, dominating conditions, Hamilton-Jacobi-Isaacs equations, locally gradient strategies.

AMS Subject Classifications. Primary 49N75; Secondary 91A24.

1 Definitions and assumptions

Let the *state* be described by the vector $\mathbf{z} \in \mathsf{Z}$ in the *playing space* $\mathsf{Z} \subseteq \mathbb{R}^n$, and obey the following equation:

$$\dot{\mathbf{z}}(t) = \mathbf{f}(\mathbf{z}(t), \mathbf{u}_p(t), \mathbf{u}_e(t)), \quad \mathbf{z}(0) = \mathbf{z}^0 \ , \tag{1}$$

where $\mathbf{f} : \mathsf{Z} \times \mathsf{U}_p \times \mathsf{U}_e \to \mathsf{Z}$, $\mathbf{z}^0 \in \mathsf{Z}$, and the control histories $\mathbf{u}_p(\cdot) = \{\mathbf{u}_p(t)\}_{t \geq 0}$ and $\mathbf{u}_e(\cdot) = \{\mathbf{u}_e(t)\}_{t \geq 0}$ belong to the sets of measurable (piecewise continuous) functions Ω_{U_p} and Ω_{U_e} taking values in the compact sets $\mathsf{U}_p \subset \mathbb{R}^{n_p}$ and $\mathsf{U}_e \subset \mathbb{R}^{n_e}$. Suppose that $\mathbf{f}$ is continuous function of all its variables, locally Lipschitz in $\mathbf{z}$, and meets the inequality $|\mathbf{f}(\mathbf{z}, \mathbf{u}_p, \mathbf{u}_e)| \leq \varkappa(1 + |\mathbf{z}|)$, $\varkappa = const > 0$, $\mathbf{z} \in \mathsf{Z}$.

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_7,

Let *strategies* be rules to select values of control variables with use of available information (e.g., *open-loop controls, pure feedbacks, memory strategies, discriminating feedbacks* [1], [2]). Let any pair of admissible strategies $\mathcal{S} = (\mathcal{S}_p, \mathcal{S}_e)$ generate through (1) the *pair of control histories* $\mathbf{u}_p^{\mathcal{S}}(\cdot, \mathbf{z}^0) = \{\mathbf{u}_p(t, \mathbf{z}^0, \mathcal{S})\}_{t\geq 0} \in \Omega_{\mathsf{U}_p}$ and $\mathbf{u}_e^{\mathcal{S}}(\cdot, \mathbf{z}^0) = \{\mathbf{u}_e(t, \mathbf{z}^0, \mathcal{S})\}_{t\geq 0} \in \Omega_{\mathsf{U}_e}$, as well as the trajectory $\mathbf{z}^{\mathcal{S}}(\cdot, \mathbf{z}^0) = \{\mathbf{z}(t, \mathbf{z}^0, \mathcal{S})\}_{t\geq 0} \in \mathsf{Z}$, that is absolutely continuous and satisfies (1) with $\mathbf{u}_p(t) = \mathbf{u}_p^{\mathcal{S}}(t, \mathbf{z}^0)$ and $\mathbf{u}_e(t) = \mathbf{u}_p^{\mathcal{S}}(t, \mathbf{z}^0)$ almost everywhere for $t > 0$.

Let P can terminate the game on any of two given manifolds $\mathsf{M}_a \subset \mathsf{Z}$ and $\mathsf{M}_b \subset \mathsf{Z}$ of class C^2 locally. Let $\mathbf{z}^0 \in \mathsf{Z}$, and $\mathcal{S} = (\mathcal{S}_p, \mathcal{S}_e)$ be a pair of the admissible strategies [1], [2]. Let $\tau_a^{\mathcal{S}}(\mathbf{z}^0)$ and $\tau_b^{\mathcal{S}}(\mathbf{z}^0)$ be the instants of reaching M_a and M_b along the corresponding trajectories:

$$\tau_l^{\mathcal{S}}(\mathbf{z}^0) = \begin{cases} \min\{t \geq 0 : \ \mathbf{z}^{\mathcal{S}}(t, \mathbf{z}^0) \in \mathsf{M}_l\} & \textit{if } \exists t \geq 0 : \ \mathbf{z}^{\mathcal{S}}(t, \mathbf{z}^0) \in \mathsf{M}_l, \\ +\infty & \textit{otherwise}, \quad l \in L = \{a, b\} \ . \end{cases} \tag{2}$$

Let $K^l : \mathsf{M}_l \to \mathbb{R}^+$ and $\mathcal{L}_{ab} : \mathsf{Z} \times \mathsf{U}_p \times \mathsf{U}_e \to \mathbb{R}^+$ be of class C^1, and the *performance indices (payoffs)* $\mathcal{P}_a$ and $\mathcal{P}_b$ be of Boltza type with the common integrand $\mathcal{L}_{ab}$

$$\mathcal{P}_l(\mathbf{z}^0, \mathcal{S}_p, \mathcal{S}_e) = \begin{cases} K^l(\mathbf{z}(\tau_l^{\mathcal{S}}(\mathbf{z}^0), \mathbf{z}^0)) + \int_0^{\tau_l^{\mathcal{S}}(\mathbf{z}^0)} \mathcal{L}_{ab}^{\mathcal{S}}(t, \mathbf{z}^0) dt & \\ & \textit{if } \tau_l^{\mathcal{S}}(\mathbf{z}^0) \textit{ is finite}, \\ +\infty & \textit{otherwise}, \quad l \in L \ , \end{cases} \tag{3}$$

where $\mathcal{L}_{ab}^{\mathcal{S}}(t, \mathbf{z}^0) = \mathcal{L}_{ab}(\mathbf{z}^{\mathcal{S}}(t, \mathbf{z}^0), \mathbf{u}_p^{\mathcal{S}}(t, \mathbf{z}^0), \mathbf{u}_e^{\mathcal{S}}(t, \mathbf{z}^0)), \ t \geq 0$. Suppose also that K^a and K^b coincide on

$$\mathsf{M}_{ab} = \mathsf{M}_a \cap \mathsf{M}_b \ .$$

By $\mathfrak{G}$, denote the game with the terminal manifold

$$\mathsf{M}_{a|b} = \mathsf{M}_a \cup \mathsf{M}_b \tag{4}$$

and the payoff

$$\mathcal{P}(\mathbf{z}^0, \mathcal{S}_p, \mathcal{S}_e) = \begin{cases} \mathcal{P}_{l^{\mathcal{S}}(\mathbf{z}^0)}(\mathbf{z}^{\mathcal{S}}(t, \mathbf{z}^0), \mathbf{u}_p^{\mathcal{S}}(t, \mathbf{z}^0), \mathbf{u}_e^{\mathcal{S}}(t, \mathbf{z}^0)) & \\ & \textit{if } l^{\mathcal{S}}(\mathbf{z}^0) \in L, \\ +\infty, & \textit{if } l^{\mathcal{S}}(\mathbf{z}^0) \textit{ undefined} \ , \end{cases}$$

where

$$l^{\mathcal{S}}(\mathbf{z}^0) = \begin{cases} l & \textit{if } \tau_l^{\mathcal{S}}(\mathbf{z}^0) = \min(\tau_a^{\mathcal{S}}(\mathbf{z}^0), \tau_b^{\mathcal{S}}(\mathbf{z}^0)) \textit{ and } \tau_l^{\mathcal{S}}(\mathbf{z}^0) \\ & \textit{is finite}, \\ \textit{undefined} & \textit{otherwise} \ . \end{cases}$$

All above-mentioned functions of time generated with use of pairs of the admissible strategies are supposed to be defined, unique, and indefinitely continuable in time.

Let us associate with $\mathfrak{G}$ two games, $\mathfrak{G}_a$ and $\mathfrak{G}_b$, that correspond to the *fixed targeted termination alternatives*, M_a *and* M_b. Let $\tau_a^{\mathcal{S}}(\mathbf{z}^0)$ and $\tau_b^{\mathcal{S}}(\mathbf{z}^0)$ be the termination instants (see (2)), and $\mathcal{P}_a$ and $\mathcal{P}_b$ be the payoffs in $\mathfrak{G}_a$ and $\mathfrak{G}_b$; see (3). Also, suppose that $\mathfrak{G}_{a|b}$ is the game with the *open-loop targeted termination alternative* that has $\mathsf{M}_{a|b}$ (see (4)) as the terminal manifold and the payoff $\mathcal{P}_{a|b}$ equal to $\mathcal{P}_{l^0(\mathbf{z}^0)}$ (see (3)) with $l^0(\mathbf{z}^0) \in L$ assigned by P. Unlike $\mathfrak{G}$ where the value of accomplished alternative $l^{\mathcal{S}}(\mathbf{z}^0)$ depends on the strategies chosen by *all* players, $l^0(\mathbf{z}^0)$ is an additional control parameter for P in $\mathfrak{G}_{a|b}$.

Let solutions of $\mathfrak{G}_a$ and $\mathfrak{G}_b$ be known, and $\mathcal{S}^l = (\mathcal{S}_p^l, \mathcal{S}_e^l)$ be the pair of optimal strategies of the players, $V^l(\mathbf{z}^0)$ be the value function, $\mathbf{z}_l(t, \mathbf{z}^0)$ be the optimal trajectory, and $\tau_l(\mathbf{z}^0)$ be the optimal duration of $\mathfrak{G}_l$ for $\mathbf{z}^0 \in \mathsf{Z}$, $l \in L$. Let D^0 be the part of playing space where both alternatives are matched:

$$\mathsf{D}^0 = \{\mathbf{z}^0 : V^a(\mathbf{z}^0) = V^b(\mathbf{z}^0)\}. \tag{5}$$

Let Z_a^0 be the part of the playing space where a dominates b:

$$\mathsf{Z}_a^0 = \{\mathbf{z}^0 :\ V^a(\mathbf{z}^0) < V^b(\mathbf{z}^0)\}\ .$$

Also let, Z_a and Z_b be the parts of the playing space where the domination condition for a or b is stable along the optimal trajectory of $\mathfrak{G}_a$ or $\mathfrak{G}_b$:

$$\mathsf{Z}_a = \{\mathbf{z}^0 :\ V^a(\mathbf{z}_a(t,\mathbf{z}^0)) < V^b(\mathbf{z}_a(t,\mathbf{z}^0)) \& \mathbf{z}_a(t,\mathbf{z}^0) \notin \mathsf{M}_b, \forall t \in [0, \tau_a(\mathbf{z}^0)]\},$$
$$\mathsf{Z}_b = \{\mathbf{z}^0 :\ V^b(\mathbf{z}_b(t,\mathbf{z}^0)) < V^a(\mathbf{z}_b(t,\mathbf{z}^0)) \& \mathbf{z}_b(t,\mathbf{z}^0) \notin \mathsf{M}_a, \forall t \in [0, \tau_b(\mathbf{z}^0)]\}\ .$$

We distinguish a *negligible violation* of the initial dominating condition for a along the optimal trajectory of $\mathfrak{G}_a$ starting at $\mathbf{z}^0$ when

$$\exists t \in [0, \tau_a(\mathbf{z}^0)] :\ V^a(\mathbf{z}_a(t,\mathbf{z}^0)) = V^b(\mathbf{z}_a(t,\mathbf{z}^0))\ ,$$

and a *substantial violation* when

$$\exists \varepsilon > 0\ \exists t' \in [0, \tau_a(\mathbf{z}^0) - \varepsilon]\ \forall t \in [t', t' + \varepsilon] :\ V^a(\mathbf{z}_a(t,\mathbf{z}^0)) > V^b(\mathbf{z}_a(t,\mathbf{z}^0))\ .$$

Let B_a and $\mathsf{Z}_{\bar{a}}$ be the parts of the playing space where the initial dominating condition for a is violated negligibly and substantially. Denote by $\overline{\mathsf{Z}}_a$ the closure of Z_a. Define Z_b^0, Z_b, $\overline{\mathsf{Z}}_b$, and $\mathsf{Z}_{\bar{b}}$ similarly. Let $\mathsf{Z}_{a|b} = \overline{\mathsf{Z}}_a \cup \overline{\mathsf{Z}}_b$, and $\mathsf{Z}_{\overline{a|b}}$ be the compliment of $\mathsf{Z} \backslash \mathsf{Z}_{a|b}$.

Let $\mathfrak{G}_1$ and $\mathfrak{G}_2$ be two games of players P_1 and E_1 and P_2 and E_2 with playing spaces Z_1 and Z_2. Let $\mathcal{S}_{P_1}(\mathbf{z})$ and $\mathcal{S}_{P_2}(\mathbf{z}_2)$ be the feedback optimal strategies of P_1 in $\mathfrak{G}_1$ and P_2 in $\mathfrak{G}_2$, and $\mathbf{g}$ be a mapping from Z_2 to Z_1. In this case, we call

$\mathfrak{G}_2$ *strategically equivalent to* $\mathfrak{G}_1$ *for the pair of players* (P_1, P_2) *in* Z_2 *according to the mapping* $\mathbf{g}$ and denote it as

$$\mathfrak{G}_2 \underset{(P_2,P_1)}{\overset{(\mathsf{Z}_2,\mathbf{g})}{\sim}} \mathfrak{G}_1$$

if $\mathcal{S}_{P_2}(\mathbf{z}_2) = \mathcal{S}_{P_1}(\mathbf{g}(\mathbf{z}_2))$. If, additionally,

$$\mathfrak{G}_2 \underset{(E_2,E_1)}{\overset{(\mathsf{Z}_2,\mathbf{g})}{\sim}} \mathfrak{G}_1 \ ,$$

we express these two relations as:

$$\mathfrak{G}_2 \underset{\{(P_2,P_1),(E_2,E_1)\}}{\overset{(\mathsf{Z}_2,\mathbf{g})}{\sim}} \mathfrak{G}_1 \ .$$

If $\mathbf{z}^0 \in \mathsf{Z}_l$, then:

$$\forall t \in [0, \tau_l(\mathbf{z}^0)] : \ l^{\mathcal{S}^l}(z_l(t, \mathbf{z}^0)) = l, \quad l \in L.$$

Therefore, it is reasonable to assume that $\mathfrak{G}$ is strategically equivalent to $\mathfrak{G}_{a|b}$ for both pairs of players $\{(P, P), (E, E)\}$ in $\mathsf{Z}_{a|b}$ according to the mapping $g_\equiv$ where $\mathbf{g}_\equiv(\mathbf{z}) = \mathbf{z}, \ \forall \mathbf{z} \in \mathsf{Z}_{a|b}$, i.e.,

$$\mathfrak{G} \underset{\{(P,P),(E,E)\}}{\overset{(\mathsf{Z}_{a|b},\mathbf{g}_\equiv)}{\sim}} \mathfrak{G}_{a|b} \ . \tag{6}$$

It means that if $\mathbf{z}^0 \in \mathsf{Z}_{a|b}$, then P *targets the alternative ensuring the least cost at* $\mathbf{z}^0$ *or any of them when both options are matched*. In accordance with (6),

$$V(\mathbf{z}^0) = V^{a|b}(\mathbf{z}^0), \ \forall \mathbf{z}^0 \in \mathsf{Z}_{a|b} \ , \tag{7}$$

where V and $V^{a|b}$ are the value functions of $\mathfrak{G}$ and $\mathfrak{G}_{a|b}$. Note that by definition of $\mathfrak{G}_{a|b}$,

$$V^{a|b}(\mathbf{z}^0) = \min(V^a(\mathbf{z}^0), V^b(\mathbf{z}^0)) \ . \tag{8}$$

Let $\mathbf{z}^0 \in \mathsf{Z}_{\overline{a|b}} \neq \emptyset$, and a pair of the admissible strategies $\mathcal{S} = (\mathcal{S}_p, \mathcal{S}_e)$ be chosen by the players. By definition, B_l separates Z_l and $\mathsf{Z}_{\bar{l}}$, $l \in L$. Let $\tau^{\mathcal{S}}_{B_a}(\mathbf{z}^0)$ and $\tau^{\mathcal{S}}_{B_b}(\mathbf{z}^0)$ be the instants when the state reaches B_a and B_b along the trajectory $\mathbf{z}^{\mathcal{S}}(\cdot, \mathbf{z}^0)$, and (see (2))

$$\tau^{\mathcal{S}}_{\min}(\mathbf{z}^0) = \min(\tau^{\mathcal{S}}_a(\mathbf{z}^0), \tau^{\mathcal{S}}_b(\mathbf{z}^0), \tau^{\mathcal{S}}_{B_a}(\mathbf{z}^0), \tau^{\mathcal{S}}_{B_b}(\mathbf{z}^0)) \ .$$

Let $\mathfrak{G}_{\overline{a|b}}$ be the game with the playing space $\mathsf{Z}_{\overline{a|b}}$, the terminal manifold $\partial \mathsf{Z}_{\overline{a|b}} = \mathsf{M}_a \cup \mathsf{M}_b \cup \mathsf{B}_a \cup \mathsf{B}_b$, and the payoff $\mathcal{P}_{\overline{a|b}}$ be described as:

$$\mathcal{P}_{\overline{a|b}}(\mathbf{z}^0, \mathcal{S}_p, \mathcal{S}_e) = \begin{cases} \mathcal{P}_l(\mathbf{z}^0, \mathcal{S}_p, \mathcal{S}_e) & \textit{if } \tau^{\mathcal{S}}_{M_l}(\mathbf{z}^0) = \tau^{\mathcal{S}}_{\min}(\mathbf{z}^0), \\ V^l(\mathbf{z}^{\mathcal{S}}(\tau^{\mathcal{S}}_{B_l}(\mathbf{z}^0), \mathbf{z}^0)) + \int_0^{\tau^{\mathcal{S}}_{B_l}(\mathbf{z}^0)} \mathcal{L}^{\mathcal{S}}_{ab}(t, \mathbf{z}^0) dt & \\ & \textit{if } \tau^{\mathcal{S}}_{B_l}(\mathbf{z}^0) = \tau^{\mathcal{S}}_{\min}(\mathbf{z}^0), \ l \in L \ . \end{cases}$$

In $\mathsf{Z}_{\overline{a|b}}$, a targeted alternative could not be fixed because neither dominating condition is stable there. Let us assume that

$$\mathfrak{G} \overset{(\mathsf{Z}_{\overline{a|b}},\mathbf{g}_{\equiv})}{\underset{\{(P,P),(E,E)\}}{\sim}} \mathfrak{G}_{\overline{a|b}} \,,$$

and study $\mathfrak{G}_{\overline{a|b}}$ under some additional hypotheses. Namely, assume that (see (5)):

$$\begin{aligned}&\mathrm{sign}[(V^a_{\mathbf{z}}(\mathbf{z}^0) - V^b_{\mathbf{z}}(\mathbf{z}^0))\mathbf{f}(\mathbf{z}^0, \mathbf{u}_p^{\mathcal{S}^a}(0,\mathbf{z}^0), \mathbf{u}_e^{\mathcal{S}^a}(0,\mathbf{z}^0))] = \\ &\mathrm{sign}[(V^b_{\mathbf{z}}(\mathbf{z}^0) - V^a_{\mathbf{z}}(\mathbf{z}^0))\mathbf{f}(\mathbf{z}^0, \mathbf{u}_p^{\mathcal{S}^b}(0,\mathbf{z}^0), \mathbf{u}_e^{\mathcal{S}^b}(0,\mathbf{z}^0))],\ \forall \mathbf{z}^0 \in \mathsf{D}^0.\end{aligned} \tag{9}$$

Also, consider only Z that has the structure shown in Fig. 1, where $\mathsf{D}^0 = \mathsf{F} \cup \mathsf{A} \cup \mathsf{D}$, and F correspond to positive, A to zero, and D to negative signs in (9); see [8] for more detail.

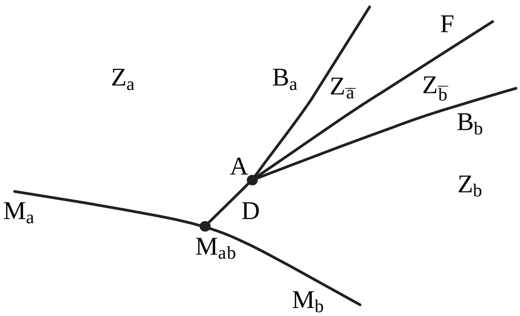

Figure 1: Decomposition of the playing space in $\mathfrak{G}$.

2 The Hamilton-Jacobi-Isaacs equations

Let $\mathbf{z}^0 \in \mathsf{F}$, $\mathcal{S} = (\mathcal{S}_p, \mathcal{S}_e)$ be the pair of strategies chosen by the players, $\nu^{\mathcal{S}} = \nu^{\mathcal{S}}(\mathbf{z}^0, \mathbf{u}_p^{\mathcal{S}}(0,\mathbf{z}^0), \mathbf{u}_e^{\mathcal{S}}(0,\mathbf{z}^0))$ be the instant direction of corresponding shift. Then,

$$V^a_{\nu^{\mathcal{S}}}(\mathbf{z}^0) > V^b_{\nu^{\mathcal{S}}}(\mathbf{z}^0) \text{ if } \nu^{\mathcal{S}} \text{ directed to } \mathsf{Z}_{\overline{a}}$$

and

$$V^b_{\nu^{\mathcal{S}}}(\mathbf{z}^0) > V^a_{\nu^{\mathcal{S}}}(\mathbf{z}^0) \text{ if } \nu^{\mathcal{S}} \text{ directed to } \mathsf{Z}_{\overline{b}}\,.$$

Therefore, E should choose the strategy $\mathcal{S}_e^F$ that being coupled with a strategy $\mathcal{S}_p^F$ of P on F provides

$$V^a_{\nu^{(\mathcal{S}_p^F,\mathcal{S}_e^F)}}(\mathbf{z}^0) = V^b_{\nu^{(\mathcal{S}_p^F,\mathcal{S}_e^F)}}(\mathbf{z}^0)\,,$$

and keep the state on F by this choice; see, e.g., [1], [3], [7]. Moreover, along the optimal trajectories starting in $\mathsf{Z}_{\overline{a|b}}$, the state first has to go on F and then directly to $\mathsf{A} \subset \mathsf{B}_a \cap \mathsf{B}_b$ where

$$\mathfrak{G}_{\overline{a|b}} \overset{(\mathsf{A},\mathbf{g}_\equiv)}{\underset{\{(P,P),(E,E)\}}{\sim}} \mathfrak{G}_l, \quad l \in L.$$

Thus, in the backward construction of the solution of $\mathfrak{G}_{\overline{a|b}}$, three additional games—$\mathfrak{G}_{\overline{a}}$, $\mathfrak{G}_{\overline{b}}$, and $\mathfrak{G}_F$—emerge, and the following relations hold:

$$\mathfrak{G}_{\overline{a|b}} \overset{(\mathsf{F},\mathbf{g}_\equiv)}{\underset{\{(P,P),(E,E)\}}{\sim}} \mathfrak{G}_F, \qquad \mathfrak{G}_{\overline{a|b}} \overset{(\mathsf{Z}_{\bar{l}},\mathbf{g}_\equiv)}{\underset{\{(P,P),(E,E)\}}{\sim}} \mathfrak{G}_{\bar{l}}, \quad l \in L.$$

Let

$$H(\mathbf{z}, \mathbf{u}_p, \mathbf{u}_e, \lambda) = \lambda \mathbf{f}(\mathbf{z}, \mathbf{u}_p, \mathbf{u}_e) + \mathcal{L}_{ab}(\mathbf{z}, \mathbf{u}_p, \mathbf{u}_e) \tag{10}$$

be the Hamiltonian. The HJI equations [2], [5]

$$\begin{aligned}
\min_{\mathbf{u}_p \in \mathsf{U}_p} \max_{\mathbf{u}_e \in \mathsf{U}_e} H(\mathbf{z}_{\bar{l}}, \mathbf{u}_p, \mathbf{u}_e, \lambda_{\bar{l}}) &= H(\mathbf{z}_{\bar{l}}, \mathbf{u}_p^{\bar{l}}(\mathbf{z}_{\bar{l}}, \lambda_{\bar{l}}), \mathbf{u}_e^{\bar{l}}(\mathbf{z}_{\bar{l}}, \lambda_{\bar{l}}), \lambda_{\bar{l}}) \\
&= H^{\bar{l}}(z_{\bar{l}}, \lambda_{\bar{l}}) = 0 \quad \text{for } \mathfrak{G}_{\bar{l}}, \\
\min_{\mathbf{u}_p \in \mathsf{U}_p} \max_{\mathbf{u}_e \in \mathsf{U}_e:\, \mathbf{z} \in \mathsf{F}} H(\mathbf{z}_F, \mathbf{u}_p, \mathbf{u}_e, \lambda_F) &= H(\mathbf{z}_F, \mathbf{u}_p^F(\mathbf{z}_F, \lambda_F), \mathbf{u}_e^F(\mathbf{z}_F, \lambda_F), \lambda_F) \\
&= H^F(z_F, \lambda_F) = 0 \quad \text{for } \mathfrak{G}_F, \\
\min_{\mathbf{u}_p \in \mathsf{U}_p} \max_{\mathbf{u}_e \in \mathsf{U}_e} H(\mathbf{z}_l, \mathbf{u}_p, \mathbf{u}_e, \lambda_l) &= H(\mathbf{z}_l, \mathbf{u}_p^l(\mathbf{z}_l, \lambda_l), \mathbf{u}_e^l(\mathbf{z}_l, \lambda_l), \lambda_l) \\
&= H^l(\mathbf{z}_l, \lambda_l) = 0 \quad \text{for } \mathfrak{G}_l, \quad l \in L \,,
\end{aligned} \tag{11}$$

differ only in the additional restriction which are imposed on the choice of E on F to ensure that the state moves along F. The equations of characteristics

$$\begin{aligned}
\dot{\mathbf{z}}_{\bar{l}} = H_\lambda^*(\mathbf{z}_{\bar{l}}, \lambda_{\bar{l}}), \quad \dot{\mathbf{z}}_F = H_\lambda^F(\mathbf{z}_F, \lambda_F), \quad \dot{\mathbf{z}}_l = H_\lambda^*(\mathbf{z}_l, \lambda_l), \\
\dot{\lambda}_{\bar{l}} = -H_{\mathbf{z}}^*(\mathbf{z}_{\bar{l}}, \lambda_{\bar{l}}), \quad \dot{\lambda}_F = -H_{\mathbf{z}}^F(\mathbf{z}_F, \lambda_F), \quad \dot{\lambda}_l = -H_{\mathbf{z}}^*(\mathbf{z}_l, \lambda_l),
\end{aligned}$$

and the boundary conditions

$$\begin{aligned}
[V^{\bar{l}}(\mathbf{z}) - V^F(\mathbf{z})]|_{\mathbf{z} \in \mathsf{F}} = 0, \\
[V^F(\mathbf{z}) - V^l(\mathbf{z})]|_{\mathbf{z} \in \mathsf{A}} = 0, \\
[V^l(\mathbf{z}) - K^l(\mathbf{z})]|_{\mathbf{z} \in \mathsf{M}_l} = 0 \,,
\end{aligned} \tag{12}$$

may be used when one attempts to solve (11) for $\mathfrak{G}_{\bar{l}}$, $\mathfrak{G}_F$ and $\mathfrak{G}_l$, $l \in L$. Then, the adjoint variables are the gradients of the value functions along the optimal trajectories

$$\lambda_{\bar{l}} = V_{\mathbf{z}}^{\bar{l}}, \quad \lambda_F = V_{\mathbf{z}}^F, \quad \lambda_l = V_{\mathbf{z}}^l,$$

and the value function V in $\mathfrak{G}$ is evaluated as:

$$V(\mathbf{z}) = \begin{cases} V^{\bar{l}}(\mathbf{z}) & \text{if } \mathbf{z} \in \mathsf{Z}_{\bar{l}}, \\ V^F(\mathbf{z}) & \text{if } \mathbf{z} \in \mathsf{F}, \\ V^l(\mathbf{z}) & \text{if } \mathbf{z} \in \mathsf{Z}_l,\ l \in L\ . \end{cases}$$

However, $V^{\bar{a}}$, $V^{\bar{b}}$, and V^F could be hardly obtained. In particular, it is related to the fact that the initially unknown value function V^F defines the boundary conditions for $V^{\bar{a}}$ and $V^{\bar{b}}$; see (12). Also, this function itself is a solution of the game $\mathfrak{G}_F$ with the state constraints (see (11)) that introduce additional complexity [1], [2], [4], [6], [7].

3 Locally gradient strategies

Let us construct approximate strategies that correspond to the locally optimal choice of targeted alternative by P not only in $\mathsf{Z}_{a|b}$ (see (7)) but in $\mathsf{Z}_{\overline{a|b}}$ as well. For this purpose we need somehow to approximate $V^{a|b}(\mathbf{z})$ (see (8)) and its gradient (where it exists). Follow [9], [10]; let us use a parameterized family of functions

$$M^\xi(v_1, v_2) = \frac{v_1^\xi v_2 + v_1 v_2^\xi}{v_1^\xi + v_2^\xi}, \quad \xi, v_1, v_2 \in \mathbb{R}^+\ ,$$

that possess some remarkable properties.

Lemma 3.1. $M^\xi(v_1, v_2)$ *provides an upper approximation for* $\min(v_1, v_2)$ *such that:*

$$\min(v_1, v_2) < M^\xi(v_1, v_2) \textit{ if } v_1 \neq v_2,$$
$$\min(v, v) = M^\xi(v, v) = v\ ,$$

and

$$\lim_{\xi \to +\infty} M^\xi(v_1, v_2) = \begin{cases} v_1 & \textit{if } v_1 < v_2 \\ v_2 & \textit{if } v_1 > v_2, \quad \xi, v_1, v_2, v \in \mathbb{R}^+\ . \end{cases} \tag{13}$$

□

The partial derivatives for $M^\xi(v_1, v_2)$ may be written as:

$$M^\xi_{v_i}(v_1, v_2) = \frac{v_{3-i}^\xi (v_i v_{3-i}^\xi + \xi v_i^\xi v_{3-i} + (1-\xi) v_i^{1+\xi})}{v_i (v_1^\xi + v_2^\xi)^2},$$
$$\xi, v_1, v_2 \in \mathbb{R}^+,\ i = 1, 2\ .$$

Lemma 3.2. $M_{v_i}^{\xi}(v_1, v_2)$ *approximates the partial derivative of* $\min(v_1, v_2)$ *(where it exists) such that:*

$$\lim_{\xi \to +\infty} M_{v_i}^{\xi}(v_1, v_2) = \begin{cases} 1 & \text{if } v_i < v_{3-i}, \\ 1/2 & \text{if } v_1 = v_2, \\ 0 & \text{if } v_i > v_{3-i}, \quad \xi, v_1, v_2 \in \mathbb{R}^+, \ i = 1, 2\,. \end{cases} \tag{14}$$

□

Assume that in (11):

$$\begin{aligned} &\mathbf{u}_p^*(\mathbf{z}, \lambda) = \mathbf{u}_p^l(\mathbf{z}, \lambda) = \mathbf{u}_p^{\bar{l}}(\mathbf{z}, \lambda), \\ &\mathbf{u}_e^*(\mathbf{z}, \lambda) = \mathbf{u}_e^l(\mathbf{z}, \lambda) = \mathbf{u}_e^{\bar{l}}(\mathbf{z}, \lambda), \quad \mathbf{z} \in \mathsf{Z}, \lambda \in \mathbb{R}^n, l \in L\,. \end{aligned} \tag{15}$$

Let (compare to (7))

$$W^{\xi}(\mathbf{z}) = M^{\xi}(V^a(\mathbf{z}), V^b(\mathbf{z})), \quad \mathbf{z} \in \mathsf{Z}_{\overline{a|b}}\,. \tag{16}$$

Let us define approximate feedback pursuit $\{\mathcal{S}_p^{\xi}(\mathbf{z})\}_{\xi>0}$ and evasion $\{\mathcal{S}_e^{\xi}(\mathbf{z})\}_{\xi>0}$ strategies as follows (see (15)):

$$\mathcal{S}_p^{\xi}(\mathbf{z}) = \mathbf{u}_p^*(\mathbf{z}, W_{\mathbf{z}}^{\xi}(\mathbf{z})), \quad \mathcal{S}_e^{\xi}(\mathbf{z}) = \mathbf{u}_e^*(\mathbf{z}, W_{\mathbf{z}}^{\xi}(\mathbf{z})), \tag{17}$$

where

$$W_{\mathbf{z}}^{\xi}(\mathbf{z}) = M_{v_1}^{\xi}(V^a(\mathbf{z}), V^b(\mathbf{z}))V_{\mathbf{z}}^a(\mathbf{z}) + M_{v_2}^{\xi}(V^a(\mathbf{z}), V^b(\mathbf{z}))V_{\mathbf{z}}^b(\mathbf{z})\,. \tag{18}$$

Let (see (10))

$$H^{\xi}(\mathbf{z}) = H(\mathbf{z}, \mathcal{S}_p^{\xi}(\mathbf{z}), \mathcal{S}_e^{\xi}(\mathbf{z}), W_{\mathbf{z}}^{\xi}(\mathbf{z})) \tag{19}$$

and $\mathbf{f}, \mathcal{L}_{ab}, \mathbf{u}_p^*$, and $\mathbf{u}_e^*$ be at least continuous functions.

Theorem 3.1. *If* $\mathbf{z}^0 \notin \mathsf{D}^0$ *then* $\lim_{\xi \to +\infty} H^{\xi}(\mathbf{z}^0) = 0$.

Proof. Let $\mathbf{z}^0 \notin \mathsf{D}^0$ and the value functions V^a and V^a take different values at $\mathbf{z}^0$; see (5). Suppose that $V^{a|b}(\mathbf{z}^0) = V^a(\mathbf{z}^0)$; see (8). By definition (see (19)):

$$\begin{aligned} H^{\xi}(\mathbf{z}^0) = W_{\mathbf{z}}^{\xi}(\mathbf{z}^0)\mathbf{f}(\mathbf{z}^0, \mathbf{u}_p^*(\mathbf{z}^0, W_{\mathbf{z}}^{\xi}(\mathbf{z}^0)), \mathbf{u}_e^*(\mathbf{z}^0, W_{\mathbf{z}}^{\xi}(\mathbf{z}^0))) + \\ \mathcal{L}_{ab}(\mathbf{z}^0, \mathbf{u}_p^*(\mathbf{z}^0, W_{\mathbf{z}}^{\xi}(\mathbf{z}^0)), \mathbf{u}_e^*(\mathbf{z}^0, W_{\mathbf{z}}^{\xi}(\mathbf{z}^0)))\,. \end{aligned} \tag{20}$$

By Lemma 3.2 (see also (18)):

$$\lim_{\xi \to +\infty} W_{\mathbf{z}}^{\xi}(\mathbf{z}^0) = V_{\mathbf{z}}^a(\mathbf{z}^0)\,.$$

Therefore, from (11) and (20) we have:

$$\lim_{\xi \to +\infty} H^{\xi}(\mathbf{z}^0) = H^a(\mathbf{z}^0, V_{\mathbf{z}}^a(\mathbf{z}^0)) = 0\,.$$

□

Let $\mathcal{S}_e$ be an admissible strategy chosen by E, and (1) generates the trajectory $\mathbf{z}^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(\cdot,\mathbf{z}^0)$ for every pair of strategies $(\mathcal{S}_p^\xi,\mathcal{S}_e)$, $\forall\xi>0$.

Theorem 3.2. *If* $\mathbf{z}^0\in\mathsf{D}^0$ *and*

$$\lim_{\xi\to+\infty} H^\xi(\mathbf{z}^0) = -k, \quad k>0\ ,$$

then there exist $\tau>0$ *and* $\xi_0>0$ *such that (see* (8)*)*

$$V^{a|b}(\mathbf{z}^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(\tau,\mathbf{z}^0)) < V^{a|b}(\mathbf{z}^0)$$

for every $\xi>\xi_0$ *and for every admissible strategy* $\mathcal{S}_e$*.*

Proof. Along with the trajectory, the pair $(\mathcal{S}_p^\xi,\mathcal{S}_e)$ determines also the players' controls $\mathbf{u}_p^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(t,\mathbf{z}^0)$ and $\mathbf{u}_e^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(t,\mathbf{z}^0)$, the integrand $\mathcal{L}_{ab}^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(t,\mathbf{z}^0)$ and the Hamiltonian

$$H(\mathbf{z}^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(t,\mathbf{z}^0),\mathbf{u}_p^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(t,\mathbf{z}^0),\mathbf{u}_p^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(t,\mathbf{z}^0),W_{\mathbf{z}}^\xi(\mathbf{z}^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(t,\mathbf{z}^0))) \tag{21}$$

as functions of time, $t\geq 0$. Denote (21) as $H^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(t,\mathbf{z}^0)$. Obviously,

$$H^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(t,\mathbf{z}^0) = \frac{d}{dt}W^\xi(\mathbf{z}^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(t,\mathbf{z}^0)) + \mathcal{L}_{ab}^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(t,\mathbf{z}^0)\ .$$

Under the assumptions made above, $H^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(t,\mathbf{z}^0)$ is a continuous function of time, $t\geq 0$. It follows from (11) and (20) that

$$H^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(t,\mathbf{z}^0) \leq H^\xi(\mathbf{z}^{(\mathcal{S}_p^\xi,\mathcal{S}_e^\xi)}(t,\mathbf{z}^0))$$

and there exist $\tau>0$ such that $\forall\xi>\xi_0$

$$H^\xi(\mathbf{z}^{(\mathcal{S}_p^\xi,\mathcal{S}_e^\xi)}(t,\mathbf{z}^0)) < 0, \quad \forall t\in[0,\tau]\ .$$

Since $\mathcal{L}_{ab}^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(t,\mathbf{z}^0)\geq 0$, the following condition holds:

$$\frac{d}{dt}W_{\mathbf{z}}^\xi(\mathbf{z}^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(t,\mathbf{z}^0)) < 0, \quad \forall t\in[0,\tau]\ .$$

Therefore (see Lemma 3.1),

$$V^{a|b}(\mathbf{z}^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(\tau,\mathbf{z}^0)) \leq W^\xi(\mathbf{z}^{(\mathcal{S}_p^\xi,\mathcal{S}_e)}(\tau,\mathbf{z}^0)) < W^\xi(\mathbf{z}^0) = V^{a|b}(\mathbf{z}^0)\ .$$

□

Theorem 1 shows that, at the states with a dominating alternative, the family $\{\mathcal{S}_p^\xi(\mathbf{z})\}_{\xi>0}$ could provide an approximation for the strategy $\mathcal{S}_p^{a|b}(\mathbf{z})$ targeted the locally least cost. According to Theorem 2, at the states with matched alternatives, P should use strategies from the family $\{\mathcal{S}_p^\xi(\mathbf{z})\}_{\xi>0}$ only if $H^\xi(\mathbf{z})<0$, $\forall\xi>\xi_0$.

4 Conclusion

In this paper, we describe the HJI equations that met by exact smooth parts of solutions and an approach for generating approximate strategies in alternative pursuit games. In the latter case, assumed that the players optimize local growth of minimum of the costs corresponding to the fixed termination alternatives. Convergent upper and low approximations of the min function from [9] and known solutions of the games with fixed alternatives are used to construct explicit expressions for the strategies. Dynamic and guaranteed properties of the obtained approximating control laws must be carefully studied for every particular game; see, e.g., [10].

References

[1] Bernhard P., *Singular Surfaces in Differential Games: An Introduction,* in P. Hargedorn, H.W. Knobloch, G.H. Olsder (Eds.), *Differential Games and Applications. Springer Lecture Notes in Information and Control Sciences,* Springer, Berlin, (1977), 1–33.

[2] Bernhard P., *Differential Games: Isaacs Equation,* in M.G. Singh (Ed.), *System and Control Encyclopedia,* Pergamon Press, UK, (1987), 1010–1017.

[3] Breakwell J.V., Bernhard P., *A Simple Game with a Singular Focal Line,* JOTA **64** (2) (1990), 419–428.

[4] Cardliaguet P., Quincampoix M., Saint-Pierre P., *Pursuit Differential Games with State Constraints*, SIAM J. Control Optim. **39** (5) (2001), 1615–1632.

[5] Isaacs R., *Differential Games,* John Wiley and Sons, New York, (1967).

[6] Melikyan A., *Equations for Singular Focal Paths in Dynamic Games Motion,* in J.A. Filar, V. Gaitsgory, F. Imado (eds.) *Proc. of the 7th International Symposium on Dynamic Games and Applications*, Kanagawa, Japan, (1996), 690–695.

[7] Melikyan A., *Generalized Characteristics of First Order PDEs: Applications in Optimal Control and Differential Games,* Birkhäuser, Boston, (1998).

[8] Shevchenko I., *On Reduction of Alternative Pursuit Games*, Game Theory and Applications, **11** (2007), 125–137.

[9] Stipanović D., Sriram, Tomlin C. J., *Strategies for Agents in Multi-Player Pursuit-Evasion Games,* in T.L. Vincent (ed.) *Proc. of the 11th International Symposium on Dynamic Games and Applications,* Tucson, Arizona, (2004), 937–952.

[10] Stipanović D., Melikyan A., Hovakimyan N., *Some Sufficient Conditions for Multi-Player Pursuit-Evasion Games with Continuous and Discrete Observations*, in P. Bernhard, V. Gaitsgory, O. Pourtallier (eds.) *Proc. of the 12th International Symposium on Dynamic Games and Applications*. (CD version). Sophia Antipolis, French Riviera, (2006), 1–11.

Some Sufficient Conditions for Multi-Player Pursuit-Evasion Games with Continuous and Discrete Observations

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

Arik Melikyan
Institute for Problems in Mechanics, Russian Academy of Sciences
Moscow, Russia
melik@ipmnet.ru

Naira Hovakimyan
Department of Aerospace and Ocean Engineering, Virginia Polytechnic Institute and State University
Blacksburg, Virginia, USA
nhovakim@vt.edu

Abstract

For a pursuit-evasion game with two pursuers and two evaders we introduce classes of functions that represent upper and lower approximations of the min function. Using these functions and a Liapunov type of analysis we formulate several sufficient conditions for the guaranteed capture and evasion for the cases with continuous and discrete observation times for the players. Both pursuers and evaders are modeled as nonlinear systems that are affine functions with respect to their control inputs.

Key words. Pursuit and evasion games, Liapunov stability, differential games.

AMS Subject Classifications. Primary 49N75; Secondary 34D20, 49N70.

1 Introduction

Differential games of the pursuit-evasion type have a long history. Optimality of the strategies for the players in a pursuit-evasion game was introduced and studied

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_8,

by Isaacs [8] by formulating the Hamilton-Jacobi-Isaacs (HJI) partial differential equation (PDE). A solution for a two-player differential game with two identical vehicles have been computed in [18] and a convergent approximation method based on the time-dependent HJI equation was formulated in [19]. Concept of solutions of pursuit-evasion games based on a strategy of extremal aiming was considered by Krasovskii and Subbotin [10]. In [1] solutions of the differential game problems were assumed in a viscosity sense [4] and some approximation schemes were proposed to compute solutions for particular differential games. Strategies for the players in some specific more than two players pursuit-evasion games, based on the solutions of the HJI equation, were presented in [6], [2], [14], [21], [11], and [22]. Application of generalized characteristics of PDEs to differential pursuit-evasion games was studied in [15,16]. Strategy that guarantees evasion for the single evader that is based on fast approximations of viscosity solutions of the modified HJI equation [19] using polytopic functions was proposed in [7].

Differential game of encounter with one pursuer and one evader with discrete observation times was introduced in [13]. Later, the case of differential games with two players where one player has incomplete information due to delays was considered in [3]. The results presented in [13] were later extended using a cost of information for more complicated models for the pursuer and the evader in [20]. Methodology for studying pursuit-evasion games with several pursuers and one evader with discrete observations that is based on the comparison with differential games with continuous observation was studied in [17].

Another approach to define strategies in the pursuit-evasion games using a Liapunov type of analysis (see, for example, [9]) was considered in [23]. Using a set of appropriate functions of the norms of relative distances between pursuers and evaders, strategies for the players in multi-player pursuit-evasion games were formulated. The players were represented by the simple kinematic models. One of the most important features presented in this work was the fact that this methodology can be applied to a pursuit-evasion game with arbitrary number of players. In this paper, we extend the results presented in [23] to include more complicated nonlinear models that are affine in control strategies for the players in the game. We restrict the game to the case of two pursuers and two evaders but the methodology may be easily extended to the case of a game with more than four players. Additionally, we propose two sets of functions that provide convergent approximations of the min function in order to formulate strategies that guarantee either capture or evasion of both evaders. Also, conditions are formulated in the case of continuous and discrete observation times [17].

The organization of the paper is as follows. In Sec. 2 we introduce two sets of functions that represent two convergent sequences of the min function. One of the convergent sequences is always less and the other is always greater than the min function. Then, we use a Liapunov type of analysis and the corresponding class of approximation functions to derive strategies that guarantee either capture or evasion of the evaders under continuous observations. In Sec. 3, we provide equivalent results under the assumption of discrete observation times along the lines of the results presented in [3].

2 Pursuit-evasion game

In this section, our first goal is to formulate strategies for two pursuers that would guarantee capture of both evaders in the pursuit-evasion game with two pursuers and two evaders. Before doing so, a lower approximation of the min function is needed. Thus, let us consider a function

$$\phi_\delta(y_1, y_2) = \alpha_\delta\ (y)\|y_1\| + (1 - \alpha_\delta(y))\|y_2\| \tag{1}$$

where y_1 and y_2 are real vectors of arbitrary dimensions and $\|\cdot\|$ denotes the standard Euclidian norm. Also, we assume that $y = [y_1^T, y_2^T]^T$ is a nonzero vector, that is, $\|y\| \neq 0$. Scalar function $\alpha_\delta(y)$ is defined by:

$$\alpha_\delta\ (y) = \frac{\|y_2\|^\delta}{\|y_1\|^\delta + \|y_2\|^\delta} = \frac{1}{1 + \beta(y)^\delta}\ ,\quad \beta\ (y) = \frac{\|y_1\|}{\|y_2\|}\ ,\quad \delta > 0\ , \tag{2}$$

where δ is a given positive constant. Then:

$$\lim_{\delta\to\infty} \alpha_\delta\ (y) = \begin{cases} 1/2\,, & \|y_1\| = \|y_2\| \\ 1\,, & \|y_1\| < \|y_2\| \\ 0\,, & \|y_1\| > \|y_2\| \end{cases} \tag{3}$$

implies

$$\lim_{\delta\to\infty} \phi_\delta(y_1, y_2) = \min\left\{\|y_1\|, \|y_2\|\right\}\ . \tag{4}$$

From Eq. (4) it follows that by increasing δ we can approximate the $\min$ function with an arbitrary accuracy. Additionally, from Eqs. (1) and (2), we obtain

$$\phi_\delta(y_1, y_2) \equiv \frac{\|y_2\|^\delta\|y_1\| + \|y_1\|^\delta\|y_2\|}{\|y_1\|^\delta + \|y_2\|^\delta},\quad \delta > 0 \tag{5}$$

which implies

$$\min\left\{||y_1||\,,\, ||y_2||\right\} \le \phi_\delta(y_1, y_2)\ (\forall\delta)(\delta > 0) \tag{6}$$

since the inequality

$$\left(\min\left\{\|y_1\|, \|y_2\|\right\} - \|y_2\|\right)\|y_1\|^\delta + \left(\min\left\{\|y_1\|, \|y_2\|\right\} - \|y_1\|\right)\|y_2\|^\delta \le 0 \tag{7}$$

is always true. Thus, not only may we approximate the $\min$ function with an arbitrary accuracy but we guarantee that the convergence is “from above." As mentioned at the beginning of this section, this fact would be of particular significance in formulating strategies for the pursuers that would guarantee the capture of both evaders. Now, let us assume that vectors y_1 and y_2 are nonzero vectors and let $a = \|y_1\|$ and $b = \|y_2\|$ be corresponding positive numbers. Then, the following function:

$$\psi(\delta) = \frac{a^\delta b + ab^\delta}{a^\delta + b^\delta} = \frac{b + a\beta^\delta}{1 + \beta^\delta}, \quad \delta > 0 \,, \tag{8}$$

where $\beta = b/a$, is a decreasing function of δ in the case when $a \neq b$. Notice that if $a = b$ then $\psi(\delta) = \min\{a, b\} = a$ for any positive δ and the case needs no further analysis. Now, assuming without any loss of generality that $b = \max\{a, b\}$, in the case when $a \neq b$, one may easily show the following:

$$\psi(\delta_1) > \psi(\delta_2) \Leftrightarrow (b - a)(\beta^{\delta_2} - \beta^{\delta_1}) > 0 \tag{9}$$

which implies that the function $\psi(\delta)$ is a decreasing function with respect to positive values of δ when $b \neq a$. Notice that $\beta = b/a > 1$ since $b > a$. Thus, in addition to the already established properties of the function $\phi_\delta(\cdot, \cdot)$, we proved that $\psi(\delta) = \phi_\delta(\|y_1\|, \|y_2\|)$ is a strictly decreasing function with respect to the positive values of δ when y_1 and y_2 are nonzero vectors such that $\|y_1\| \neq \|y_2\|$.

Now, let us consider a pursuit-evasion game with two evaders (denoted by the index set $\{1, 2\}$) and two pursuers (denoted by the index set $\{3, 4\}$). The agents' dynamic models are given by:

$$\frac{dx_i}{dt} = f(x_i, u_i) = h_i(x_i) + g_i(x_i)u_i, \quad i = 1, 2, 3, 4 \,, \tag{10}$$

where x_i denotes the *i*-th agent position and u_i denotes its input. The inputs are assumed to be continuous functions of the state vector $x = [x_1^T, x_2^T, x_3^T, x_4^T]^T$ that satisfy the following constraints:

$$||u_i|| \leq \mu_i, \;\; \mu_i > 0, \;\; \forall i \in \{1, 2, 3, 4\} \,. \tag{11}$$

In addition, let us consider the following function:

$$v(x) = \phi_\delta(x_1 - x_3, x_1 - x_4) + \phi_\delta(x_2 - x_3, x_2 - x_4) \,. \tag{12}$$

In order to define the strategies for the players, let us consider the problem of optimizing the derivative of the function $v(x)$ [23]:

$$\min_{\substack{\|u_3\| \leq \mu_3 \\ \|u_4\| \leq \mu_4}} \; \max_{\substack{\|u_1\| \leq \mu_1 \\ \|u_2\| \leq \mu_2}} \left\{ \frac{dv}{dt} \right\} \,. \tag{13}$$

Due to the fact that the dynamic systems (10) are affine in control, the control laws for the pursuers and the evaders corresponding to (13) are given by:

$$\begin{aligned} u_i^o(x) &= \arg\max_{\|u_i\|\le\mu_i}\left\{\tfrac{dv}{dt}\right\} = \arg\max_{\|u_i\|\le\mu_i}\left\{\tfrac{\partial v}{\partial x_i} g_i(x_i)u_i\right\} \\ &= \mu_i g_i^T(x_i)\tfrac{\partial v}{\partial x_i}^T \left\|\tfrac{\partial v}{\partial x_i} g_i(x_i)\right\|^{-1},\ i\in\{1,2\} \\ u_i^o(x) &= \arg\min_{\|u_i\|\le\mu_i}\left\{\tfrac{dv}{dt}\right\} = \arg\min_{\|u_i\|\le\mu_i}\left\{\tfrac{\partial v}{\partial x_i} g_i(x_i)u_i\right\} \\ &= -\mu_i g_i^T(x_i)\tfrac{\partial v}{\partial x_i}^T \left\|\tfrac{\partial v}{\partial x_i} g_i(x_i)\right\|^{-1},\ i\in\{3,4\}\,. \end{aligned} \tag{14}$$

It is important to note that without loss of generality we assume that the control laws in Eq. (14) are well defined, that is, $\|\frac{\partial v}{\partial x_i} g_i(x_i)\| \neq 0$, $i\in\{1,2,3,4\}$, along the agents' trajectories. From Eq. (14), it follows that the evaders are maximizing and the pursuers are trying to minimize the growth of the function $v(x)$ which is an upper approximation of the min function. The soft capture is defined in the standard way as the following.

Definition 2.1. The soft capture of the i-th evader by the j-th pursuer, for a given capture radius R, is defined by the condition $\|x_i - x_j\| \le R$.

From this point, we will use term “capture" whenever soft capture occurs. Now, let us assume that the pursuers choose their strategies according to Eq. (14) and formulate the following theorem.

Theorem 2.1. *Assume that the initial condition x_0, at the initial time t_0, is such that the players are outside of the capture region defined by a positive number R and*

$$\frac{dv}{dt} \le -\gamma,\ \ \gamma > 0\,, \tag{15}$$

along the trajectories of the agents' dynamic systems (10) *for the input strategies $u_i = u_i^o(x)$, $i\in\{1,2,3,4\}$ defined in* (14)*. Then the capture of both evaders is guaranteed in a finite time for any feedback strategies of the evaders that satisfy constraints in* (11) *if the pursuers choose strategies $u_i = u_i^o(x)$, $i\in\{3,4\}$ given in* (14).

Proof. First notice that for $u_i = u_i^o(x)$, $i\in\{3,4\}$, and any feedback strategies $u_i(x)$, $i\in\{1,2\}$ of the evaders, from Eqs. (10), (11), (13), (14), and (15) it follows that:

$$\frac{dv}{dt}(u_1(x),u_2(x),u_3^o(x),u_4^o(x)) \le \frac{dv}{dt}(u_1^o(x),u_2^o(x),u_3^o(x),u_4^o(x)) \le -\gamma\,. \tag{16}$$

By integrating Eq. (16) between t_0 and t, we obtain:

$$v(x(t)) \le -\gamma(t-t_0) + v(x_0) \Rightarrow v(x(t)) \le R\ \text{ for }\ t \ge t_c = t_0 + \frac{v(x_0)-R}{\gamma}\,. \tag{17}$$

Since the initial condition implies that the players are outside of the capture region, from Eqs. (5) and (12) it follows that $v(x_0)$ is finite. Thus, t_c is a finite number too. Then, for $t = t_c$ we obtain:

$$\begin{aligned} v(x(t_c)) = \phi_\delta(x_1(t_c) - x_3(t_c), x_1(t_c) - x_4(t_c)) \\ +\phi_\delta(x_2(t_c) - x_3(t_c), x_2(t_c) - x_4(t_c)) \le R \ , \end{aligned} \tag{18}$$

which implies

$$\begin{aligned} \phi_\delta(x_1(t_c) - x_3(t_c), x_1(t_c) - x_4(t_c)) \le R \\ \phi_\delta(x_2(t_c) - x_3(t_c), x_2(t_c) - x_4(t_c)) \le R \end{aligned} \tag{19}$$

since $\phi_\delta(\cdot,\cdot)$ is a nonnegative function. Using inequality (6) we obtain:

$$\begin{aligned} \min\{\|x_1(t_c) - x_3(t_c)\|, \|x_1(t_c) - x_4(t_c)\|\} \le R \\ \min\{\|x_2(t_c) - x_3(t_c)\|, \|x_2(t_c) - x_4(t_c)\|\} \le R \ . \end{aligned} \tag{20}$$

Finally, notice that inequalities in (20) mean capture of the first and the second evader at the finite time instant t_c, respectively. Q.E.D. □

In order to define strategies for the evaders, we consider the following function:

$$\varphi_\delta(y_1, y_2) = \frac{\|y_1\|\|y_2\|}{\sqrt[\delta]{\|y_1\|^\delta + \|y_2\|^\delta}}, \ \ \delta > 0 \ , \tag{21}$$

where δ is a given positive constant. Again, y_1 and y_2 are assumed to be real vectors of arbitrary dimensions such that $y = [y_1^T, y_2^T]^T$ is a nonzero vector. Then, the following limit property:

$$\lim_{\delta \to \infty} \sqrt[\delta]{1 + c^\delta} = 1 \ \text{ if } \ c \in [0, 1] \tag{22}$$

implies that

$$\lim_{\delta \to \infty} \varphi_\delta(y_1, y_2) = \min\left\{\|y_1\|, \|y_2\|\right\} \ . \tag{23}$$

Additionally, it is trivial to show that the function defined in (21) is a lower approximation of the min function, that is:

$$\varphi_\delta(y_1, y_2) \le \min\{\|y_1\|, \|y_2\|\}, \ (\forall\delta)(\delta > 0) \ . \tag{24}$$

Now, from Eqs. (23) and (24) it follows that the min function can be approximated arbitrarily close "from below" by the function $\varphi_\delta(\cdot,\cdot)$ for an appropriate choice of δ. Again, let us assume that vectors y_1 and y_2 are nonzero vectors and let $a = \|y_1\|$ and $b = \|y_2\|$ be corresponding positive numbers. Then, similar to Eq. (8), we define the following function:

$$\chi(\delta) = \frac{ab}{\sqrt[\delta]{a^\delta + b^\delta}}, \ \ \delta > 0 \ . \tag{25}$$

At this point we claim that the function $\chi(\delta)$ is an increasing function of the variable δ for $\delta > 0$. This can be easily proved by using the following inequality:

$$\sqrt[\delta_1]{a^{\delta_1} + b^{\delta_1}} > \sqrt[\delta_2]{a^{\delta_2} + b^{\delta_2}}, \text{ for } a > 0, b > 0, \delta_2 > \delta_1 > 0\,, \tag{26}$$

that is stated as Problem 1289(d) on p. 128 in [5] and a solution is provided in [12] on p. 281 as a solution to Problem 187(d). Similar to the analysis provided for the function $\phi_\delta(\cdot,\cdot)$ defined in (5), we have shown that $\chi(\delta) = \varphi_\delta(\|y_1\|, \|y_2\|)$ is a strictly increasing function with respect to the positive values of δ when y_1 and y_2 are nonzero vectors. Notice that the additional condition $\|y_1\| \neq \|y_2\|$ is not needed here.

In order to define guaranteed strategies for the escape of the evaders, we propose the following functions:

$$v_i(x) = \varphi_\delta(x_i - x_3, x_i - x_4),\ i \in \{1,2\}\,. \tag{27}$$

Corresponding to functions $v_i(x)$, $i \in \{1,2\}$ defined in (27) we compute the strategies:

$$\begin{aligned} u_i^o\,(x) &= \operatorname*{arg\,max}_{\|u_i\|\le\mu_i}\left\{\tfrac{dv_i}{dt}\right\} = \operatorname*{arg\,max}_{\|u_i\|\le\mu_i}\left\{\tfrac{\partial v_i}{\partial x_i} g_i(x_i)u_i\right\} \\ &= \mu_i g_i^T(x_i)\tfrac{\partial v_i}{\partial x_i}^T \left\|\tfrac{\partial v_i}{\partial x_i} g_i(x_i)\right\|^{-1},\ i \in \{1,2\} \\ u_j^o\,(x) &= \operatorname*{arg\,min}_{\|u_j\|\le\mu_j}\left\{\tfrac{dv_i}{dt}\right\} = \operatorname*{arg\,min}_{\|u_j\|\le\mu_j}\left\{\tfrac{\partial v_i}{\partial x_j} g_j(x_j)u_j\right\} \\ &= -\mu_j g_j^T(x_j)\tfrac{\partial v_i}{\partial x_j}^T \left\|\tfrac{\partial v_i}{\partial x_j} g_j(x_j)\right\|^{-1},\ j \in \{3,4\},\ i \in \{1,2\}\,. \end{aligned} \tag{28}$$

Then, similar to Theorem 2.1, we formulate the sufficient conditions for the guaranteed evasion of the i-th evader ($i \in \{1,2\}$) as the following.

Theorem 2.2. *Assume that the initial condition x_0, at the initial time t_0, is such that the players are outside of the region defined by the set $\{x : v_i(x) \geq R\}$ and*

$$\frac{dv_i}{dt} \geq 0,\ \ \forall t \geq t_0 \tag{29}$$

along the trajectories of the agents' dynamic systems (10) *for the input strategies $u_j = u_j^o(x)$, $j \in \{1,2,3,4\}$ defined in* (28) *with respect to the function v_i in* (27). *If the i-th evader chooses its strategy to be $u_i = u_i^o(x)$ as defined in* (28), *then it will avoid being captured by the pursuers for any choice of their feedback strategies that satisfy Eq.* (11).

Proof. If we choose $u_i = u_i^o(x)$ defined in (28) for both evaders and any feedback strategies $u_i(x)$, $i \in \{3,4\}$ for the pursuers, from Eqs. (10), (11), (13), (28), and (29) it follows that:

$$\frac{dv_i}{dt}(u_1^o(x), u_2^o(x), u_3(x), u_4(x)) \geq \frac{dv_i}{dt}(u_1^o(x), u_2^o(x), u_3^o(x), u_4^o(x)) \geq 0 \,. \tag{30}$$

By integrating Eq. (29) between t_0 and t we obtain:

$$v_i(x(t)) \geq v_i(x_0), \;\; \forall t \geq t_0 \,. \tag{31}$$

From Eq. (27) and inequalities (24) and (31) we obtain:

$$\min\{\|x_i(t) - x_3(t)\|, \|x_i(t) - x_4(t)\|\} \geq v_i(x(t)) \geq v_i(x_0) \geq R, \;\; \forall t \geq t_0 \,. \tag{32}$$

Q.E.D. □

Notice that one of the differences between Theorems 2.1 and 2.2 is in the specification of the initial conditions. In Theorem 2.1, the set of initial conditions is exactly defined in terms of the min function, where in Theorem 2.2 it is given as an approximation in terms of functions $\varphi_\delta(\cdot,\cdot)$ due to Eq. (27). For example, for $\delta = 1$ from Eqs. (5) and (21) it follows that $\varphi_1(y_1, y_2) = 0.5\phi_1(y_1, y_2)$ for any y_1 and y_2 such that $y = [y_1^T, y_2^T]^T$ is a nonzero vector. Then, assuming the players start the game outside of the capture region defined by a positive number ρ, we would have the following sequence of inequalities:

$$\begin{aligned} \rho \;&\leq \min\{\|x_i(t_0) - x_3(t_0)\|, \|x_i(t_0) - x_4(t_0)\|\} \\ &\leq \phi_1(x_i(t_0) - x_3(t_0), x_i(t_0) - x_4(t_0)) \\ &= 2\varphi_1(x_i(t_0) - x_3(t_0), x_i(t_0) - x_4(t_0)) \\ &= 2v_i(x_0), \;\; i \in \{1, 2\} \end{aligned} \tag{33}$$

implying $v_i(x_0) \geq \frac{\rho}{2}$. Thus, to satisfy inequality (32) one would have to set $\rho = 2R$. By increasing δ one may show that ρ would converge to R (R represents the exact capture region) which means that the specification of the initial condition region would be influenced by the choice of δ in the case of the guaranteed strategies for the evaders.

At the end of this section it is important to note that the results of Theorem 2.1 can be easily extended to the case of a pursuit-evasion game with arbitrary number of players as long as the number of pursuers is larger or equal to the number of evaders by following the methodology proposed in [23]. In order to apply the results of Theorem 2.1 to the case when the number of evaders is larger than the number of pursuers, one would need additional conditions that might be of discontinuous nature and this topic will be considered in our future research. Since the results of Theorem 2.2 are formulated for each evader independently, it is a straightforward consequence that these conditions can be extended to include the case of a pursuit-evasion game with an arbitrary number of pursuers and an arbitrary number of evaders.

3 The game with discrete observations

In this section, following the approach of [3], we assume that a fixed set of observation time-instants is given:

$$t_0 \leq t_1 \leq t_2 \leq \cdots \leq t_i \leq \dots \; . \tag{34}$$

Assume that the pursuers can observe the evader's state vectors only at certain time instants t_i, i.e., they know vectors $x_1(t_i), x_2(t_i)$ for $i = 0, 1, 2, \dots$, but they have continuous measurement of their own states $x_3(t), x_4(t)$. Thus, the control strategies of the pursuers can use only the values of $x_1(t_i), x_2(t_i)$ and $x_3(t), x_4(t)$. Without the loss of generality it is sufficient to assume that the pursuers also know the vectors $x_3(t_i), x_4(t_i)$ only. Thus, the most flexible feedback control of the pursuers is of the following type:

$$\begin{aligned} &u_j(x(t_i), t), \quad t_i \leq t \leq t_{i+1}, \quad i = 0, 1, \dots, \quad j = 3, 4 \\ &x(t_i) = (x_1(t_i), x_2(t_i), x_3(t_i), x_4(t_i)) \; . \end{aligned} \tag{35}$$

Such a control is called piecewise open-loop control. At time instant t_i the pursuers observe the vector $x(t_i)$, substitute it into the function (35), and apply it on the interval $[t_i, t_{i+1}]$. The resulting piecewise open-loop control is:

$$u_j(t) = u_j(x(t_i), t), \quad t_i \leq t \leq t_{i+1}, \quad i = 0, 1, \dots \; . \tag{36}$$

These piecewise open-loop control strategies are meaningful if some knowledge of $x(t_{i+1}) = (x_1(t_{i+1}), x_2(t_{i+1}), x_3(t_{i+1}), x_4(t_{i+1}))$ is available. Based on the knowledge of $x(t_i) = (x_1(t_i), x_2(t_i), x_3(t_i), x_4(t_i))$, it is straightforward to see that:

$$x_j(t_{i+1}) \in D_j(x_j(t_i)), \quad j = 1, 2, 3, 4 \; ,$$

where $D_j(x_j(t_i))$ is the j-th player's reachable set at the time t_{i+1}, when starting the motion at the instant t_i from the position $x_j(t_i)$. Introducing the so-called jump vectors p_{ji}, one can rewrite this in the form

$$\begin{aligned} &x_j(t_{i+1}) = x_j(t_i) + p_{ji}, \quad p_{ji} \in \bar{D}_j(x_j(t_i)) = D_j(x_j(t_i)) - x_j(t_i) \\ &\quad j = 1, 2, 3, 4 \; . \end{aligned} \tag{37}$$

Here the set $D_j(x_j(t_i)) - x_j(t_i)$ is the set of all vectors y such that $y + x_j(t_i) \in D_j(x_j(t_i))$. The above relations can be treated as a discrete analog of the dynamic Eqs. (10).

The vectors p_{ji} of the set $\bar{D}_j(x_j(t_i))$ have the following representation. Integrating both sides of (10) on the segment $[t_i, t_{i+1}]$ one obtains:

$$x_j(t_{i+1}) = x_j(t_i) + \int_{t_i}^{t_{i+1}} h_j(x_j(t))dt + \int_{t_i}^{t_{i+1}} g_j(x_j(t))u_j(t)dt \; .$$

Comparing with (37) one can see that the jump vector can be written in the form:

$$p_{ji} = \int_{t_i}^{t_{i+1}} h_j(x_j(t))dt + \int_{t_i}^{t_{i+1}} g_j(x_j(t))u_j(t)dt \,.$$

Note that the open-loop control $u(t), t_i \leq t \leq t_{i+1}$, generating a given p_{ji} may be nonunique: the set $\bar{D}_j(x_j(t_i))$ consists of all the vectors permitting such representation.

The system (37) tracks the state vectors for discrete instances only. For an intermediate instant t, $t_i < t < t_{i+1}$, the pursuers have the precise information about their own positions $x_3(t), x_4(t)$ and an estimation for $x_1(t), x_2(t)$ based on the last observation $x_1(t_i), x_2(t_i)$.

Based on the dynamics (37) and using the vectors $x = (x_1, x_2, x_3, x_4)$, $p = (p_1, p_2, p_3, p_4)$ one can suggest the following discrete version of the conditions (13), (15):

$$\min_{\{p_3,p_4\}} \max_{\{p_1,p_2\}} v(x+p) - v(x) \leq -\gamma_d, \quad p_j \in \bar{D}_j(x_j), \quad j = 1,2,3,4 \,. \tag{38}$$

Note that the parameter γ in (15) has the dimension of velocity, while the dimension of γ_d here is the distance.

A statement similar to Theorem 2.1, giving a sufficient capture condition, can be formulated for the case of discrete observations.

Theorem 3.1. *Assume that the initial condition x_0 at the initial time t_0 is such that the players are outside of the capture region defined by a positive number R. Then the capture of both evaders is guaranteed in finite time if the pursuers choose the piecewise open-loop strategies of the form* (35) *based on the jump vectors p_j defined in* (38)*.*

The proof of Theorem 3.1 is the finite-difference analog of the proof of Theorem 2.1. Since each step reduces the value of the function $v(x)$ by γ_d, the number of steps required for the capture is defined as the smallest integer N satisfying the inequality:

$$N\gamma_d \geq v(x_0) - R \,.$$

Now the estimation for the capture time is: $t_c \leq t_N$.

The pursuers can choose to minimize the number of observations. In (38), the instant t_{i+1} is considered as a fixed parameter. Now we assume that it can vary, and will search for the instant t^*_{i+1} farthest from t_i, i.e., giving the maximal interval $\Delta_i = t_{i+1} - t_i$, such that the condition (38) still holds:

$$\max_{t_{i+1}} \min_{\{p_3,p_4\}} \max_{\{p_1,p_2\}} v(x+p) - v(x) \leq -\gamma_d, \quad p_j \in \bar{D}_j(x_j), \quad j = 1,2,3,4 \,. \tag{39}$$

Note that the domain $\bar{D}_j(x_j)$ here depends upon t_{i+1}, and the value of Δ_i depends upon x, $\Delta_i = \Delta_i(x)$. Generally, maximizing the time t^*_{i+1} leads to the equality in (39).

Successive observation moments can be identified on the following basis:

$$t_{i+1} = t_i + \Delta_i(x(t_i)), \quad i = 0, 1, \dots . \tag{40}$$

Generally, this may reduce the number of observations N required for the capture.

Under certain conditions the number of observations required for evaders may be only one observation at the initial time t_0. We formulate a sufficient condition for this.

Theorem 3.2. *Assume that the initial condition x_0 at the initial time t_0 is such that the players are outside of the region defined by the set $\{x : v_i(x) \geq R\}$ and there exist open-loop controls for evaders $u_1(t)$, $u_2(t)$, $t \geq t_0$, such that*

$$\frac{dv}{dt}(x, u_1(t), u_2(t), u_3, u_4) \geq 0 \tag{41}$$

for all values of u_3, u_4 and all x along the trajectories and for all $t \geq t_0$. Then both evaders avoid being captured by the pursuers.

The proof uses the estimations obtained in Theorem 2.2. In more complicated situations, further observation points t_i could be required. Note that using an open-loop control, as in Theorem 3.2, means that only one observation at the initial time t_0 is made.

Conclusions

In this paper, a number of sufficient conditions that guarantee either capture or evasion of two evaders in a four-player pursuit-evasion game is provided. The players' dynamics are represented by nonlinear models affine in control. Conditions are formulated using a Liapunov type of analysis by considering functions that are convergent approximations of the min function. Guaranteed strategies are provided in both cases when either continuous or discrete observations are available to the players.

References

[1] M. Bardi and I. Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, Birkhäuser, Boston, 1997.

[2] J. V. Breakwell and P. Hagedorn, *Point capture of two evaders in succession*, Journal of Optimization Theory and Applications **27** (1979), 89–97.

[3] F. L. Chernousko and A. A. Melikyan, *Some differential games with incomplete information*, Lecture Notes in Computer Science, vol. 27, Springer-Verlag, Berlin, 1975, pp. 445–450.

[4] M. G. Crandall and P.-L. Lions, *Viscosity solutions of Hamilton-Jacobi equations*, Transactions of American Mathematical Society **277** (1983), 1–42.

[5] B. P. Demidovich, *Sbornik zadach i uprazhnenii po matematicheskomu analizu*, Nauka, Moscow, Russia, 1977 (in Russian).

[6] P. Hagedorn and J. V. Breakwell, *A differential game with two pursuers and one evader*, Journal of Optimization Theory and Applications **18** (1976), 15–29.

[7] I. Hwang, D. M. Stipanović, and C. J. Tomlin, *Polytopic approximations of reachable sets applied to linear dynamic games and to a class of nonlinear systems*, Advances in Control, Communication Networks, and Transportation Systems: In Honor of Pravin Varaiya (E.H. Abed, ed.), Systems & Control: Foundations & Applications, Birkhäuser, Boston, 2005, pp. 1–20.

[8] R. Isaacs, *Differential games: A mathematical theory with applications to warfare and pursuit, control and optimization*, John Wiley and Sons, Inc., New York, 1965.

[9] H. K. Khalil, *Nonlinear systems*, third ed., Prentice Hall, Upper Saddle River, NJ, 2002.

[10] N. N. Krasovskii and A. I. Subbotin, *Game-theoretical control problems*, Springer-Verlag, New York, 1988.

[11] A. Y. Levchenkov and A. G. Pashkov, *Differential game of optimal approach of two inertial pursuers to a noninertial evader*, Journal of Optimization Theory and Applications **65** (1990), 501–518.

[12] I. I. Liashko, *Matematicheskii analiz*, vol. 1, Vyshcha shkola, Kiev, Ukraine, 1983 (in Russian).

[13] A. A. Melikyan, *On minimal observations in a game of encounter*, PMM **37** (1973), no. 3, 407–414 (in Russian).

[14] A. A. Melikyan, *Optimal interaction of two pursuers in a game problem*, Tekhnicheskaya kibernetika (1981), no. 2, 49–56 (in Russian).

[15] A. A. Melikyan, *Generalized characteristics of first order PDEs*, Birkhäuser, Boston, 1998.

[16] A. A. Melikyan, N. H. Hovakimyan, and L. Harutyunyan, *Games of simple pursuit and approach on two-dimensional cone*, Journal of Optimization, Theory and Applications **98** (1998), no. 3, 515–543.

[17] A. A. Melikyan and A. Pourtallier, *Games with several pursuers and one evader with discrete observations*, Game Theory and Applications (L. A. Petrosjan and V. V. Mazalov, eds.), vol. 2, Nova Science Publishers, New York, 1996, pp. 169–184.

[18] A. W. Merz, *The game of two identical cars*, Journal of Optimization Theory and Applications **9** (1972), 324–343.

[19] I. Mitchell, A. M. Bayen, and C. J. Tomlin, *A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games*, IEEE Transactions on Automatic Control **50** (2005), 947–957.

[20] G. J. Olsder and O. Pourtallier, *Optimal selection of observation times in a costly information game*, New Trends in Dynamic Games and Applications (G. J. Olsder, ed.), Annals of the International Society of Dynamic Games, vol. 3, Birkhäuser, Boston, 1995, pp. 227–246.

[21] A. G. Pashkov and S. D. Terekhov, *A differential game of approach with two pursuers and one evader*, Journal of Optimization Theory and Applications **55** (1987), 303–311.

[22] L. A. Petrosjan, *Differential games of pursuit*, Series on Optimization, vol. 2, World Scientific, Singapore, 1993.

[23] D. M. Stipanović, S. Shankaran, and C. J. Tomlin, *Strategies for agents in multi-player pursuit-evasion games*, Proceedings of the Eleventh International Symposium on Dynamic Games and Applications (Tucson, Arizona), 2004.

A Numerical Approach to the 'Princess and Monster' Game on an Interval

Steve Alpern*
London School of Economics
Houghton Street, London WC2A 2AE, UK
alpern@lse.ac.uk

Robbert Fokkink
Delft Institute of Applied Mathematics
Delft University
P.O. Box 5031, 2600 GA Delft The Netherlands
r.j.fokkink@tudelft.nl

Roy Lindelauf
Faculty of Economics and Business Administration
Tilburg University
P.O. Box 90153, 5000 LE Tilburg The Netherlands
r.lindelauf@uvt.nl

Geert Jan Olsder
Delft Institute of Applied Mathematics
Delft University
P.O. Box 5031, 2600 GA Delft, The Netherlands
g.j.olsder@tudelft.nl

Abstract

Rufus Isaacs introduced Princess and Monster games in the final chapter of his classic book. The value of the Princess and Monster game on an interval is as of yet unknown. We present some numerical results to estimate this value.

Key words. Search games, linear programming.

AMS Subject Classifications. 91A05.

In the final chapter of his classic book *Differential Games*, Rufus Isaacs introduced the 'Princess and Monster' games. A Monster and a Princess may move about in a restricted space, more specifically in a network, and the Monster tries to catch the Princess. They are not able to see each other and that is why this type of game is known as a Search Game [3,7,9]. It is different from the more famil-

*Prof. Alpern is currently on visit at Delft University.

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_9,

iar Game of Pursuit [5], in which both players have visual contact. In most of the Search Games that have been solved so far, the Princess is immobile; e.g., [6,12]. The only Search Game with a mobile Princess on a graph that has been solved is the Princess and Monster game on a circle, and this was done a long time ago [1,13]. In a complementary paper [4] we have shown that the Princess and Monster game on an interval $[-1, 1]$ is not trivial (not trivial in the sense that for the Monster it is not optimal to start at one random end and then go as fast as possible to the other; if the Monster would move in this way, then we say that he is a *sweeper*, as in [10]) and that the value[1] $\mathcal{V}$ of the game exists and is bounded by $15/11 < \mathcal{V} < 13/9$. These bounds were obtained by analytical considerations and computations that can be checked by hand. In this paper, we consider a restricted game that has a value $\mathcal{V}_r \leq \mathcal{V}$. By numerical simulations we show that $\mathcal{V}_r \approx 1.373$. Here, we concentrate on a numerical approach of the problem. Analytical results can be found in [4].

1 Rules of the game

The rules of the game are as follows. The Monster $\mathcal{M}$ and the Princess $\mathcal{P}$ may choose an arbitrary initial point on the closed interval $[-1, 1]$. The Monster moves at speed bounded by 1, so the trajectory of $\mathcal{M}$, $\mathcal{M}(t)$ is a continuous function with Lipschitz constant 1. The Princess may move at arbitrary speed. In [4] we have shown that $\mathcal{M}$ always moves at maximum speed and that $\mathcal{P}$ need not move at speed greater than 1. If a player moves at speed 1 then we say that the player *runs*.

The minimax theorem in [2] implies that the value of the game $\mathcal{V}$ exists. The precise optimal strategies for the Monster and the Princess are not known, but we have derived some properties of optimal strategies in [4]. We have shown that it is optimal for the Princess to never cross the mid point and if she reaches an end, then she should stay there. If the Monster reaches an end, however, then he should turn and run to the opposite end.

2 A restricted search game on the interval

We have indicated in [4] that $\mathcal{V}$ can in principle be computed in an iterative manner, but the convergence is very slow and not numerically feasible. In this paper we take a different approach and restrict the number of pure strategies of the players.

For the Princess we allow the following pure strategies:

P The Princess chooses an infinitesimal small $\varepsilon > 0$. She either hides at an end point and remains immobile, or she chooses an arbitrarily initial point in the intervals $[-1+\varepsilon, -\varepsilon] \cup [\varepsilon, 1-\varepsilon]$ and remains immobile there until the

[1] To be precise, the value of the game is determined up to ε: given any $\varepsilon > 0$ there exists a mixed strategy for the Princess that guarantees her a payoff of $V - \varepsilon$. We ignore this ε in our notation.

sweeper coming from the nearest end is ε close. Then she runs to the middle and turns when she gets ε close, where she turns and runs back to the nearest end.

We note that an optimal strategy for the Princess is always optimal up to $\varepsilon > 0$, as remarked already in the footnote above. The smaller ε, the nearer the Princess reaches the value of the game.

For the Monster we allow pure strategies of two types (the strategies of $\mathcal{M}$ and $\mathcal{P}$ are indicated by M and P, respectively):

M_1 Choose an arbitrary initial point on $[-1, +1]$ and choose a direction, right or left. $\mathcal{M}$ runs in that direction until the end and then he runs back. In particular, if $\mathcal{M}$, as a sweeper, starts at an end then runs to the opposite end.

M_2 Choose an arbitrary initial point on $[-1, +1]$ and choose a direction. $\mathcal{M}$ runs in that direction until he meets the sweeper coming from the opposite end. Then $\mathcal{M}$ turns and joins the sweeper. Subsequently, he keeps running in this direction till the end (at either $+1$ or -1) has been reached and then returns to the opposite end.

We are unable to prove that the strategy P is optimal for the Princess, but we conjecture that it is. If it is, then the game is solved.

Theorem 2.1. *If the Princess uses a mixed strategy that consists of the pure strategies in P, then the optimal response of the Monster is to use a mixed strategy that consists of pure strategies in M_1 and M_2.*

It is possible to prove this theorem following the approach of Sec. 5 in [4]. We only give a sketch of the proof here. If the Princess uses these pure strategies, then she remains immobile and runs only if the sweeper is ε-close, or she hides at an end and remains immobile all the time. Therefore, at time $t > 0$, either the Princess is immobile and $\mathcal{P}(t) = \pm 1$ or $\mathcal{P}(t) \in (-1 + t + \varepsilon, 1 - t - \varepsilon)$, or she is running and $\mathcal{P}(t) = \pm(-1 + t + \varepsilon)$. The Monster knows this. If he starts at an end, then his only sensible strategy is to run to the other end. If he starts in the middle, then he should run in one direction to increase the chance of catching the Princess while she is still immobile. Then if he is ε-close to a sweeper, the Monster may turn and then it is sensible only to run to the other side, or he may continue his run to the end, turn there and run to the other side. If we ignore ε, which we may in the limit, then we get the pure strategies given above for the Monster.

The strategies in P are of two types. The Princess either hides at an end and stays put, let's call this strategy E, or she uses a strategy I of hiding at an internal point and moving close to the central point at time $1 - \varepsilon$, staying away from the sweepers for as long as possible. If the Princess only plays E then the Monster only plays the sweeper strategy and the value of the game is 1. It is better for the Princess to mix E and I, which she does in the strategy P. In this way, the Monster is forced to use pure strategies that get to the central point before the sweeper does.

The value of this version of the search game, with the given restricted classes of strategies, is indicated by $\mathcal{V}_r$. Because of Thm. 2.1 it satisfies $\mathcal{V}_r \leq \mathcal{V}$.

3 Approximation of $\mathcal{V}_r$ by Discretization

We analyze the restricted game by numerical simulations. As a first very rough approximation of $\mathcal{V}$, discretize the interval $[-1, 1]$ and take two grid points only -1 and $+1$: the mesh of this simple grid is $\Delta x = 2$. Discretize time accordingly into time steps of $\Delta t = 2$. It is not hard to see that it is optimal for both players to choose a grid point at random. The Monster moves to the other grid point, the Princess remains where she is. The value of this simple discretized game is 1.

As a second approximation discretize by three grid points $-1, 0, 1$. Discretize time accordingly into time-steps of $\Delta t = 1$ allowing the Princess an ε-advantage: the Princess may move on time $n-\varepsilon$ while the Monster moves on time n for $n \in \mathbb{N}$. There is an obvious symmetry in the game: each player choses -1 and $+1$ equally likely. Let's call this the end point strategy E as opposed to the mid point strategy C (of "Center") in which the initial point is 0. If the Monster chooses an end point, then he runs the opposite end as quickly as possible. If the Princess chooses an end point, then she stays there. If the Monster chooses the mid point, then he runs to a random end and runs back. If the Princess chooses the mid point, she runs to a random side at time $1 - \epsilon$. So we get a 2×2 matrix game (consistently ignoring ε):

	E	C
E	1	1.5
C	2	0

in which the Monster chooses a row and the Princess a column. The value of this game is slightly larger: $6/5$.

We discretize the game. The Princess and the Monster may only choose an initial position on a fixed equidistant grid with $2n$ points, both endpoints included in this counting ($n = 1, 2, \ldots$). Hence, the mesh size Δ equals $2/(2n-1)$. The case with $n = 1$ coincides with the first approximation given above. The Princess moves at the time steps $\Delta - \varepsilon, 2\Delta - \varepsilon, \ldots$ and the Monster moves at time steps $0, \Delta, 2\Delta, \ldots$. The game is over as soon as they occupy the same grid point. In particular, if the Princess and the Monster choose the same initial grid point, then the game is immediately over and the capture time is 0.

Remark *It is obvious that the probability for both players choosing the same initial point is positive for the finite grid case. Because of the assumption that in such a case capture is immediate, this is to the disadvantage for the Princess and this suggests that the value of this game is a lower bound for $\mathcal{V}_r$ and with increasing n it will converge from below to $\mathcal{V}_r$.*

We discretize the interval by putting a symmetric grid with respect to the mid point. There are n equidistant grid points smaller than 0 and there are n equidistant grid points greater than 0 at:

$$\{-1, -1+\Delta, \ldots, -1+(n-1)\Delta\} \cup \{1, 1-\Delta, \ldots, 1-(n-1)\Delta\}.$$

We denote the value of the discretized game by $\mathcal{V}_n$. Our analysis of the game in [4] can be used to demonstrate that $\mathcal{V}_n \to \mathcal{V}_r$ as $n \to \infty$. We only sketch the argument here.

It is not difficult to show that, now that the hiding place of the Princess is limited to a grid, the Monster starts his search from a grid point. One of our results is that the Monster runs all the time, so at the discrete time steps the Monster is at one of the grid points. Hence, the discretization does not limit the Monster, which implies that $\mathcal{V}_n \leq \mathcal{V}_r$. On the other hand, any continuous strategy of the Princess from P can be shadowed on the grid by a discrete strategy that is within a mesh distance. The continuity properties of the payoff then imply that $\mathcal{V}_n \uparrow \mathcal{V}_r$ for an increasing number n of grid points.

Table 1: Numerical approximation of $\mathcal{V}_r$

n	$\mathcal{V}_n$
1	1
2	1.266
4	1.330
8	1.354
16	1.365
32	1.370
64	1.373

The matrix in the discretized game has size $8n \times 2n$ and we are unable to compute $\mathcal{V}_n$ for larger n. Our results seem to suggest that the limit value is $\mathcal{V}_r = 11/8$.

The results of our simulations show that the Princess uses three types of strategies: (a) she hides at either end point, with a positive probability ≈ 0.127 at each end point; (b) she hides at $\pm\varepsilon$ until time $1-\varepsilon$ and runs to a random end with total probability ≈ 0.236 evenly divided over the two points; (c) she takes an initial position in $(-1, 1)$ according to a continuous probability distribution, as depicted in Fig. 1.

The Monster also uses two discrete strategies and one continuous strategy. He uses the sweeper strategy with probability ≈ 0.80, evenly divided between the two sweep options. He starts at $\pm(-1+\varepsilon)$ and runs to the opposite end and back with probability ≈ 0.075 for each option, or he picks an initial position in $(-1, 1)$ according to a continuous probability distribution, as depicted in Fig. 2. If the initial position < 0, then he runs to the right end and subsequently back; if the initial position is > 0, then he runs first to the left end and then back. Note that the Monster only uses the pure strategies of M_1 and he only uses half of these: if

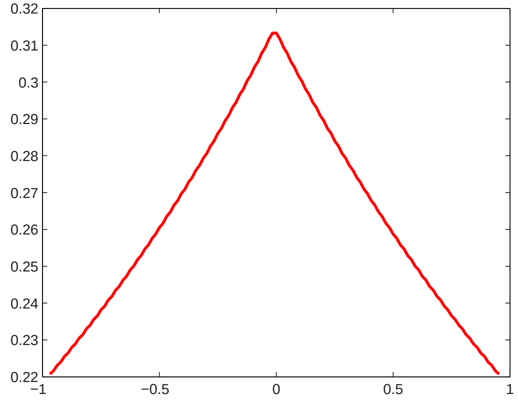

Figure 1: Probability density of the continuous Princess strategy.

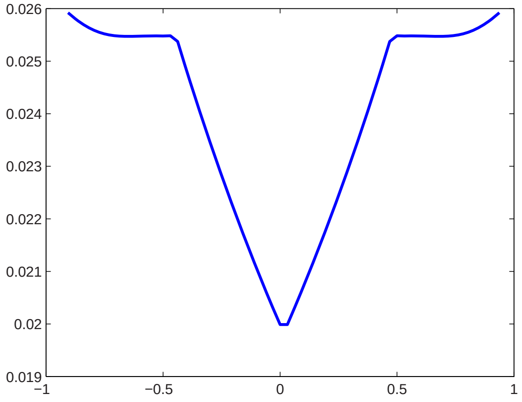

Figure 2: Probability density of the continuous Monster strategy.

$\mathcal{M}(0) < 0$ then he runs to $+1$ and back to -1; if $\mathcal{M}(0) > 0$ then he runs to -1 and back to $+1$

The probability density in Fig. 1 suggests a piecewise linear function, but there is no obvious analytical function for the probability density in Fig. 2. Finding an analytical solution of the restricted game appears to be a difficult problem.

4 Approximation by means of Taylor expansions

We now take a different turn and try to approximate the analytical solution of the restricted game by developing Taylor expansions. The starting assumptions here are that both players use a continuous initial distribution, with delta functions (concentrated masses) at both ends. For the Monster we assume a continuous initial density distribution M_l on $(-1,+1)$ for immediate left turns after the start and a continuous initial density distribution M_r on $(-1,+1)$ for immediate right turns after the start. Because of symmetry, $M_l(x) = M_r(-x)$. The remaining mass, $1 - \int_1^{+1}(M_l(x) + M_r(-x))dx$, is split into two equal parts which will refer to the mass concentrations at both ends. The functions P_l and P_r are likewise defined for the Princess, the only difference being that the Princess does not immediately after the start turn left or right, but first does not move and then, at moment $\min\{1+x, 1-x\} - \epsilon$, with x being the initial position of $\mathcal{P}$, turns left or right. This ϵ-value is arbitrarily small (and positive) and indicates the fact that the Princess is followed on her heels by the Monster if both run in the same direction. Obvious constraints are that all these densities are nonnegative.

Assume the functions P_l and M_l are linear, i.e.,

$$P_l(x) = c + dx,\ M_l(x) = a + bx,$$

on the interval $[-1,+1]$ and where the constants a, b must be chosen by the Monster, and c, d must be chosen by the Princess, subject to the constraints already mentioned. Thus we face the game

$$\min_{a,b}\max_{c,d} T,$$

where T denotes the time of capture. Expressing T in terms of the variables a, b, c, d and subsequently solving the minmax problem (using Maple) we find

$$a = \frac{654}{4327},\ b = \frac{624}{4327},\ c = \frac{2240}{4327},\ d = \frac{140}{4327},$$

and the value of the game is 1.345.

Please note that the admissible class of strategies for the Monster does not include the optimal strategies as obtained in the discretized game (where, for instance, $\mathcal{M}$, starting from a point in $(-1,0)$, would only move to the right and not to the left). Note also, that the class of admissible strategies for the Princess are different from the one before (after an initial rest, she simply runs to one of the two ends and stays there). In spite of the fact that the optimal strategies (as dealt with in the previous section) are outside the current admissible classes used, it is surprising to see how close the numerically obtained values are.

This method can easily be adapted to include different classes of strategies for the Monster and Princess, such as for instance to include the possibilities of the previous section. Moreover, the method can be extended to include second and even higher-order terms in the expansions of M_l and P_l.

5 Conclusion

In this paper, we present computer simulations of a search game with a mobile hider and we take some preliminary steps to develop Taylor expansions of the optimal mixed strategies. A solution of this game has eluded us so far.

We have given the Princess a mixed strategy P that contains pure strategies of two types: immobile hiding at an end point E or mobile hiding in an internal point I. Our numerical simulations show that in this restricted game the Princess plays E with probability around ≈ 0.25 and plays I with probability ≈ 0.75. The response of the Monster is to act as a sweeper S or to start out from an internal point and running up and down the interval, possibly turning upon meeting the sweeper. Let's call this J. The strategies I and J are mixed, but let's think of them as pure strategies for a moment so that we can set up a 2×2 zero-sum game matrix.

Our numerical simulations show that the Monster plays S with probability ≈ 0.8 and J with probability ≈ 0.2. The payoff of S against E and I is easy to compute. The payoff of J against E is not so easy to compute. Our simulations indicate that in the strategy J the Monster starts very near one end and that he runs to the other end and back, as indicated by the probability density in Fig. 2. The payoff of this strategy against E is very near 3, so let's estimate it at that. The payoff of J against I is unclear, but if the value of the game is $11/8$, which our numerical simulations suggest, then the matrix of the game should be as follows:

	E	I
S	1	1.5
J	3	5/6

In this matrix game, the Princess plays E with probability 0.25 and the Monster plays S with probability 0.8125, in agreement with our numerical results.

Our simulations clearly imply that strategy M_2 is irrelevant and only those pure strategies in M_1 are used in which the Monster first runs towards 0. It is probably not easy to give a solid proof of these facts, but it is possible to give some intuition for it and we would like to thank one of the anonymous referees for supplying this intuition to us. The strategies in M_1 and M_2 are designed against a Princess that hides in an internal point. Such a Princess moves from the internal point towards 0 and the Monster catches her if he moves to 0 as well. That is why only half of the strategies in M_1 are used. The Princess is near 0 at a time near 1, so the Monster should make sure that he is at 0 near time 1. In other words, the Monster preferably starts near an end point, as in the distribution of Fig. 2. Once the Monster has started out near an end point, he has an option of turning once he meets a sweeper from the opposite end, as in M_2. If he does, then it takes him longer to find a Princess that is hiding at an end point. Also, if this Monster meets the sweeper and has not caught the Princess, then he knows that the Princess is out of reach and will be found at an end point. Thus, it makes sense for the Monster to continue his way, as he does in strategy M_1.

Acknowledgements

This research was carried out while the first author visited Delft University on NWO visitor grant B61-590.

References

[1] S. Alpern (1974). *The search game with mobile hider on the circle*, Proceedings of the Conference on Differential Games and Control Theory, Kingston, Rhode Island, July 1973. In Differential Games and Control Theory (E. O. Roxin, P. T. Liu, and R.L Sternberg, eds), 181-200. M. Dekker, New York.

[2] S. Alpern and S. Gal. *A mixed strategy minimax theorem without compactness*, S.I.A.M. J. Control Optim. **33** (1998), no. 2 311-323, .

[3] S. Alpern and S. Gal. *The Theory of Search Games and Rendezvous*. Kluwer International Series in Operations Research and Management Sciences, 319 pp, Kluwer, Boston, 2003.

[4] S. Alpern, R. Fokkink, R. Lindelauf, G.J. Olsder (2006). *The Princess and Monster Game on an interval*, SIAM J. of Control Optim., to appear.

[5] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory* (second edition), SIAM, Philadelphia, 1999.

[6] A. Dagan and S. Gal, *Searching a network from an arbitrary starting point*, Networks, in press.

[7] S. Gal, *Search Games*. Academic Press, New York, 1980.

[8] S. Gal, *On the optimality of a simple strategy for searching graphs*, Int. J. Game Theory **29** (2000), 533-542.

[9] A. Y. Garnaev, *Search Games and Other Applications of Game Theory*. Springer-Verlag, Berlin, 2000.

[10] J. V. Howard, *Rendezvous search on the interval and the circle*, Operations Res. **47** (1999), 550-557.

[11] R. Isaacs, *Differential Games*. Wiley, New York, 1965.

[12] J. H. Reijnierse and J. A. M. Potters, *Search games with immobile hider*, Int. J. Game Theory **21** (1993), 385-394.

[13] M. I. Zeliken *On a differential game with incomplete information*, Soviet Math. Doklady **13** (1973), 228-231.

Numerical Approximation and Optimal Strategies for Differential Games with Lack of Information on One Side

Pierre Cardaliaguet
Laboratoire de Mathématiques,
Université de Bretagne Occidentale,
UMR 6205, 6 Av. Le Gorgeu,
BP 809, 29285 Brest
Pierre.Cardaliaguet@univ-brest.fr

Abstract

We investigate two-player zero-sum differential games in which the first player has more information on the terminal payoff than the second one. For this game we explain how to build numerical schemes for the value function and use these schemes to construct ϵ-optimal strategies for the players.

Key words. Differential games, lack of information, approximation, optimal strategies

AMS Subject Classifications. Primary 91A23; Secondary 49N70

1 Introduction

Since Aumann and Maschler´s seminal paper [1], repeated games with lack of information have attracted a lot of attention. In this zero-sum two-player model, one of the payers has a private information on some parameter of the repeated game while the other player has only some belief on this parameter.

In a series of papers [3], [4], we have explained how to adapt Aumann and Maschler theory to differential games with lack of information on one or two sides and have characterized this value in terms of “dual solution” of some Hamilton-Jacobi equation. In this paper we concentrate on differential games with lack of information on one side and address two issues: first the numerical approximation of the value; second the construction of ϵ-optimal strategies of the players. Both problems are strongly connected: indeed we use the numerical approximation of the value to build ϵ-optimal strategies.

Let us point out that the two problems cannot be deduced in a straightforward way from classical approaches. Indeed, the key point in the numerical approximation of classical differential games is that the value satisfies a dynamic programming principle (see [7] and the references therein, for instance). This is not the case here: all we have is that the value function and its convex conjugate

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_10,

$\mathbf{V}^*$ (see below) satisfy a subdynamic programming, which is strict in general for both functions. This is linked with the fact that the informed player has to use his information without revealing it too quickly, while the non informed player has to use the observation of his opponent's controls to learn something about the unknown parameter. Note that we have somehow to describe such behaviour in the construction of optimal strategies for the players.

In order to describe our results, we now briefly recall the framework and main results of [3]. The dynamics of the differential game is given by:

$$\begin{cases} x'(t) = f(x(t), u(t), v(t)) , & u(t) \in U, \; v(t) \in V, \\ x(t_0) = x_0 \end{cases} \tag{1}$$

where U and V are compact subsets of some finite dimensional spaces and $f : R^N \times U \times V \to R^N$ is Lipschitz continuous. The terminal time of the game is denoted by T. The game starts at time $t_0 \in [0, T]$ from the initial position x_0.

Let $g_i : R^N \to R$ be I terminal payoffs (where $I \geq 2$), $p = (p_i)_{i=1,\dots,I}$ belong to the set $\Delta(I)$ of probabilities on $\{1, \dots, I\}$. At the initial time t_0, the parameter i is chosen at random according to the probability p and is communicated to Player I only. Then the players control system (1) in order, for Player I, to minimize the terminal payoff $g_i(x(T))$, and for Player II to maximize it. Note that Player II does not really know which terminal payoff he is actually optimizing because he ignores which index i has been chosen. The key assumption is that both players observe their opponent's control. We also suppose that both players have a full knowledge of the dynamics f, of the payoffs g_i ($i \in \{1, \dots, I\}$) and of the probability p. So the unique difference in their information structure is that Player I knows the index i chosen according to probability p, while Player II does not. However Player II can try to guess his missing information by looking at his opponent's behavior.

In [3], we proved that this game—under Isaacs' condition recall below—has a value: namely the equality

$$\inf_{(\alpha_i) \in (\mathcal{A}_r(t_0))^I} \sup_{\beta \in \mathcal{B}_r(t_0)} \sum_{i=1}^{I} p_i \mathbf{E}_{\alpha_i \beta} \left(g_i \left(X_T^{t_0, x_0, \alpha_i, \beta} \right) \right) = \tag{2}$$

$$\sup_{\beta \in \mathcal{B}_r(t_0)} \inf_{(\alpha_i) \in (\mathcal{A}_r(t_0))^I} \sum_{i=1}^{I} p_i \mathbf{E}_{\alpha_i \beta} \left(g_i \left(X_T^{t_0, x_0, \alpha_i, \beta} \right) \right)$$

holds. In the above expressions, $\alpha_i \in \mathcal{A}_r(t_0)$ (for $i = 1, \dots, I$) are I random strategies for Player I, $\beta \in \mathcal{B}_r(t_0)$ is a random strategy for Player II and $\mathbf{E}_{\alpha_i \beta} \left(g_i \left(X_T^{t_0, x_0, \alpha_i, \beta} \right) \right)$ is the payoff associated with the pair of strategies (α_i, β) for the terminal payoff g_i: these notions are explained in the next section. Player I chooses his strategies $\hat{\alpha} = (\alpha_i)$ according to the value of the index i, while Player II, on the contrary, plays a strategy β independent of i. This reflects the asymmetry of information of the players. In (2) the sum $\sum_i p_i(\cdot)$ is just the expectation of the payoff when g_i is chosen according to the probability p. We denote by $\mathbf{V}(t_0, x_0, p)$ the value of the game given by (2).

The first result of this paper, described in details in Sec. 3, is about the approximation of the value $\mathbf{V}$. Let us fix a large integer L and let us denote by $\tau = T/L$ the time-step of discretization. We set

$$t_k = k\tau \qquad \text{for} \quad k \in \{0, \ldots, L\}$$

and we define by backward induction on k the function $\mathbf{V}_\tau(t_k, \cdot, \cdot) : R^N \times \Delta(I) \to R$:

$$\text{for } k = L, \qquad \mathbf{V}_\tau(T, x, p) = \sum_i p_i g_i(x)$$

and (assuming $\mathbf{V}_\tau(t_{k+1}, \cdot, \cdot)$ is built)

$$\mathbf{V}_\tau(t_k, x, p) = \mathrm{Vex}_p \left(\min_{u \in U} \max_{v \in V} \mathbf{V}_\tau(t_{k+1}, x + \tau f(x, u, v), p) \right), \tag{3}$$

where $\mathrm{Vex}_p(\cdot)$ denote the convex hull with respect to p. Then we prove that the function $\mathbf{V}_\tau$ uniformly converges to $\mathbf{V}$ as $\tau \to 0^+$. This time discretization of the value is surprizingly close to the discretization of the value of the classical (i.e., with symmetric information) zero-sum differential game with terminal payoff $\sum_i p_i g_i(x)$ (see [2]): the only difference is the fact that we have to take the convex hull with respect to p.

The second result of the paper is a construction of ϵ-optimal strategies for the players. For Player I it is based on the previous time discretization. Let us recall that, if we were in a classical differential games, a possible construction would be the following: suppose the state of the system has reached some $x \in R^N$ at some time t_k, Player I would just play on the time interval $(t_k, t_{k+1}]$ the constant control $u \in U$ which minimizes $\min_{u \in U} \max_{v \in V} \mathbf{V}_\tau(t_{k+1}, x + \tau f(x, u, v))$ (recall that in this case $\mathbf{V}_\tau$ is built by (3) without the convex hull).

In our game with lack of information on one side, Player's I strategy is much more involved. The key point—which is completely new in the framework of differential games—is that the construction of his strategy requires the joint construction of a random processes (p_t) living on $\Delta(I)$. An interpretation of this random process is the following: even if Player I knows which parameter i has been chosen by nature, he also knows that Player II ignores it. The random process p_t somehow represents the *a-posteriori* information he allows Player I to get on the parameter i.

Let us now describe how Player I builds his strategy: when the system has reached some state x at time t_k and when p_t equals some p, Player I first chooses some $\lambda_j \geq 0$, $p^j \in \Delta(I)$ and $u^j \in U$ in such a way that (recall (3))

$$\sum_j p_j = 1,\ \sum_j \lambda_j p^j = p\ ,$$

$$\mathbf{V}_\tau(t_k, x, p) = \sum_j \lambda_j \left(\min_{u\in U} \max_{v \in V} \mathbf{V}_\tau(t_{k+1}, x + \tau f(x,u,v), p^j) \right)\ ,$$

u^j being optimal in $\min_{u\in U}\max_{v\in V} \mathbf{V}_\tau(t_{k+1}, x + \tau f(x,u,v), p^j)$. Then he chooses randomly the control u^j with probability $\lambda^j p^j_i / p_i$ (where i is the parameter chosen by nature) and plays this constant control u^j on the time-interval $(t_k, t_{k+1}]$. If u^j has been chosen, he also updates the process p_t by setting $p_t = p^j$ on $(t_k, t_{k+1}]$. The probability $\lambda^j p^j_i/p_i$ is precisely the one under which the process (p_t) is a martingale for Player II. Such a construction is strongly related with the so-called splitting method in repeated game theory [1], [9]. It turns out that the strategy built in this way is ϵ-optimal for Player I as soon as the time-step τ of the discretization is sufficiently small. The rigorous construction and its optimality are explained in details in Sec. 4.

For Player II, ϵ-optimal strategies are built by using the approximation of the convex conjugate of the value function. Surprisingly, Player II also has to built a random process, but this time in the dual variable $\hat{p} \in R^I$. This is done in Sec. 5.

2 Main notations and assumptions

In this section we introduce the main notations and assumptions needed in the paper. We also recall the main results of [3]: existence of a value for the game for lack of information on one side, as well as the characterization of the value.

Notations: Throughout the paper, $x.y$ denotes the scalar product in the space R^N, R^I, or R^J (depending on the context) and $|\cdot|$ the Euclidean norm. The ball of center x and radius r will be denoted by $B_r(x)$. The set $\Delta(I)$ is the set of probabilities measures on $\{1,\ldots,I\}$, always identified with the simplex of R^I:

$$p = (p_1, \ldots, p_I) \in \Delta(I) \quad \Leftrightarrow \quad \sum_{i=1}^I p_i = 1 \text{ and } p_i \geq 0 \text{ for } i = 1, \ldots I\ .$$

The dynamics of the game is given by:

$$\begin{cases} x'(t) = f(x(t), u(t), v(t))\ , \qquad u(t) \in U,\ v(t) \in V\ . \\ x(t_0) = x_0. \end{cases} \tag{4}$$

Throughout the paper we assume that:

$$\left\{\begin{array}{rl} i) & U \text{ and } V \text{ are compact subsets of some finite dimensional spaces,} \\ ii) & f : R^N \times U \times V \to R^N \text{ is bounded, continuous, Lipschitz} \\ & \text{continuous with respect to the } x \text{ variable,} \\ iii) & \text{for } i = 1, \ldots, I,\ g_i : R^N \to R \text{ is Lipschitz continuous} \\ & \text{and bounded.} \end{array}\right. \tag{5}$$

We also assume that Isaacs condition holds and define the Hamiltonian of our primal HJ equation:

$$H(x, \xi) := \min_{u \in U} \max_{v \in V} f(x, u, v).\xi = \max_{v \in V} \min_{u \in U} f(x, u, v).\xi \tag{6}$$

for any $(x, \xi) \in R^N \times R^N$.

For any $t_0 < T$, the set of open-loop controls for Player I is defined by:

$$\mathcal{U}(t_0) = \{u : [t_0, T] \to U \text{ Lebesgue measurable}\} \ .$$

Open-loop controls for Player II are defined symmetrically and denoted by $\mathcal{V}(t_0)$. For any $(u, v) \in \mathcal{U}(t_0) \times \mathcal{V}(t_0)$ and any initial position $x_0 \in R^N$, we denote by $t \to X_t^{t_0, x_0, u, v}$ the solution to (4).

A *pure* strategy for Player I at time t_0 is a map $\alpha : \mathcal{V}(t_0) \to \mathcal{U}(t_0)$ which is nonanticipative with delay, i.e., there is a partition $t_0 < t_1 < \cdots < t_k = T$ such that, for any $v_1, v_2 \in \mathcal{V}(t_0)$, if $v_1 \equiv v_2$ a.e. on $[t_0, t_i]$ for some $i \in \{0, \ldots, k-1\}$, then $\alpha(v_1) \equiv \alpha(v_2)$ a.e. on $[t_0, t_{i+1}]$.

A random strategy for Player I is a pair $((\Omega_\alpha, \mathcal{F}_\alpha, \mathbf{P}_\alpha), \alpha)$, where $(\Omega_\alpha, \mathcal{F}_\alpha, \mathbf{P}_\alpha)$ is a probability space (chosen by Player I in the set $\{([0,1]^k, \mathcal{F}_k, \mathbf{P}_k),\ k \in N^*\}$ where $\mathcal{F}_k$ is the Borel σ-algebra on $[0,1]^k$ and $\mathbf{P}_k$ is the Lebesgue measure) and $\alpha : \Omega_\alpha \times \mathcal{V}(t_0) \to \mathcal{U}(t_0)$ satisfying:

(i) α is measurable from $\Omega_\alpha \times \mathcal{V}(t_0)$ to $\mathcal{U}(t_0)$, with Ω_α endowed with the $\sigma-$field $\mathcal{F}_\alpha$ and $\mathcal{U}(t_0)$ and $\mathcal{V}(t_0)$ with the Borel $\sigma-$field associated with the L^1 distance,

(ii) there is a partition $t_0 < t_1 < \cdots < t_k = T$ such that, for any $v_1, v_2 \in \mathcal{V}(t_0)$, if $v_1 \equiv v_2$ a.e. on $[t_0, t_i]$ for some $i \in \{0, \ldots, k-1\}$, then $\alpha(\omega, v_1) \equiv \alpha(\omega, v_2)$ a.e. on $[t_0, t_{i+1}]$ for any $\omega \in \Omega_\alpha$.

We denote by $\mathcal{A}(t_0)$ the set of pure strategies and by $\mathcal{A}_r(t_0)$ the set of random strategies for Player I. By abuse of notations, an element of $\mathcal{A}_r(t_0)$ is simply noted α—instead of $((\Omega_\alpha, \mathcal{F}_\alpha, \mathbf{P}_\alpha), \alpha)$—, the underlying probability space being always denoted by $(\Omega_\alpha, \mathcal{F}_\alpha, \mathbf{P}_\alpha)$.

In order to take into account the fact that Player I knows the index i of the terminal payoff, an admissible strategy for Player I is actually a $I-$uple $\hat{\alpha} = (\alpha_1, \ldots, \alpha_I) \in (\mathcal{A}_r(t_0))^I$.

Pure and random strategies for Player II are defined symmetrically; $\mathcal{B}(t_0)$ (resp. $\mathcal{B}_r(t_0)$) denotes the set of pure strategies (resp. random strategies). Generic elements of $\mathcal{B}_r(t_0)$ are denoted by β, with associated probability space $(\Omega_\beta, \mathcal{F}_\beta, \mathbf{P}_\beta)$.

Lemma 2.1 (Lemma 2.2 of [3]). *For any pair $(\alpha, \beta) \in \mathcal{A}_r(t_0) \times \mathcal{B}_r(t_0)$ and any $\omega := (\omega_1, \omega_2) \in \Omega_\alpha \times \Omega_\beta$, there is a unique pair $(u_\omega, v_\omega) \in \mathcal{U}(t_0) \times \mathcal{V}(t_0)$, such that*

$$\alpha(\omega_1, v_\omega) = u_\omega \text{ and } \beta(\omega_2, u_\omega) = v_\omega \; . \tag{7}$$

Furthermore, the map $\omega \to (u_\omega, v_\omega)$ is measurable from $\Omega_\alpha \times \Omega_\beta$ endowed with $\mathcal{F}_\alpha \otimes \mathcal{F}_\beta$ into $\mathcal{U}(t_0) \times \mathcal{V}(t_0)$ endowed with the Borel $\sigma-$field associated with the L^1 distance.

Remark 2.1. In [3] we have worked with a slightly different notion of nonanticipative strategy. It is not difficult to see that the above result as well as Theorem 2.1 remain true with the definition of strategy given here.

Notations: Given any pair $(\alpha, \beta) \in \mathcal{A}_r(t_0) \times \mathcal{B}_r(t_0)$, we denote by $(X_t^{t_0,x_0,\alpha,\beta})$ the map $(t, \omega) \to (X_t^{t_0,x_0,u_\omega,v_\omega})$ defined on $[t_0, T] \times \Omega_\alpha \times \Omega_\beta$, where (u_ω, v_ω) satisfies (7). The expectation $\mathbf{E}_{\alpha\beta}$ is the integral over $\Omega_\alpha \times \Omega_\beta$ against the probability measure $\mathbf{P}_\alpha \otimes \mathbf{P}_\beta$. In particular, if $\phi : R^N \to R$ is some bounded continuous map and $t \in (t_0, T]$, we have:

$$\mathbf{E}_{\alpha\beta}\left(\phi\left(X_t^{t_0,x_0,\alpha,\beta}\right)\right) := \int_{\Omega_\alpha \times \Omega_\beta} \phi\left(X_t^{t_0,x_0,u_\omega,v_\omega}\right) d\mathbf{P}_\alpha \otimes \mathbf{P}_\beta(\omega) \; , \tag{8}$$

where (u_ω, v_ω) is defined by (7).

For $p \in \Delta(I)$, $(t_0, x_0) \in [0, T) \times R^N$, the payoff associated with a strategy $\hat{\alpha} = (\alpha_i)_{i=1,\ldots,I} \in (\mathcal{A}_r(t_0))^I$ of Player I and a strategy $\beta \in \mathcal{B}_r(t_0)$ of Player II is defined by:

$$\mathcal{J}(t_0, x_0, \hat{\alpha}, \beta, p) = \sum_{i=1}^{I} p_i \mathbf{E}_{\alpha_i\beta}\left(g_i\left(X_T^{t_0,x_0,\alpha_i,\beta}\right)\right) \; , \tag{9}$$

where $\mathbf{E}_{\alpha_i\beta}$ is defined by (8).

Let us now recall the main result of [3].

Theorem 2.1 (Existence of the value [3]). *Assume that conditions (5) on f and on the g_i hold and that Isaacs assumption (6) is satisfied. Then the following equality holds:*

$$\inf_{\hat{\alpha} \in (\mathcal{A}_r(t_0))^I} \sup_{\beta \in \mathcal{B}_r(t_0)} \mathcal{J}(t_0, x_0, \hat{\alpha}, \beta, p) \; = \tag{10}$$

$$\sup_{\beta\in\mathcal{B}_r(t_0)}\ \inf_{\hat\alpha\in(\mathcal{A}_r(t_0))^I} \mathcal{J}(t_0,x_0,\hat\alpha,\beta,p)\ . \tag{11}$$

We denote by $\mathbf{V}(t_0,x_0,p)$ *the common value of both expressions.*

In order to give the characterization of $\mathbf{V}$, we have to recall that the (Fenchel) conjugate w^* of a map $w:[0,T]\times R^N\times\Delta(I)\to R$ is defined by:

$$w^*(t,x,\hat p)=\max_{p\in\Delta(I)} p.\hat p-w(t,x,p) \qquad \forall (t,x,\hat p)\in[0,T]\times R^N\times R^I\ .$$

In particular, $\mathbf{V}^*$ denotes the conjugate of $\mathbf{V}$. If now w is defined on the dual space $[0,T]\times R^N\times R^I$, we also denote by w^* its conjugate with respect to $\hat p$ given by:

$$w^*(t,x,p)=\max_{\hat p\in R^I} p.\hat p-w(t,x,\hat p) \qquad \forall (t,x,p)\in[0,T]\times R^N\times\Delta(I)\ .$$

Proposition 2.1 (Characterization of the value, [3])**.** *Assume that conditions (5) on* f *and on the* g_i *hold and that Isaacs assumption (6) is satisfied. Then the value function* $\mathbf{V}$ *is the unique function defined on* $[0,T]\times R^N\times\Delta(I)$ *such that:*

(i) $\mathbf{V}$ *is Lipschitz continuous in all its variables, convex with respect to* p *and such that*

$$\mathbf{V}(T,x,p)=\sum_{i=1}^I p_i g_i(x) \qquad \forall (x,p)\in R^N\times\Delta(I)\ ,$$

(ii) for any $p\in\Delta(I)$*,* $(t,x)\to\mathbf{V}(t,x,p)$ *is a viscosity subsolution of the primal HJ equation*

$$w_t+H\left(x,Dw\right)=0 \text{ in } [0,T]\times R^N \tag{12}$$

where H *is defined by (6),*

(iii) for any $\hat p\in R^I$*,* $(t,x)\to\mathbf{V}^*(t,x,\hat p)$ *is a viscosity subsolution of the dual HJ equation*

$$w_t+H^*\left(x,Dw\right)=0 \text{ in } [0,T]\times R^N \tag{13}$$

where $H^*(x,\xi)=-H(x,-\xi)$ *for any* $(x,\xi)\in R^N\times R^N$*.*

We say that $\mathbf{V}$ *is the unique dual solution of the HJ equation (12) with terminal condition* $\mathbf{V}(T,x,p)=\sum_{i=1}^I p_i g_i(x)$*.*

Remark 2.2.

(1) We recall that the notion of viscosity solutions was introduced by Crandall-Lions in [5] and first used in the framework of differential games in [7] (see also [2]). The idea of introducing the dual game and the Fenchel conjugate of the value functions comes back to De Meyer [6] for repeated games.

(2) A function w satisfying conditions (i) and (ii) only is called a dual subsolution of (12), whereas a function w satisfying (i) and (iii) only is called a dual supersolution of (12). The reason for this terminology is the following comparison principle given in [3]: if w_1 is a dual subsolution of (12) and w_2 is a dual supersolution of (12), then $w_1\le w_2$.

(3) Theorem 2.1 is given in [3] for differential games with lack of information on both sides. One easily checks that the characterization given in [3] is the same as the one given here in the special case of lack of information on one side.

We shall also need below the following reformulation of $\mathbf{V}^*$ (Lemma 4.2 of [3]):

$$\begin{aligned}\mathbf{V}^*(t,x,\hat{p}) &= \inf_{\beta\in\mathcal{B}_r(t_0)} \sup_{\alpha\in\mathcal{A}_r(t_0)} \max_{i\in\{1,\ldots,I\}} \left\{\hat{p}_i - \mathbf{E}_{\alpha\beta}\left(g_i(X_T^{t,x,\alpha,\beta})\right)\right\} \\ &= \inf_{\beta\in\mathcal{B}_r(t_0)} \sup_{\alpha\in\mathcal{A}(t_0)} \max_{i\in\{1,\ldots,I\}} \left\{\hat{p}_i - \mathbf{E}_{\alpha\beta}\left(g_i(X_T^{t,x,\alpha,\beta})\right)\right\}.\end{aligned} \qquad (14)$$

The function $\mathbf{V}^*$ is the value function of the **dual game**.

3 Approximation of the value function

Let us fix a large integer L and let us denote by $\tau = T/L$ the time-step of discretization. We set

$$t_k = k\tau \qquad \text{for} \quad k \in \{0,\ldots,L\}\ .$$

We define by backward induction on k the function $\mathbf{V}_\tau(t_k,\cdot,\cdot) : R^N \times \Delta(I) \to R$ by:

$$\text{for } k = L, \qquad \mathbf{V}_\tau(T,x,p) = \sum_i p_i g_i(x)$$

and (assuming $\mathbf{V}_\tau(t_{k+1},\cdot,\cdot)$ is built)

$$\mathbf{V}_\tau(t_k,x,p) = \mathrm{Vex}_p\left(\min_{u\in U}\max_{v\in V} \mathbf{V}_\tau(t_{k+1}, x+\tau f(x,u,v), p)\right), \qquad (15)$$

where $\mathrm{Vex}_p(\cdot)$ denote the convex hull with respect to p.

One easily checks the following.

Lemma 3.1. *Under assumption (5) on f and on the g_i, the map $(t_k,x,p) \to \mathbf{V}_\tau(t_k,x,p)$ is Lipschitz continuous with a Lipschitz constant independent of τ.*

Theorem 3.1. *Assume that conditions (5) on f and on the g_i hold and that Isaacs assumption (6) is satisfied. Then the map $\mathbf{V}_\tau$ converges uniformly to the value function $\mathbf{V}$ on compact subsets of $[0,T] \times R^N \times \Delta(I)$, in the following sense:*

$$\lim_{\substack{\tau\to 0^+,\ t_k\to t,\\ x'\to x,\ p'\to p}} \mathbf{V}_\tau(t_k,x',p') = \mathbf{V}(t,x,p) \quad \forall (t,x,p) \in [0,T]\times R^N \times \Delta(I)\ .$$

Proof of Theorem 3.1: By Lemma 3.1, the $(\mathbf{V}_\tau)$ are all Lipschitz continuous with a Lipschitz constant independent of τ. Let w be any cluster point in the topology of uniform convergence on compact subsets of $[0,T]\times R^N \times \Delta(I)$ of $\mathbf{V}_\tau$ as $\tau\to 0^+$. By construction of $\mathbf{V}_\tau$ and Lemma 3.1, w is Lipschitz continuous in all its variables, convex with respect to p and satisfies

$$w(T,x,p)=\sum_{i=1}^{I} p_i g_i(x) \qquad \forall (x,p)\in R^N\times\Delta(I)\ . \tag{16}$$

We have to prove that w is a subsolution of (12) for any p and its conjugate a subsolution of (13) for any p.

Let ϕ be a test function such that $w(\cdot,\cdot,p)-\phi$ has a strict local maximum at (t_0,x_0). Then from standard arguments (see [2]), there are (t_k,x_k) converging to (t_0,x_0) such that $\mathbf{V}_\tau(\cdot,\cdot,p)-\phi$ has a local maximum at (t_k,x_k). Then, for any $x\in R^N$,

$$\mathbf{V}_\tau(t_{k+1},x,p)-\phi(t_{k+1},x)\le \mathbf{V}_\tau(t_k,x_k,p)-\phi(t_k,x_k).$$

Thus,

$$\begin{array}{rcl} 0 &=& Vex_p\left[\min_{u\in U}\max_{v\in V}\mathbf{V}_\tau(t_{k+1},x_k+\tau f(x,u,v),p)\right]-\mathbf{V}_\tau(t_k,x_k,p)\\ &\le& \min_{u\in U}\max_{v\in V}\mathbf{V}_\tau(t_{k+1},x_k+\tau f(x,u,v),p)-\mathbf{V}_\tau(t_k,x_k,p)\\ &\le& \min_{u\in U}\max_{v\in V}\phi(t_{k+1},x_k+\tau f(x,u,v))-\phi(t_k,x_k)\ , \end{array}$$

which implies by standard arguments (see [2]) that

$$\frac{\partial\phi}{\partial t}(t_0,x_0)+\min_{u\in U}\max_{v\in V} f(x_0,u,v).\frac{\partial\phi}{\partial x}(t_0,x_0)\ \ge\ 0\ . \tag{17}$$

Hence, w is a dual subsolution of (12).

For proving that w is a supersolution in the dual sense, we first note that $\mathbf{V}_\tau^*$ uniformly converges to w^* on compact subsets of $[0,T]\times R^N\times R^I$. We also underline that (15) becomes, for any $x\in R^N$,

$$\begin{array}{rcl} \mathbf{V}_\tau^*(t,x,\hat p) &=& \max_{p\in\Delta(I)} p.\hat p-\min_{u\in U}\max_{v\in V}\mathbf{V}_\tau(t_{k+1},x+\tau f(x,u,v),p)\\ &\le& \max_{u\in U}\min_{v\in V} V_\tau^*(t_{k+1},x+\tau f(x,u,v),\hat p). \end{array}$$

Let us now fix $\hat p\in R^N$, and assume that $w^*(\cdot,\cdot,\hat p)-\phi$ has a strict local maximum at (t_0,x_0). Again from standard arguments, there are (t_k,x_k) converging to (t_0,x_0) such that $V_\tau^*(\cdot,\cdot,\hat p)-\phi$ has a local maximum at (t_k,x_k). Then

$$\begin{array}{rcl} 0 &\le& \max_{u\in U}\min_{v\in V} V_\tau^*(t_{k+1},x_k+\tau f(x,u,v),\hat p)-\mathbf{V}_\tau^*(t_k,x_k,\hat p)\\ &\le& \max_{u\in U}\min_{v\in V}\phi(t_{k+1},x_k+\tau f(x,u,v))-\phi(t_k,x_k,\hat p) \end{array}$$

from which we deduce that

$$\frac{\partial\phi}{\partial t}(t_0,x_0)+\max_{u\in U}\min_{v\in V} f(x_0,u,v).\frac{\partial\phi}{\partial x}(t_0,x_0)\ \ge\ 0\ .$$

Therefore, w is a subsolution of (13), i.e., a supersolution of (12) in the dual sense.

Since w is a sub- and supersolution of (12) in the dual sense and satisfies the terminal condition (16), we have $w = \mathbf{V}$. But this holds true for any uniform limit of $\mathbf{V}_\tau$, as $\tau \to 0^+$. Hence, $\mathbf{V}_\tau$ uniformly converges to $\mathbf{V}$. □

Remark 3.1. We note for later use that an alternative discretization of the value function $\mathbf{V}$ can be given by the relation

$$\mathbf{V}_\tau(t_k, x, p) = \mathrm{Vex}_p \left(\min_{u \in U} \max_{w \in \mathrm{Co} f(x,u,V)} \mathbf{V}_\tau(t_{k+1}, x + \tau w, p) \right) \tag{18}$$

(instead of (15)) where $\mathrm{Co} f(x,u,V)$ stands for the convex hull of the set $f(x,u,V)$.

The proof is the same as for Theorem 3.1 because we get in (17), for any $(x,\xi) \in R^N \times R^N$,

$$\begin{aligned} \min_{u \in U} \max_{w \in \mathrm{Co} f(x,u,V)} w.\xi &= \min_{u \in U} \max_{w \in f(x,u,V)} w.\xi \\ &= \min_{u \in U} \max_{v \in V} f(x,u,v).\xi \ . \end{aligned}$$

4 Construction of optimal strategies for Player I

4.1 Construction of the strategy

In order to describe precisely the strategy given in the introduction, we have to start with two definitions.

Definition 4.1 (Feedback)**.** A pair of maps $(u^*, \mathbf{p}^*)$ is a feedback if there is a partition $t_0 = 0 < t_1 < \cdots < t_L = T$ of $[0,T]$, a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and L random maps

$$(u^*(\cdot, t_k, \cdot, \cdot), \mathbf{p}^*(\cdot, t_k, \cdot, \cdot)) : \Omega \times R^N \times \Delta(I) \to U \times \Delta(I)$$

(for $k \in \{0, \ldots, L-1\}$) such that:

(i) for any $x_0, \ldots, x_{L-1}$ in R^N and any $p_0, \ldots, p_{L-1}$ in $\Delta(I)$, the family of random variables $\{(u^*(\cdot, t_j, x_j, p_j), \mathbf{p}^*(\cdot, t_j, x_j, p_j))\}_{j=1,\ldots,L-1}$ are mutually independent.

(ii) for any $k \in \{0, \ldots, L-1\}$, the map $(\omega, x, p) \to (u^*(\omega, t_k, x, p), \mathbf{p}^*(\omega, t_k, x, p))$ is measurable when R^N, $\Delta(I)$ and U are endowed with the Borel σ-algebra.

Definition 4.2 (Strategy associated with a feedback)**.** Given a feedback $(u^*, \mathbf{p}^*)$, an initial time $\bar{t} \in [0,T]$, an initial position $\bar{x} \in R^N$ and an initial probability $\bar{p} \in \Delta(I)$, one defines a random strategy $\alpha \in \mathcal{A}_r(\bar{t})$ by setting:

- $(\Omega_\alpha, \mathcal{F}_\alpha, \mathbf{P}_\alpha) := (\Omega, \mathcal{F}, \mathbf{P})$, where $(\Omega, \mathcal{F}, \mathbf{P})$ is the probability space on which $(u^*, \mathbf{p}^*)$ is defined.

- α and the process (p_{t_k}) are defined by induction on the time intervals of the form $[t_k, t_{k+1})$, where $t_0 = 0 < \cdots < t_L = T$ is the partition associated with the feedback $(u^*, \mathbf{p}^*)$:
  - Let k_0 be such that $\bar{t} \in [t_{k_0}, t_{k_0+1})$. Then we set $\alpha(\omega, v)(t) = u^*(\omega, t_{k_0}, \bar{x}, \bar{p})$ on $[\bar{t}, t_{k_0+1})$ and $p_{t_{k_0}} = \bar{p}$.
  - Let $k \geq k_0 + 1$. Assume that we have defined $p_{t'_k}$ for all $k' \leq k$ and α on $[\bar{t}, t_k)$. Then we extend α to $(t_k, t_{k+1}]$ by setting

$$\alpha(\omega, v)(t) = u^*(\omega, t_k, X_{t_k}^{\bar{t},\bar{x},\alpha,v}, p_{t_k}(\omega)) \text{ for } t \in (t_k, t_{k+1}],\ v \in \mathcal{V}(\bar{t})$$

and we update the process (p_{t_k}) by setting

$$p_{t_{k+1}}(\omega) = \mathbf{p}^*(\omega, t_k, X_{t_k}^{\bar{t},\bar{x},\alpha,v}, p_{t_k}(\omega)).$$

We say that α is the random strategy associated to the feedback $(u^*, \mathbf{p}^*)$ at $(\bar{t}, \bar{x}, \bar{p})$.

Remark 4.1. The above notion of strategy is strongly inspired by Krasowskii-Subbotin step-by-step motions associated with feedbacks [8].

We are now ready to explain the construction of the feedbacks which give rise to ϵ-optimal strategies for Player I. We actually have to built I feedbacks, because Player I can take advantage of his knowledge of the index i. As before, we consider a large integer L, $\tau = T/L$ denotes the time-step. We also set $\Omega = [0,1]^L$, $\mathcal{F}$ the Borel σ-algebra and $\mathbf{P}$ the Lebesgue measure on Ω.

For each $i \in \{1, \ldots, I\}$, $t_k = k\tau$ (with $k \in \{0, \ldots, L-1\}$), $x \in R^N$ and $p \in \Delta(I)$, we define $(u_i^*(\omega, t_k, x, p), \mathbf{p}_\mathbf{i}^*(\omega, t_k, x, p))$ as follows: let $\lambda = (\lambda_j) \in \Delta(I+1)$ and $p^j \in \Delta(I)$ for $j = 1, \ldots, (I+1)$ be such that

$$\sum_{j=1}^{I+1} \lambda_j p^j = p$$

and

$$\sum_{j=1}^{I+1} \lambda_j \left(\min_{u \in U} \max_{w \in \mathrm{Co}f(x,u,V)} \mathbf{V}_\tau(t_{k+1}, x + \tau w, p^j) \right) \leq \mathbf{V}_\tau(t_k, x, p) + \tau^2\ .$$

Let $u^j \in U$ be τ^2-optimal in $\min_{u \in U} \max_{w \in \mathrm{Co}f(x,u,V)} \mathbf{V}_\tau(t_{k+1}, x + \tau w, p^j)$. We also choose the functions $\lambda = \lambda(x,p)$, $p^j = p^j(x,p)$ and $u^j = u^j(x,p)$ in such a way that they are Borel measurable with respect to (x,p) (this is why we only require them to be τ^2-optimal). Then we set

$$(u_i^*(\omega, t_k, x, p), \mathbf{p}_\mathbf{i}^*(\omega, t_k, x, p)) = (u^j, p^j)$$
$$\text{if} \quad \omega_k \in \left[\sum_{j' < j} \lambda_{j'} p_i^{j'} / p_i, \sum_{j' \leq j} \lambda_{j'} p_i^{j'} / p_i \right)$$

for $j \in \{1, \ldots, I+1\}$ and $i \in \{1, \ldots, I\}$ if $p_i \neq 0$, and

$$(u_i^*(\omega, t_k, x, p), \mathbf{p}_\mathbf{i}^*(\omega, t_k, x, p)) = (u^1, p^1)$$

for any ω otherwise. If $p_i \neq 0$, this means that (u^j, p^j) is chosen with probability $\lambda_j p_i^j / p_i$.

Let us also point out that by construction, the feedback $(u_i^*, \mathbf{p}_\mathbf{i}^*)$ depends on $i \in \{1, \ldots, I\}$ and on $\tau = 1/L$. For simplicity we do not explicitly note this latter dependence.

4.2 Optimality of the strategy

The following result states that the strategy built in the previous way is ϵ-optimal.

Theorem 4.1. *Assume that conditions (5) on f and on the g_i hold and that Isaacs assumption (6) is satisfied. For any $\epsilon > 0$, $\bar{t} \in [0, T)$, $\bar{x} \in R^N$ and $\bar{p} \in \Delta(I)$, there is some $\tau_0 > 0$ such that the strategy $\hat{\alpha}_\tau = (\alpha_{\tau,i})$ associated to the feedbacks $(u_i^*, \mathbf{p}_\mathbf{i}^*)$ at $(\bar{t}, \bar{x}, \bar{p})$ is ϵ-optimal for Player I in the game $\mathbf{V}(\bar{t}, \bar{x}, \bar{p})$ for any $\tau \in (0, \tau_0)$.*

Remark 4.2. By "$\hat{\alpha}_\tau$ is ϵ-optimal for Player I in the game $\mathbf{V}(\bar{t}, \bar{x}, \bar{p})$" we mean that $\hat{\alpha}_\tau$ is ϵ-optimal in (10).

Proof of Theorem 4.1: Let us fix $\tau = T/L$ (with $L > 1$) and let $(u_i^*, \mathbf{p}_\mathbf{i}^*)$ be the feedback built above. We first claim that there is some constant M such that, for any $k \in \{0, \ldots, L\}$, $\bar{x} \in R^N$, $\bar{p} \in \Delta(I)$, we have

$$\sup_{\beta \in \mathcal{B}(t_k)} \mathcal{J}(t_k, \bar{x}, \hat{\alpha}_\tau, \beta, \bar{p}) \leq \mathbf{V}_\tau(t_k, \bar{x}, \bar{p}) + M(L-k)\tau^2, \tag{19}$$

where $\hat{\alpha}_\tau = (\alpha_{\tau,i})$ is the random strategy associated to the feedbacks $(u_i^*, \mathbf{p}_i^*)$ at $(t_k, \bar{x}, \bar{p})$.

In order to define the constant M, let us note that there is a constant M_1 which only depends on $\|f\|_\infty$ and on $\|Df\|_\infty$ (see assumption (5) on f) such that, for any $k \in \{0, \ldots, L-1\}$ and any $x \in R^N$, for any $u \in U$ and any control $v \in \mathcal{V}(t_k)$, there is some $w \in \mathrm{Co} f(x, u, V)$ such that

$$\left| X_{t_{k+1}}^{t_k, x, u, v} - (x + \tau w) \right| \leq M_1 \tau \ .$$

Let also denote by M_2 a Lipschitz constant of $\mathbf{V}_\tau$ (which does not depend on τ). Then we set $M = (M_1 M_2 + 2)$.

We show inequality (19) by backward induction. It obviously holds true for $k = L$. Let us assume that it holds true for some $k+1 \in \{1, \ldots, L\}$. Let $\bar{x} \in R^N$

and $\bar{p} \in \Delta(I)$ be fixed. We note that, from the definition of the feedback $(u_i^*, \mathbf{p_i^*})$ (see the notations of the previous subsection), for any $\beta \in \mathcal{B}(t_k)$, we have

$$X_T^{t_k, \bar{x}, \alpha_{\tau,i}, \beta} = X_T^{t_{k+1}, x_j, \alpha_{ij}, \beta_j} \text{ with probability } \lambda_j p_i^j / \bar{p}_i,$$

where

- $x_j = X_{t_{k+1}}^{t_k, \bar{x}, u^j, \beta(u^j)}$,
- α_{ij} is the random strategy associated to the feedback $(u_i^*, \mathbf{p}_i^*)$ at (t_{k+1}, x_j, p^j)
- and $\beta_j \in \mathcal{B}(t_{k+1})$ is defined, for any $u \in \mathcal{U}(t_{k+1})$, by $\beta_j(u) = \beta(u')$ where $u' \in \mathcal{U}(t_k)$ is given by $u'(s) = u^j$ on $[t_k, t_{k+1})$ and $u'(s) = u(s)$ on $[t_{k+1}, T]$.

From the definition of M_1, there is some $w^j \in \mathrm{Co}f(\bar{x}, u^j, V)$ such that:

$$|x_j - (\bar{x} + \tau w^j)| \leq M_1 \tau^2 \ . \tag{20}$$

Hence,

$$\begin{aligned} \mathcal{J}(t_k, \bar{x}, \hat{\alpha}_\tau, \beta, \bar{p}) \ &= \ \sum_{i=1}^{I} \bar{p}_i \mathbf{E}_{\alpha_{\tau,i}} \left[g_i(X_T^{t_k, \bar{x}, \alpha_{\tau,i}, \beta} \right] \\ &= \ \sum_{i=1}^{I} \sum_{j=1}^{I+1} \bar{p}_i \lambda_j p_i^j / \bar{p}_i \mathbf{E}_{\alpha_{ij}} \left[g_i(X_T^{t_{k+1}, x_j, \alpha_{ij}, \beta_j}) \right] \\ &\leq \ \sum_{j=1}^{I+1} \lambda_j \sup_{\beta' \in \mathcal{B}(t_{k+1})} \sum_{i=1}^{I} p_i^j \mathbf{E}_{\alpha_{ij}} \left[g_i(X_T^{t_{k+1}, x_j, \alpha_{ij}, \beta'}) \right] . \end{aligned}$$

From the induction assumption we have:

$$\begin{aligned} &\sup_{\beta' \in \mathcal{B}(t_{k+1})} \sum_{i=1}^{I} p_i^j \mathbf{E}_{\alpha_{ij}} \left[g_i(X_T^{t_{k+1}, x_j, \alpha_{ij}, \beta'}) \right] \\ &\leq \mathbf{V}_\tau(t_{k+1}, x_j, p^j) + M(L - (k+1))\tau^2. \end{aligned}$$

Therefore, using first (20) and the M_2-Lipschitz continuity of $\mathbf{V}_\tau$, then the choice of u^j, λ_j, and p^j,

$$\begin{aligned} \mathcal{J}(t_k, \bar{x}, \hat{\alpha}_\tau, \beta, \bar{p}) \ &\leq \ \sum_{j=1}^{I+1} \lambda_j \left(\mathbf{V}_\tau(t_{k+1}, x_j, p^j) + M(L - (k+1))\tau^2 \right) \\ &\leq \ \sum_{j=1}^{I+1} \lambda_j \left(\mathbf{V}_\tau(t_{k+1}, \bar{x} + \tau w^j, p^j) + M_1 M_2 \tau^2 + M(L - (k+1))\tau^2) \right) \\ &\leq \ \sum_{j=1}^{I+1} \lambda_j \max_{w \in \mathrm{Co}f(x, u^j, V)} \left(\mathbf{V}_\tau(t_{k+1}, \bar{x} + \tau w, p^j) + M_1 M_2 \tau^2 \right. \\ &\qquad\qquad\qquad \left. + M(L - (k+1))\tau^2) \right) \\ &\leq \ \mathbf{V}_\tau(t_k, \bar{x}, \bar{p}) + (M_1 M_2 + 2)\tau^2 + M(L - (k+1))\tau^2. \end{aligned}$$

So inequality (19) holds because $M = (M_1 M_2 + 2)$.

Let us now fix $\bar{t}$, $\bar{x}$, and $\bar{p}$ and let $\hat{\alpha}_\tau = (\alpha_{\tau,i})$ be the admissible strategy associated with the feedbacks $(u_i^*, \mathbf{p}_i^*)$ at $(\bar{t}, \bar{x}, \bar{p})$. Let k be such that $\bar{t} \in [t_{k-1}, t_k)$. Then we have:

$$\begin{aligned} \sup_{\beta \in \mathcal{B}_r(\bar{t})} \mathcal{J}(\bar{t}, \bar{x}, \hat{\alpha}_\tau, \beta, \bar{p}) \ &= \ \sup_{\beta \in \mathcal{B}(\bar{t})} \mathcal{J}(\bar{t}, \bar{x}, \hat{\alpha}_\tau, \beta, \bar{p}) \\ &\leq \ \max_{|x - \bar{x}| \leq M\tau} \sup_{\beta \in \mathcal{B}(t_k)} \mathcal{J}(t_k, x, \hat{\alpha}_x, \beta, \bar{p}), \end{aligned}$$

where $\hat{\alpha}_x$ is the random strategy associated with the feedbacks $(u_i^*, \mathbf{p}_i^*)$ at $(t_k, x, \bar{p})$ and M is a bound on $|f|$. Hence by (19),

$$\sup_{\beta\in\mathcal{B}_r(\bar{t})} \mathcal{J}(\bar{t},\bar{x},\hat{\alpha}_\tau,\beta,\bar{p}) \ \le\ \max_{|x-\bar{x}|\le M\tau} \mathbf{V}_\tau(t_k,x,\bar{p}) + M(L-k)\tau^2.$$

Letting $\tau = T/L \to 0^+$, $\max_{|x-\bar{x}|\le M\tau} \mathbf{V}_\tau(t_k,x,\bar{p})$ converges to $\mathbf{V}(\bar{t},\bar{x},\bar{p})$ and $(L-k)\tau^2$ converges to 0. Therefore, for τ small enough, say $\tau \in (0,\tau_0)$, we have

$$\sup_{\beta\in\mathcal{B}_r(\bar{t})} \mathcal{J}(\bar{t},\bar{x},\hat{\alpha}_\tau,\beta,\bar{p}) \ \le\ \mathbf{V}(\bar{t},\bar{x},\bar{p}) + \epsilon\ .$$

□

5 Approximation of the dual game and construction of a ϵ-optimal strategy for Player II

In this section we build a ϵ-optimal strategy of the non informed player. The main idea is that it is enough to play a ϵ-optimal strategy in the dual game. As for the first player, one needs for this to build a time-approximation of the conjugate of the value function. Then we show that ϵ-optimal strategies in the dual game $\mathbf{V}^*(t,x,\hat{p})$ are ϵ-optimal strategies in the initial game $\mathbf{V}(t,x,p)$ provided $\hat{p} \in \partial_p \mathbf{V}(t,x,p)$. We finally build the ϵ-optimal strategy for Player II.

Let us fix as before a large integer L, set $\tau = T/L$ and $t_k = k\tau$ for $k \in \{0,\ldots,L\}$. We define by backward induction the function $\mathbf{C}_\tau$ (with $\mathbf{C}$ as "conjugate")

$$\text{for } k = L, \qquad \mathbf{C}_\tau(T,x,p) = \max_i\{\hat{p}_i - g_i(x)\}$$

and (assuming $\mathbf{C}_\tau(t_{k+1},\cdot,\cdot)$ is built)

$$\mathbf{C}_\tau(t_k,x,\hat{p}) = \mathrm{Vex}_{\hat{p}}\left(\min_{v\in V}\ \max_{w\in \mathrm{Co}f(x,U,v)} \mathbf{C}_\tau(t_{k+1},x+\tau w,\hat{p})\right) \tag{21}$$

where $\mathrm{Vex}_{\hat{p}}(\cdot)$ denote the convex hull with respect to $\hat{p}$ and $\mathrm{Co}f(x,U,v)$ the convex hull of the set $f(x,U,v)$.

One easily checks the following.

Lemma 5.1. *Assume that conditions (5) on f and on the g_i hold and that Isaacs assumption (6) is satisfied. Then the map $(t_k,x,\hat{p}) \to \mathbf{C}_\tau(t_k,x,\hat{p})$ is Lipschitz continuous with a Lipschitz constant independent of τ.*

Theorem 5.1. *Assume that conditions (5) on f and on the g_i hold and that Isaacs assumption (6) is satisfied. Then the map $\mathbf{C}_\tau$ converges uniformly to the conjugate of the value function $\mathbf{V}^*$ on compact subsets of $[0,T]\times R^N \times R^I$, i.e.,*

$$\lim_{\substack{\tau\to 0^+,\ t_k\to t,\\ x'\to x,\ \hat{p}'\to\hat{p}}} \mathbf{C}_\tau(t_k,x',\hat{p}') = \mathbf{V}^*(t,x,\hat{p}) \quad \forall (t,x,\hat{p}) \in [0,T]\times R^N\times R^I\ .$$

Remark 5.1. As in the construction of the approximation of the primal game, we could use instead of formula (21) the following:

$$\mathbf{C}_\tau(t_k,x,\hat{p}) = \mathrm{Vex}_{\hat{p}}\left(\min_{v\in V}\max_{u\in U}\mathbf{C}_\tau(t_{k+1},x+\tau f(x,u,v),\hat{p})\right).$$

Proof of Theorem 5.1: The proof is exactly the same as that of Theorem 3.1. We show that, if w is any cluster point in the topology of uniform convergence on compact subsets of $[0,T]\times R^N\times R^I$ of $\mathbf{C}_\tau$ as $\tau\to 0^+$, then w is Lipschitz continuous in all its variables, is convex with respect to $\hat{p}$ and satisfies

$$w(T,x,\hat{p}) = \max_{i=,\dots,I}\{\hat{p}_i - g_i(x)\} \qquad \forall (x,p)\in R^N\times R^I\ . \tag{22}$$

Moreover, arguing as in the proof of Theorem 3.1, one can check that w is a subsolution of the dual HJ Eq. (13) and w^* is a subsolution of the primal HJ Eq. (12), which shows that $w^*=\mathbf{V}$, i.e., $w=\mathbf{V}^*$. Hence, any accumulation point of $\mathbf{V}_\tau$ is equal to $\mathbf{V}^*$, which shows our claim. □

We now explain that Player II can use ϵ-optimal strategies in the dual game in order to play in the initial game.

Lemma 5.2. *Assume that conditions (5) on f and on the g_i hold and that Isaacs assumption (6) is satisfied. Let $\hat{p}\in\partial_p\mathbf{V}(t_0,x_0,p)$ and $\bar{\beta}\in\mathcal{B}_r(t_0)$ be ϵ-optimal in the dual game $\mathbf{V}^*(t_0,x_0,\hat{p})$ (i.e., in (14)). Then $\bar{\beta}$ is ϵ-optimal in the game $\mathbf{V}(t_0,x_0,p)$ (i.e., in (11)).*

Proof of Lemma 5.2: Let $\bar{\beta}$ be ϵ-optimal in the dual game $\mathbf{V}^*(t_0,x_0,\hat{p})$. Then we have

$$\begin{aligned}\mathbf{V}^*(t_0,x_0,\hat{p}) &= \inf_\beta\sup_\alpha \max_i\{\hat{p}_i - \mathbf{E}_\beta(g_i(X_T^{t_0,x_0,\alpha,\beta}))\}\\ &\geq \max_i\{\hat{p}_i - \inf_\alpha \mathbf{E}_{\bar{\beta}}(g_i(X_T^{t_0,x_0,\alpha,\bar{\beta}}))\} - \epsilon.\end{aligned}$$

Hence,

$$\begin{aligned}\mathbf{V}(t_0,x_0,p) &= \hat{p}.p - \mathbf{V}^*(t_0,x_0,\hat{p})\\ &\leq \hat{p}.p - \max_i\{\hat{p}_i - \inf_\alpha \mathbf{E}_{\bar{\beta}}(g_i(X_T^{t_0,x_0,\alpha,\bar{\beta}}))\}+\epsilon\\ &\leq \hat{p}.p - \textstyle\sum_i p_i\{\hat{p}_i - \inf_\alpha \mathbf{E}_{\bar{\beta}}(g_i(X_T^{t_0,x_0,\alpha,\bar{\beta}}))\}+\epsilon\\ &\leq \textstyle\sum_i p_i\{\inf_\alpha \mathbf{E}_{\bar{\beta}}(g_i(X_T^{t_0,x_0,\alpha,\bar{\beta}}))\}+\epsilon\\ &\leq \inf_{\hat{\alpha}}\mathcal{J}(t_0,x_0,\hat{\alpha},\bar{\beta},p)+\epsilon.\end{aligned}$$

So $\bar{\beta}$ is ϵ-optimal in the game $\mathbf{V}(t_0,x_0,p)$. □

We are now ready to describe an ϵ-optimal strategy for Player II. As for Player I, Player II has to build a random process $(\hat{p}_{t_k})$ (now in R^I) jointly with his strategy. For this he constructs a feedback $(v^*,\hat{\mathbf{p}}^*)$ (defined here on $\Omega\times\{t_k\}\times R^N\times R^I$) by using the time discretization of the game explained above.

Let L be a large integer, let $\tau = T/L$ be the time-step and let us set $\Omega = [0,1]^L$, $\mathcal{F}$ the Borel σ-algebra, and $\mathbf{P}$ the Lebesgue measure on Ω. Let $k \in \{0,\ldots,L-1\}$, $\hat{p} = (\hat{p}_i) \in R^I$, $x \in R^N$ and $\omega \in \Omega$ be fixed. We define $(v^*(\omega,t_k,x,\hat{p}), \hat{\mathbf{p}}^*(\omega,t_k,x,\hat{p}))$ as follows : let $\lambda = (\lambda_j) \in \Delta(I+1)$ and $\hat{p}^j \in R^I$ for $j = 1,\ldots,(I+1)$ be such that

$$\sum_{j=1}^{I+1} \lambda_j p^j = \hat{p}$$

and

$$\sum_{j=1}^{I+1} \lambda_j \left(\min_{v\in V} \max_{w\in \mathrm{Co}f(x,U,v)} \mathbf{C}_\tau(t_{k+1}, x+\tau w, \hat{p}^j) \right) \le \mathbf{C}_\tau(t_k,x,\hat{p}) + \tau^2 \ .$$

Let $v^j \in V$ be τ^2-optimal in $\min_{v\in V} \max_{w\in \mathrm{Co}f(x,U,v)} \mathbf{C}_\tau(t_{k+1}, x+\tau w, \hat{p}^j)$. We choose the functions $\lambda = \lambda(x,\hat{p})$, $\hat{p}^j = \hat{p}^j(x,\hat{p})$ and $v^j = v^j(x,\hat{p})$ in such a way that they are Borel measurable with respect to $(x,\hat{p})$. Then we set

$$(v^*(\omega,t_k,x,\hat{p}), \hat{\mathbf{p}}^*(\omega,t_k,x,\hat{p})) = (u^j, \hat{p}^j) \quad \text{if} \quad \omega_k \in \left[\sum_{j'<j} \lambda_{j'}, \sum_{j'\le j} \lambda_{j'} \right)$$

for $j \in \{1,\ldots,I+1\}$. This means that Player II plays the control v^j on $[t_k,t_{k+1})$ with probability λ_j, while the random process $(\hat{p}_{t_k})$ moves from $\hat{p}$ to $\hat{p}^j$.

We point out that the feedback $(v^*,\hat{\mathbf{p}}^*)$ depends on $\tau > 0$, although we do not write it explicitly. As in Definition 4.2, we can associate with such a feedback a random strategy (which is now built jointly with a random process $(\hat{p}_{t_k})$ on R^I).

Theorem 5.2. *Assume that conditions (5) on f and on the g_i hold and that Isaacs assumption (6) is satisfied. Let $\epsilon > 0$, $\bar{t} \in [0,T]$, $\bar{x} \in R^N$ and $\bar{p} \in \Delta(I)$. Let also $\hat{p} \in \partial_p \mathbf{V}(\bar{t},\bar{x},\bar{p})$. Then there is some $\tau_0 > 0$ such that the strategy β_τ associated to the feedback $(v^*,\hat{\mathbf{p}}^*)$ at $(\bar{t},\bar{x},\bar{p})$ is ϵ-optimal for Player II in the game $\mathbf{V}(\bar{t},\bar{x},p)$ for any $\tau \in (0,\tau_0)$.*

Proof of Theorem 5.2: From Lemma 5.2 we just have to prove that β_τ is ϵ-optimal in the dual game $\mathbf{V}^*(\bar{t},\bar{x},\hat{p})$. Following the proof of Theorem 4.1, it is enough to show that there is some constant M such that, for any $k \in \{0,\ldots,L\}$, any $\bar{x} \in R^N$ and any $\hat{p} \in R^I$, we have

$$\sup_{\alpha\in\mathcal{A}(t_k)} \mathcal{J}_1(t_k,\bar{x},\alpha,\beta_\tau,\hat{p}) \le \mathbf{C}_\tau(t_k,\bar{x},\hat{p}) + M(L-k)\tau^2, \qquad (23)$$

where β_τ is the random strategy associated with the feedback $(v^*,\hat{\mathbf{p}}^*)$ at $(t_k,\bar{x},\bar{p})$ and where we have set, for any $\bar{t} \in [0,T]$, $\bar{x} \in R^N$, $\hat{p} \in R^I$ and $(\alpha,\beta) \in$

$\mathcal{A}_r(\bar{t}) \times \mathcal{B}_r(\bar{t})$,

$$\mathcal{J}_1(\bar{t}, \bar{x}, \alpha, \beta, \hat{p}) = \max_{i=1,\ldots,I} \left(\hat{p}_i - \mathbf{E}_{\alpha\beta}\left[g_i(X_T^{\bar{t},\bar{x},\alpha,\beta})\right]\right) .$$

In order to define the constant M, we note as in the proof of Theorem 4.1 that there is a constant M_1 which only depends on $\|f\|_\infty$ and on $\|Df\|_\infty$ such that, for any $k \in \{0, \ldots, L-1\}$, $x \in R^N$, $u \in \mathcal{U}(t_k)$ and $v \in V$, there is some $w \in \mathrm{Co}f(x, U, v)$ such that:

$$\left|X_{t_{k+1}}^{t_k,x,u,v} - (x + \tau w)\right| \le M_1 \tau .$$

Let M_2 be a Lipschitz constant of $\mathbf{C}_\tau$ (which does not depend on τ). Then we set $M = (M_1 M_2 + 2)$.

We show inequality (23) by backward induction. In obviously holds true for $k = L$. Let us assume that it holds true for some $k + 1 \in \{1, \ldots, L\}$.

Let $\bar{x} \in R^N$ and $\hat{p} \in R^I$ be fixed. From the definition of $(v^*, \hat{\mathbf{p}}^*)$ we have, for any $\alpha \in \mathcal{A}(t_k)$,

$$X_T^{t_k,\bar{x},\alpha,\beta_\tau} = X_T^{t_{k+1},x_j,\alpha_j,\beta_j} \text{ with probability } \lambda_j,$$

where

- $x_j = X_{t_{k+1}}^{t_k,\bar{x},\alpha,v^j}$,
- β_j is the random strategy associated to the feedback $(v^*, \hat{\mathbf{p}}^*)$ at $(t_{k+1}, x_j, \hat{p}^j)$
- and $\alpha_j \in \mathcal{A}(t_{k+1})$ is defined, for any $v \in \mathcal{V}(t_{k+1})$ by $\alpha_j(v) = \alpha(v')$, where $v'(s) = v^j$ on $[t_k, t_{k+1})$ and $v'(s) = v(s)$ on $[t_{k+1}, T]$.

Let $w^j \in \mathrm{Co}f(\bar{x}, U, v^j)$ be such that

$$|x_j - (\bar{x} + \tau w^j)| \le M_1 \tau^2 . \tag{24}$$

Then:

$$\begin{aligned}
\mathcal{J}_1(t_k, \bar{x}, \alpha, \beta_\tau, \hat{p}) \ &= \ \max_i \left(\hat{p}_i - \mathbf{E}_{\beta_\tau}\left[g_i(X_T^{t_k,\bar{x},\alpha,\beta_\tau}\right]\right) \\
&= \ \max_i \left(\sum_{j=1}^{I+1} \lambda_j (\hat{p}_i^j - \mathbf{E}_{\beta_j}\left[g_i(X_T^{t_{k+1},x_j,\alpha_j,\beta_j})\right])\right) \\
&\le \ \sum_{j=1}^{I+1} \lambda_j \sup_{\alpha' \in \mathcal{A}(t_{k+1})} \max_i \left(\hat{p}_i^j - \mathbf{E}_{\beta_j}\left[g_i(X_T^{t_{k+1},x_j,\alpha',\beta_j})\right]\right) .
\end{aligned}$$

From the induction assumption we have:

$$\sup_{\alpha' \in \mathcal{A}(t_{k+1})} \max_i \left(\hat{p}_i^j - \mathbf{E}_{\beta_j}\left[g_i(X_T^{t_{k+1},x_j,\alpha',\beta_j})\right]\right)$$
$$\le \mathbf{C}_\tau(t_{k+1}, x_j, \hat{p}^j) + M(L - (k+1))\tau^2 .$$

Therefore, using first (24) and the M_2-Lipschitz continuity of $\mathbf{C}_\tau$, then the choice of v^j, λ_j and $\hat{p}^j$, we get as in the proof of Theorem 4.1

$$\begin{aligned}
\mathcal{J}_1(t_k, \bar{x}, \alpha, \beta_\tau, \hat{p}) \ &\le \ \sum_{j=1}^{I+1} \lambda_j \left(\mathbf{C}_\tau(t_{k+1}, x_j, \hat{p}^j) + M(L - (k+1))\tau^2\right) \\
&\le \ \mathbf{C}_\tau(t_k, \bar{x}, \hat{p}) + (M_1 M_2 + 2)\tau^2 + M(L - (k+1))\tau^2,
\end{aligned}$$

when inequality (23) holds thanks to the choice of $M = (M_1 M_2 + 2)$. □

References

[1] Aumann, R. J.; Maschler, M. B., Repeated games with incomplete information. With the collaboration of Richard E. Stearns. MIT Press, Cambridge, MA, 1995.

[2] Bardi, M., Capuzzo-Dolcetta, I., Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations. Systems and Control: Foundations and Applications. Birkhäuser, Boston, 1997.

[3] Cardaliaguet, P., *Differential games with asymmetric information*, SIAM J. Control Optim. 46 (2007), no. 3, 816–838.

[4] Cardaliaguet, P., *Representation formulae for differential games with asymmetric information*, to appear in JOTA.

[5] Crandall M.G., Lions P.L., *Viscosity solutions of Hamilton-Jacobi equations.* Trans. Amer. Math. Soc. (1983), 277, 1-42.

[6] De Meyer, B., *Repeated games, duality and the central limit theorem.* Math. Oper. Res. 21 (1996), no. 1, 237–251.

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

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

[9] Sorin, S., A first course on zero-sum repeated games. Mathématiques & Applications (Berlin), 37. Springer-Verlag, Berlin, 2002.

Fully-Discrete Schemes for the Value Function of Pursuit-Evasion Games with State Constraints

Emiliano Cristiani
Dipartimento di Matematica
SAPIENZA – Università di Roma
Rome, Italy
`emiliano.cristiani@gmail.com`

Maurizio Falcone
Dipartimento di Matematica
SAPIENZA – Università di Roma
Rome, Italy
`falcone@mat.uniroma1.it`

Abstract

We deal with the approximation of a generalized Pursuit-Evasion game with state constraints. Our approach is based on the Dynamic Programming principle and on the characterization of the lower value v of the game via the Isaacs equation. Our main result is the convergence of the fully-discrete scheme for Pursuit-Evasion games under continuity assumptions on v and some geometric assumptions on the dynamics and on the set of constraints $\overline{\Omega}$. We also analyze the Tag-Chase game in a bounded convex domain when the two players have the same velocity and we prove that in the constrained case the time of capture is finite. Some hints to improve the efficiency of the algorithm on serial and parallel machines will be also given. An extensive numerical section will show the accuracy of our method for the approximation of the value function v and of the corresponding optimal trajectories in a number of different configurations.

Key words. Differential games, pursuit-evasion games, state constraints, Isaacs equation, fully-discrete scheme, feedback controls, Tag-Chase game, parallel algorithms.

AMS Subject Classifications. Primary 65M12; Secondary 49N70, 49L25.

1 Introduction

In this paper, we present and analyze a numerical approximation scheme for 2-player Pursuit-Evasion games with state constraints. The scheme is based on the dynamic programming approach and the convergence results are obtained in the

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_11,

framework of viscosity solutions (see the survey papers [3,7,17] for a general introduction to this topic without constraints).

Our main result shows that the solution of the fully-discrete problem converges to the time-discrete value function as the mesh size Δx goes to zero provided a technical “consistency" assumption on the triangulation is satisfied. Moreover, an *a priori* error bound on that approximation is proved in Theorem 3.1 and a very easy sufficient condition guaranteeing consistency is shown (Corollary 3.3). The proof of the main result is obtained by extending to games a technique presented in [18] for the minimum time problem without state constraints and adapting the approach for state-constrained control problems presented in [10].

Note that the convergence of the fully-discrete solution to the solution of the continuous problem in the free (*i.e.*, unconstrained) case is proved in [6], but this result cannot be directly extended to the constrained case. In [10], a convergence result is proved for constrained control problems, but it strictly relies on the fact that the time-discrete value function is continuous so we cannot apply the same ideas here.
In order to prove convergence of the approximate solution of the fully-discrete scheme to the value function, we have coupled our result with the convergence result obtained by Bardi et al. [8] (see also [21,20]) in the framework of Pursuit-Evasion games. This allows to conclude that, under suitable assumptions, the convergence of the fully-discrete solution converge to the solution of the continuous problem when the time and space steps, Δt and Δx, go to zero (although a precise estimate of the order of convergence is still missing).

It should be noted that very few results on *constrained* differential games are available although several interesting problems with state constraints have been studied in the classical books by Isaacs [19] and Breakwell [9]. The aim of those contributions is mainly to analyze the games under study in order to determine directly the optimal strategies for the players *avoiding* in this way the Isaacs equation. The main theoretical contributions to the characterization of the value function for state constrained problems are, by our knowledge, the papers by Alziary de Roquefort [1], Bardi *et al.* [8], and Cardaliaguet *et al.* [12]. From the numerical point of view, the list of contributions is even shorter. The first examples of computed optimal trajectories for Pursuit-Evasion games appeared in the work by Alziary de Roquefort [2]. In Bardi *et al.* [6], there are some interesting tests in $\Omega \subset \mathbb{R}^2$ with state constraints and discontinuous value function. In [4], the effect of the boundary conditions for the free problem in $\mathbb{R}^4$ is studied. In the paper Cardaliaguet *et al.* [11], a modified viability kernel algorithm (see [13] for more details on this approach) is presented and a convergence proof for that approximation scheme is given. Finally, let us also mention the paper of Pesch *et al.* [22] where the optimal trajectories are computed by means of neural networks (again avoiding the solution of the Isaacs equation).

Another contribution of this paper is the analysis of the constrained Tag-Chase game where the two players run one after the other in a bounded convex domain *with the same velocity*. In this case, the value function is discontinuous and most of the theoretical results we know for the Tag-Chase game do not hold. We prove that the time of capture is finite if the capture occurs whenever the distance between the Pursuer and the Evader is less than a positive parameter ε (the radius of capture). This shows that the presence of constraints can change dramatically the result of the game. In fact, in the unconstrained Tag-Chase game where both the players have the same velocity, the Evader always wins (and the time of capture is $+\infty$). The paper is organized as follows. In Sec. 2, we set up our problem, introduce the notations and present our approximation scheme. Section 3 is devoted to the convergence analysis, we prove first some properties of the discrete value functions v_h and v_h^k corresponding, respectively, to the solutions of the time-discrete and fully-discrete schemes. The final convergence result is obtained coupling the error estimate of Theorem 3.1 with the results in [8]. In Sec. 4, we deal with the Tag-Chase game in a convex domain showing that this problem has a finite capture time also when the two players have the same maximal velocity. Section 5 presents some hints for the construction of the algorithms and, in particular, it deals with two features which allow to reduce the computational cost for the solution of the Isaacs equation: a high-dimensional interpolation technique and the symmetry properties of the Tag-Chase game played in a square domain. Finally, Sec. 6 presents several tests for different geometric configurations of the state constraints (convex and non-convex) as well as for various choices of the relative velocities of the two players. We analyze the results in terms of the value function v but also in terms of the optimal trajectories that one can compute using v.

2 The fully-discrete approximation scheme

Let us start introducing the problem and our notations. The system describing the dynamics is

$$\begin{cases} \dot{y}(t) = f(y(t), a(t), b(t)) \,, \ \ t > 0 \\ y(0) = x \,, \end{cases} \tag{1}$$

where $y(t) \in \mathbb{R}^{2n}$ is the state of the system, $a(\cdot) \in \mathcal{A}$ and $b(\cdot) \in \mathcal{B}$ are, respectively, the controls of the first and the second player, $\mathcal{A}$ and $\mathcal{B}$ being the sets of *admissible strategies* defined as

$$\mathcal{A} = \{a(\cdot) : [0, +\infty) \to A, \text{ measurable}\} \,,$$

$$\mathcal{B} = \{b(\cdot) : [0, +\infty) \to B, \text{ measurable}\} \,,$$

and A and B are given compact sets of $\mathbb{R}^m$. We will always assume that

$$\begin{cases} f : \mathbb{R}^{2n} \times A \times B \to \mathbb{R}^{2n} \text{ is continuous w.r. to all the variables and} \\ \text{there exists } L > 0 \text{ such that } |f(y_1, a, b) - f(y_2, a, b)| \leq L|y_1 - y_2| \\ \text{for all } y_1, y_2 \in \mathbb{R}^{2n}, \ a \in A, \ b \in B. \end{cases} \tag{2}$$

We will denote the solution of (1) by $y(t;x,a(\cdot),b(\cdot))$.
A target set $\mathcal{T} \subset \mathbb{R}^{2n}$ is given and it is assumed to be closed. The first player, called the *Pursuer* and denoted by P, wants to drive the system to $\mathcal{T}$. The second player, called the *Evader* and denoted by E, wants to drive the system away.

We deal with the natural extension of the minimum time problem, so we define the *payoff* of the game as the first (if any) time of arrival $T(x)$ on the target $\mathcal{T}$ for the trajectory solution of (1) starting at x. Note that, as usual, we set $T(x) = +\infty$ if the trajectory will not reach the target. The two players are opponents since the first player wants to minimize the payoff associated to the solution of the system whereas the second player wants to maximize it.

From now on, we restrict our analysis to Pursuit-Evasion games although some results are still valid in a more general context. We denote the coordinate of the space by $x = (x_P, x_E)$ where $x_P, x_E \in \mathbb{R}^n$. Each player can control only his own dynamics, *i.e.*, f has the form $f(x,a,b) = \big(f_P(x_P,a), f_E(x_E,b)\big)$. The state of the system is $y(t) = \big(y_P(t), y_E(t)\big)$ and a typical target has the form

$$\mathcal{T} = \{(x_P, x_E) \in \mathbb{R}^{2n} : |x_P - x_E| \le \varepsilon\}, \quad \varepsilon \ge 0$$

so in the unconstrained case the target is unbounded. As we said in the Introduction, we want to construct a numerical approximation for Pursuit-Evasion games *with state constraints*. This means that player P has to steer the system to the target satisfying the constraint $y_P(t) \in \overline{\Omega}_1$ for every t whereas player E must satisfy $y_E(t) \in \overline{\Omega}_2$ for every t, where Ω_1, Ω_2 are given bounded sets. The whole problem is set in $\overline{\Omega} \subset \mathbb{R}^{2n}$ where $\Omega := \Omega_1 \times \Omega_2$. Note that one player cannot force the other to respect or ignore the state constraints just because every player can affect only his dynamics and he is completely responsible for his strategy/trajectory. In the constrained game it is natural to replace $\mathcal{T}$ with $\mathcal{T} \cap \overline{\Omega}$.

The analysis of the continuous model with state constraints via dynamic programming techniques which is the basis for our approximation can be found in [8]. Let us start giving the time-discrete and the corresponding fully-discrete version of the differential game with state constraints. We will consider a discrete version of the dynamics based on the Euler scheme, namely

$$\begin{cases} y_{n+1} = y_n + hf(y_n, a_n, b_n) \\ y_0 = x\,, \end{cases}$$

where $h = \Delta t$ is a positive time step and we denote by $y(n;x,\{a_n\},\{b_n\})$ its solution at time nh corresponding to the initial condition $x = (x_P, x_E)$ and to the discrete strategies $\{a_n\},\{b_n\}$. The state constraints obviously require that $y(n;x,\{a_n\},\{b_n\}) \in \overline{\Omega}$ for all $n \in \mathbb{N}$.
We define the *constrained admissible strategies* for each player

$$\mathcal{A}_x := \{a(\cdot) \in \mathcal{A} : y_P(t;x,a(\cdot)) \in \overline{\Omega}_1\,,\ \text{for all } t \ge 0\}$$

$$\mathcal{B}_x := \{b(\cdot) \in \mathcal{B} : y_E(t;x,b(\cdot)) \in \overline{\Omega}_2\,,\ \text{for all } t \geq 0\}$$

and their time-discrete version

$$\mathcal{A}_x^h := \{\{a_n\} : a_n \in A \text{ and } y_P(n;x,\{a_n\}) \in \overline{\Omega}_1\,,\ \text{for all } n \in \mathbb{N}\}$$

$$\mathcal{B}_x^h := \{\{b_n\} : b_n \in B \text{ and } y_E(n;x,\{b_n\}) \in \overline{\Omega}_2\,,\ \text{for all } n \in \mathbb{N}\}.$$

Note that the constrained strategies now depend on x and on the state constraints. We will always assume that $\mathcal{A}_x \neq \emptyset$, $\mathcal{B}_x \neq \emptyset$, $\mathcal{A}_x^h \neq \emptyset$, and $\mathcal{B}_x^h \neq \emptyset$ for all $x \in \overline{\Omega}$ and h sufficiently small.

In the same way, we have to define the sets of admissible controls for every point $x \in \overline{\Omega}$. Let us start with the continuous problem. Following [20,21], we will select the subsets of admissible controls, denoted by $A(y)$ and $B(y)$, for every $y = (y_P, y_E) \in \overline{\Omega \setminus \mathcal{T}}$,

$$A(y) = \{a \in A : \exists\, r > 0 \text{ such that } y_P(t;y'_P,a) \in \overline{\Omega}_1 \text{ for } t \in [0,r] \text{ and } y'_P \in B(y_P,r) \cap \overline{\Omega}_1\}, \quad (3)$$

$$B(y) = \{b \in B : \exists\, r > 0 \text{ such that } y_E(t;y'_E,b) \in \overline{\Omega}_2 \text{ for } t \in [0,r],\ \text{and } y'_E \in B(y_E,r) \cap \overline{\Omega}_2\}. \quad (4)$$

For the time-discrete dynamics we define an analogue of subsets of $A(y)$ and $B(y)$ as follows:

$$A_h(y) := \{a \in A : y_P + hf_P(y_P,a) \in \overline{\Omega}_1\}\,, \qquad y \in \overline{\Omega \setminus \mathcal{T}}, \quad (5)$$

$$B_h(y) := \{b \in B : y_E + hf_E(y_E,b) \in \overline{\Omega}_2\}\,, \qquad y \in \overline{\Omega \setminus \mathcal{T}}. \quad (6)$$

The meaning of the above definitions is very clear: in order to guarantee that his trajectory satisfies his own state constraints over a time interval h, player P (respectively, player E) has to choose his control in $A_h(y)$ (respectively, $B_h(y)$). These subsets describe at every point $y \in \overline{\Omega \setminus \mathcal{T}}$ the “allowed directions" for each player, naturally they depend also on h, the dynamics and the constraints. Note that $A_h(y) \equiv A$ (respectively, $B_h(y) \equiv B$) in the unconstrained case.
We will also assume that

$$\exists\, h_0 > 0 \text{ s.t. } A_h(x) \neq \emptyset \text{ and } B_h(x) \neq \emptyset \text{ for all } (h,x) \in (0,h_0] \times \overline{\Omega}\,. \quad (7)$$

Definition 2.1. A *discrete strategy* for the first player is a map $\alpha_x : \mathcal{B}_x^h \to \mathcal{A}_x^h$. It is *nonanticipating* if $\alpha_x \in \Gamma_x^h$, where

$$\Gamma_x^h := \{\alpha_x : \mathcal{B}_x^h \to \mathcal{A}_x^h : b_n = \tilde{b}_n \text{ for all } n \leq n' \text{ implies } \alpha_x[\{b_k\}]_n = \alpha_x[\{\tilde{b}_k\}]_n \text{ for all } n \leq n'\}. \quad (8)$$

Let us define the reachable set as the set of all points from which the system can reach the target

$$\mathcal{R}^h := \{x \in \mathbb{R}^n : \text{ for all } \{b_n\} \in \mathcal{B}^h_x \text{ there exists } \alpha_x \in \Gamma^h_x \text{ and } \bar{n} \in \mathbb{N} \\ \text{such that } y(\bar{n}; x, \alpha_x[\{b_n\}], \{b_n\}) \in \mathcal{T}\}. \quad (9)$$

Then we define

$$n_h(x, \{a_n\}, \{b_n\}) := \begin{cases} \min\{n \in \mathbb{N} : y(n; x, \{a_n\}, \{b_n\}) \in \mathcal{T}\} & x \in \mathcal{R}^h \\ +\infty & x \notin \mathcal{R}^h \end{cases}.$$

We will consider for our approximation the discrete lower value of the game, which is

$$T_h(x) := \inf_{\alpha_x \in \Gamma^h_x} \sup_{\{b_n\} \in \mathcal{B}^h_x} h n_h(x, \alpha_x[\{b_n\}], \{b_n\})$$

and its Kružkov transform

$$v_h(x) := 1 - e^{-T_h(x)}, \quad x \in \overline{\Omega}. \quad (10)$$

Note that a similar construction can be done for the upper value of the game. The Dynamic Programming Principle (DPP) for Pursuit-Evasion games with state constraints is proved in [8] which also gives a characterization of the lower and upper value of the game in terms of the Isaacs equation. From the discrete version of the DPP (see [8]), we can conclude that the time-discrete value function v_h is the unique bounded solution of

$$\begin{cases} v_h(x) = \max\limits_{b \in B_h(x)} \min\limits_{a \in A_h(x)} \{\beta v_h(x + hf(x, a, b))\} + 1 - \beta & x \in \overline{\Omega} \setminus \mathcal{T} \\ v_h(x) = 0 & x \in \mathcal{T} \cap \overline{\Omega} \end{cases} \quad (\text{HJI}_h\text{–}\Omega)$$

where $\beta := e^{-h}$ and the $maxmin$ is obviously computed on the sets of admissible controls for the constrained game. In order to achieve the fully-discrete equation we build a regular triangulation of $\overline{\Omega}$ denoting by X the set of its nodes $x_i, i = 1, \dots, N$ and by $\mathcal{S}$ the set of simplices $S_j,\ j = 1, \dots, L$. $V(S_j)$ will denote the set of the vertices of a simplex S_j and the space discretization step will be denoted by k where $k := \max_j\{diam(S_j)\}$. Let us define $D \equiv (\overline{\Omega} \setminus \mathcal{T}) \cap X$.
The fully-discrete approximation scheme for the constrained case is

$$\begin{cases} v^k_h(x_i) = \max\limits_{b \in B_h(x_i)} \min\limits_{a \in A_h(x_i)} \left\{\beta v^k_h(x_i + hf(x_i, a, b))\right\} + 1 - \beta & x_i \in D \\ v^k_h(x_i) = 0 & x_i \in \mathcal{T} \cap X \\ v^k_h(x) = \sum_j \lambda_j(x) v^k_h(x_j),\ 0 \le \lambda_j(x) \le 1, \quad \sum_j \lambda_j(x) = 1 & x \in \overline{\Omega}. \end{cases} \quad (\text{HJI}^k_h\text{–}\Omega)$$

As in the unconstrained problem, the choice of linear interpolation is not an obligation and it was made here just to simplify the presentation. Let us denote by W^k

the set

$$W^k := \left\{ w \in C(\overline{\Omega}) : \nabla w(x) = \text{constant for all } x \in S_j \, , j = 1, \dots, L \right\} .$$

The proof of the following theorem can be obtained by simple adaptations of the standard proof for the free fully-discrete scheme (see, e.g., [6]).

Theorem 2.2. Equation (HJI$_h^k$–Ω) has a unique solution $v_h^k \in W^k$ such that $v_h^k : \overline{\Omega} \to [0,1]$.

Sketch of the proof. The right-hand side of the first equation is (HJI$_h^k$–Ω) defines a map $F : \mathbb{R}^N \to \mathbb{R}^N$ where N is the cardinality of the set of nodes in the triangulation. The proof relies on the fact that F is a contraction map so there exists a unique fixed point V^* and $v_h^k(x_i) = V_i^*$, $i = 1, \dots, N$. □

3 Convergence of the fully-discrete numerical scheme

The convergence of the fully-discrete scheme will be based on two ingredients. The first is an *a priori* bound for $v_h^k - v_h$ which will be obtained studying the properties of v_h on a family of approximate “reachable sets". This bound is proved for a *general dynamics* f and does not depend on the regularity of v. Then, we couple this bound with the convergence result in [8] where they prove that v_h converges to v uniformly in $\overline{\Omega \setminus \mathcal{T}}$.
Let us define $\mathcal{R}_0 := \mathcal{T}$ and

$$\mathcal{R}_n := \left\{ x \in \overline{\Omega} \setminus \bigcup_{j=0}^{n-1} \mathcal{R}_j : \text{ for all } b \in B_h(x) \text{ there exists } \hat{a}_x(b) \in A_h(x) \right.$$
$$\left. \text{such that } x + hf(x, \hat{a}_x(b), b) \in \mathcal{R}_{n-1} \right\}, \quad n \geq 1. \quad (11)$$

See [18] for an analogous definition in the framework of the minimum time problem.

Remark 3.1. By definition, the shape of the sets $\{\mathcal{R}_n\}_{n\in\mathbb{N}}$ strictly depends on f, Ω, A and B. Moreover, the following properties hold true:

1. $\mathcal{R}_n \cap \mathcal{R}_m = \emptyset$ for all $n \neq m$;
2. If $\mathcal{R}_p = \emptyset$ for some $p \in \mathbb{N}$, then $\mathcal{R}_q = \emptyset$ for any $q \geq p$;
3. The sets $\{\mathcal{R}_n\}_{n\in\mathbb{N}}$ are the level sets of v_h and v_h has jump discontinuities on the boundary of each of them.

In the sequel will always assume that

$$\overline{\Omega} = \bigcup_{j=0}^{\infty} \mathcal{R}_j . \quad (12)$$

Note that (12) can be interpreted as a sort of small time controllability assumption and that it is not really restrictive since if there exists a point $x \in \overline{\Omega} \backslash \bigcup_j \mathcal{R}_j$ this means that player P cannot win the game from that point (*i.e.*, he cannot drive the system to the target) and then $v_h(x) = 1$.

We introduce two important assumptions on the triangulation. Let $S \in \mathcal{S}$ be a simplex, the first assumption is (see Fig. 1):

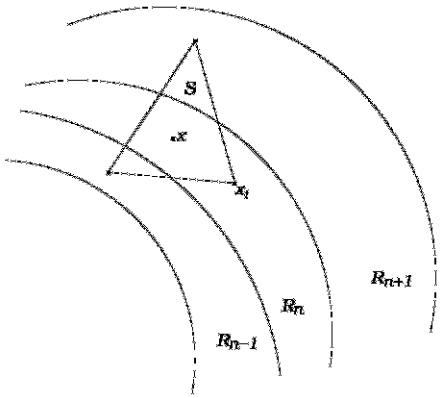

Figure 1: A simplex S crossing $\mathcal{R}_n$, $\mathcal{R}_{n-1}$, and $\mathcal{R}_{n+1}$.

$$x \in S \cap \mathcal{R}_n \Rightarrow V(S) \subset \mathcal{R}_{n-1} \cup \mathcal{R}_n \cup \mathcal{R}_{n+1}. \tag{13}$$

It means that the space discretization cannot be too coarse with respect to the time discretization. The second assumption is the ”consistency” of the triangulation (see Fig. 1).

Definition 3.2. We say that a triangulation is ”consistent” if $S \cap \mathcal{R}_n \neq \emptyset$ implies that there exists at least one vertex $x_i \in V(S)$ such that $x_i \in \mathcal{R}_n$.

The above assumption requires that every simplex of the triangulation cannot cross a level set $\mathcal{R}_n$ without having a vertex in it and, as we will see, is crucial for the convergence of the scheme. This condition will be always satisfied for k sufficiently small as we will see in Corollary 3.4. Let v_h and v_h^k denote, respectively, the solution of (HJI$_h$–Ω) and (HJI$_h^k$–Ω). We now state the main result of the paper.

Theorem 3.3. Let Ω be an open bounded set. Let (2), (12), (13) hold true and let the triangulation be ”consistent”. Then, for $n \geq 1$:

a) $v_h(x) \leq v_h(y)\,, \quad$ for any $x \in \bigcup_{j=0}^{n} \mathcal{R}_j\,, \quad$ for any $y \in \overline{\Omega} \backslash \bigcup_{j=0}^{n} \mathcal{R}_j$;

b) $v_h(x) = 1 - e^{-nh}$, for any $x \in \mathcal{R}_n$;

c) $v_h^k(x) = 1 - e^{-nh} + O(k) \sum_{j=0}^{n} e^{-jh}$ for any $x \in \mathcal{R}_n$;

d) There exists a constant $C > 0$ such that

$$|v_h(x) - v_h^k(x)| \leq \frac{Ck}{1 - e^{-h}} , \quad \text{for any } x \in \mathcal{R}_n.$$

Proof. $a)$ By induction. For $n = 0$ the statement is true since

$$0 = v_h(x) \leq v_h(y) \quad \text{for all } x \in \mathcal{T} , \quad \text{for all } y \in \overline{\Omega} \backslash \mathcal{T}.$$

Let the statement be true up to $n-1$. Suppose by contradiction that

$$\text{there exists } x \in \bigcup_{j=0}^{n} \mathcal{R}_j \text{ and } y \in \overline{\Omega} \backslash \bigcup_{j=0}^{n} \mathcal{R}_j \text{ such that } v_h(y) < v_h(x).$$

Therefore, there exists a (discrete) trajectory that starts from $\overline{\Omega} \backslash \bigcup_{j=0}^{n} \mathcal{R}_j$ and reaches the target in less then n time steps passing through $\mathcal{R}_n$. The contradiction follows by the definition of $\mathcal{R}_n$.

$b)$ By the definition of $\mathcal{R}_n$, for any $x \in \mathcal{R}_n$ we can find $n+1$ points $x^{(q)}$, $q = 0, \ldots, n$ such that $x^{(0)} = x$ and $x^{(q)} \in \mathcal{R}_{n-q}$. Introducing for simplicity the notations $a_q := a_{x^{(q)}}$ and $b_q := b_{x^{(q)}}$, we can write the sequence of the points $x^{(q)}$ more explicitly as

$$x^{(q+1)} = x^{(q)} + hf(x^{(q)}, \hat{a}_q(b_q^*), b_q^*) ,$$

where we use the * to indicate the optimal choice. As a consequence, the state of the system can reach the target in n steps and then $v_h(x) \leq 1 - e^{-nh}$. Suppose by contradiction that $v_h(x) < 1 - e^{-nh}$. As in $b)$, this means that the state has reached the target starting at x in less then n time steps but this is impossible since $x \in \mathcal{R}_n$.

$c)$ By construction we have $v_h^k(x_i) = 0$ for all $x_i \in \mathcal{R}_0 \cap X$. We now consider a generic point $x \in \mathcal{R}_0$ and let S be the simplex containing x. Since the triangulation is "consistent", S must have at least a vertex $x_{i_0} \in \mathcal{R}_0$ and then $v_h^k(x) = O(k)$ for all $x \in \mathcal{R}_0$ since $v_h^k \in W^k$. This implies, for all $x_i \in \mathcal{R}_1 \cap X$,

$$v_h^k(x_i) = \beta v_h^k(x_i + hf(x_i, a^*, b^*)) + 1 - \beta = \beta O(k) + 1 - \beta ,$$

since $x_i + hf(x_i, a^*, b^*) \in \mathcal{R}_0$. We now consider a generic point $x \in \mathcal{R}_1$. By the same arguments there exists at least one vertex $x_{i_1} \in \mathcal{R}_1$ such that

$$v_h^k(x) = v_h^k(x_{i_1}) + O(k) = \beta O(k) + 1 - \beta + O(k) = 1 - \beta + (1 + \beta)O(k).$$

For any $x_i \in \mathcal{R}_2 \cap X$,

$$v_h^k(x_i) = \beta(1 - \beta + (1 + \beta)O(k)) + 1 - \beta = 1 - \beta^2 + (\beta + \beta^2)O(k) ,$$

and, for any $x \in \mathcal{R}_2$ it exists $x_{i_2} \in \mathcal{R}_2$ such that

$$v_h^k(x) = v_h^k(x_{i_2}) + O(k) = 1 - \beta^2 + (1 + \beta + \beta^2)O(k).$$

Continuing by recursion we obtain, for any $x \in \mathcal{R}_n$

$$v_h^k(x) = 1 - \beta^n + O(k) \sum_{j=0}^{n} \beta^j.$$

$d)$ By $b)$ and $c)$ there exists a positive constant C_1 such that

$$|v_h(x) - v_h^k(x)| = C_1 k \sum_{j=0}^{n} \beta^j \leq \frac{C_1 k}{1 - \beta} = \frac{C_1 k}{1 - e^{-h}}.$$

□

Corollary 3.4. Let Ω be an open-bounded set. Let (2), (12) hold true. Moreover assume that

$$\min_{x,a,b} |f(x, a, b)| \geq f_0 > 0 \quad \text{and} \quad 0 < k \leq f_0 h. \tag{14}$$

Then, for $k \to 0^+$, v_h^k converges to v_h uniformly in $\overline{\Omega}$ for any $h > 0$ fixed.

Proof. First note that condition (14) is a sufficient condition for (13) and for the *consistency* of the triangulation. Therefore we can apply Theorem 3.3 and we easily conclude. □

Remark 3.5. The result in the above corollary does not rely on the fact that we use a split dynamic $f = (f_P, f_E)$ but we notice that it is not clear how to get the Hamilton-Jacobi-Isaacs equation associated to the problem in the case of a general dynamics with state constraints. In [21] there is an attempt in this direction but unfortunately the case considered there does not include Pursuit-Evasion games.

In order to obtain uniform convergence of v_h^k to the solution of the continuous problem when h and k tend to 0^+, we couple our result with those in [8] which are restricted to Pursuit-Evasion games. Let us denote by v the value function for the continuous problem as defined in [8]. In order to provide sufficient conditions for the continuity of v, we need to introduce further assumptions. Whenever we say that $\omega : [0, +\infty) \to [0, +\infty)$ is a *modulus* we mean that ω is nondecreasing, it is continuous at zero, and $\omega(0) = 0$. The first assumption is about the behavior of the value function v near the target $\mathcal{T}$.

$$\text{There is a modulus } \omega \text{ such that } v(x) \leq \omega(d(x, \mathcal{T})) \quad \text{for all } x \in \overline{\Omega} \tag{C1}$$

where $d(x, \mathcal{T}) = \inf_{z \in \mathcal{T}}\{|x - z|\}$.
The second is a small time controllability assumption for the Pursuer.

$$\begin{cases} \text{There is } \omega_P(\cdot, R) \text{ modulus for all } R > 0 \text{ such that for all} \\ w_1, w_2 \in \overline{\Omega}_1 \text{ there are } a(\cdot) \in \mathcal{A}_{w_1} \text{ and a time } t_{w_1,w_2} \text{ satisfying} \\ y_P(t_{w_1,w_2}; w_1, a(\cdot)) = w_2 \text{ and } 0 \le t_{w_1,w_2} \le \omega_P(|w_1 - w_2|, |w_2|). \end{cases} \tag{C2}$$

The third is a small time controllability assumption for the Evader.

$$\begin{cases} \text{There is } \omega_E(\cdot, R) \text{ modulus for all } R > 0 \text{ such that for all} \\ z_1, z_2 \in \overline{\Omega}_2 \text{ there are } b(\cdot) \in \mathcal{B}_{z_1} \text{ and a time } t_{z_1,z_2} \text{ satisfying} \\ y_E(t_{z_1,z_2}; z_1, b(\cdot)) = z_2 \text{ and } 0 \le t_{z_1,z_2} \le \omega_E(|z_1 - z_2|, |z_2|). \end{cases} \tag{C3}$$

The proof of the next theorem can be found in [8].

Theorem 3.6. Assume that (2), (C1), (C2), and (C3) hold. Then, the value function v is continuous in $\overline{\Omega}_1 \times \overline{\Omega}_2$.

Let us introduce the following regularity hypothesis on the boundary of $\mathcal{T}$.

$$\begin{cases} \text{For each } x \in \partial\mathcal{T} \text{ there are } r, \theta > 0 \text{ and } \Xi \in \mathbb{R}^{2n} \text{ such that} \\ \bigcup_{0<t<r} B(x' + t\Xi, t\theta) \subset \Omega\backslash\mathcal{T} \text{ for any } x' \in B(x, r) \cap \overline{\Omega\backslash\mathcal{T}}. \end{cases} \tag{15}$$

Note that a comparison principle for sub- and super-solutions for the same Hamiltonian is proved in [8] under additional regularity assumptions on $\partial(\Omega\backslash\mathcal{T})$. More precisely, the assumptions needed are the uniform interior cone conditions for Ω_1, Ω_2, and $\mathcal{T}$.
We have the following

Theorem 3.7. Let Ω be an open-bounded set. Let (2), (7), (C1), (C2), (C3), and (15) hold true. Finally, assume that

$$f_P(x_P, A(x)) \quad \text{and} \quad f_E(x_E, B(x)) \quad \text{are convex sets.} \tag{16}$$

Then, for $h \to 0^+$, v_h converges to v uniformly in $\overline{\Omega}$.

Proof. The assumption (16) guarantees that the value function v_h for the time-discrete problem defined in (10) coincides with that used in [8]. Moreover, assumptions of Theorem 3.6 are fulfilled so that $v \in C(\overline{\Omega})$. Then, the proof follows by Theorem 4.2 in [8]. □

Coupling the previous results we can prove our convergence result for the approximation of Pursuit-Evasion games.

Corollary 3.8. Let the assumptions of Corollary 3.4 and Theorem 3.7 hold true. Moreover, assume that $k = O(h^{1+\alpha})$, for $\alpha > 0$. Then v_h^k converges to v uniformly in $\overline{\Omega}$ for h tends to 0^+.

Proof. Since $1 - e^{-h} = O(h)$ for h tending to 0^+ we have for any $x \in \Omega$:

$$|v_h^k(x) - v(x)| \le |v_h^k(x) - v_h(x)| + |v_h(x) - v(x)| \le O(h^\alpha) + \|v_h(x) - v(x)\|_\infty .$$

□

As we said in the Introduction, a convergence theorem has been proved in [11,12] for a different approximation scheme based on viability theory. The approach is different in several respects. The first is that the techniques used in the proof are based on the characterization of the epigraph of the value function of the game in terms of a *Discriminating Kernel* for a suitable problem. By this technique the authors can easily deal with semicontinuous Hamiltonians and construct a discrete Discriminating Kernel algorithm. This technique is based on an external approximation of the epigraph of the value function via a sequence of closed sets D_p, $p \in \mathbb{N}$ (see [13] p. 224). This construction is rather expensive for games and can be hard to pursue particularly in high-dimension even if one can try to localize the algorithm near the boundary of the Discriminating Kernel (*i.e.,* nearby the graph of the value function).

4 The Tag-Chase game with state constraints

The Tag-Chase game is a particular case of Pursuit-Evasion games. We consider two boys, P and E, which run one after the other in the same 2-dimensional domain, so that the game is set in $\overline{\Omega} = \overline{\Omega}_1^2 \subset \mathbb{R}^4$ where Ω_1 is an open-bounded set of $\mathbb{R}^2$. We denote by (x_P, x_E) the coordinates of $\overline{\Omega}$ where $x_P,\ x_E \in \overline{\Omega}_1$. P (respectively, E) can run in every direction with velocity V_P (respectively, V_E) so that the dynamics of the game is

$$\begin{cases} \dot{x}_P = V_P\, a & a \in B_2(0,1) \\ \dot{x}_E = V_E\, b & b \in B_2(0,1)\ , \end{cases}$$

where $B_2(0,1) = \{z \in \mathbb{R}^2 : |z| \le 1\}$. The case $V_P > V_E$ is completely studied in [1,2]. The value function $T = -\ln(1-v)$, which represents the capture time, is continuous and bounded in its domain of definition. Moreover, the convergence result we obtained in Sec. 3 applies to this case.

On the other hand, the most interesting case is certainly $V_P = V_E$, *i.e.,* when the players have the same dynamics and no advantage is given to any of them. In this case it is easily seen that the value function T is discontinuous (at least on $\partial\mathcal{T}$) and then all theoretical results based on the continuity of the value function does not hold.

In this section we will give an answer to the following question: ”if $V_P = V_E$, is the capture time finite?”.

If the Tag-Chase game is played without constraints on the state and both players play optimally, it is immediately seen that the distance between P and E remains constant and then capture never happens (the optimal strategy for E is to move for ever as fast as he can along the line joining P and E in the opposite direction with respect to the position of P). On the contrary, if the state is constrained in a bounded domain, such a restriction seems to play a key role against the Evader, as the following proposition shows.

Proposition 4.1. Let Ω_1 be open and bounded. Moreover, let the target be

$$\mathcal{T} = \{(x_P, x_E) \in \overline{\Omega} : \ |x_P - x_E| \leq \varepsilon\}, \quad \varepsilon \geq 0. \tag{17}$$

Then,

1. If $V_P > V_E$, then the capture time $t_c = T(x_P, x_E) = -\ln(1 - v(x_P, x_E))$ is finite and bounded by
$$t_c \leq \frac{|x_P - x_E|}{V_P - V_E}.$$
2. If $V_P = V_E$, $\varepsilon \neq 0$ and Ω_1 is convex then the capture time t_c is finite.

Proof.

1. This first part of the proof can be found in [1]. We fix a strategy for P and leave E free to decide his optimal strategy. First, P reaches the starting point of E covering the distance $|x_P - x_E|$ and then he follows the E's trace. The conclusion follows by elementary computations.
2. The basic idea of the proof is the same of the previous case but we have to change the strategy for the Pursuer in order to have a finite upper bound. P runs after E always along the line joining P and E (P can do it by the convexity of Ω_1) while E chooses his own optimal trajectory as before. We can characterize the strategy of E by a function $\theta(t) : [0, +\infty) \to [0, 2\pi)$ which represents at every time the smallest angle between the velocity vector of E and the line joining P and E (see Fig. 2). Let us denote by $d_{PE}(t)$ the

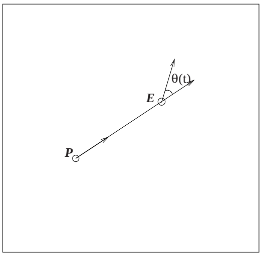

Figure 2: Trajectories of P and E in Proof of Proposition 4.1.

distance between P and E at time t. We claim that, for any fixed t,

$$\theta(t) \neq 0 \Rightarrow d'_{PE}(t) < 0 \tag{18}$$

where $' = \frac{d}{dt}$. Due to the state constraints, $\theta(t)$ cannot be equal to 0 for a time interval longer then $diag(\Omega_1)/V_P$ and after that must be different from 0 for a finite time interval because E must change his trajectory at least when he touches $\partial\Omega_1$. Therefore, if (18) holds then $d_{PE}(t) \to 0$ for $t \to \infty$ and then for any $\varepsilon > 0$ there exists a time $\bar{t}$ such that $d_{PE}(\bar{t}) \leq \varepsilon$ (the capture occurs). In order to prove (18), let us define the two vectors $E(t)$ and $P(t)$ which are, respectively, the position of P and E at time t and the vector $r(t) := E(t) - P(t)$. By definition, we have $d_{PE}(t) = |r(t)|$. Without loss of generality, suppose that at time t, $P(t)$ is in the origin and $E(t)$ lies on the x-axis and

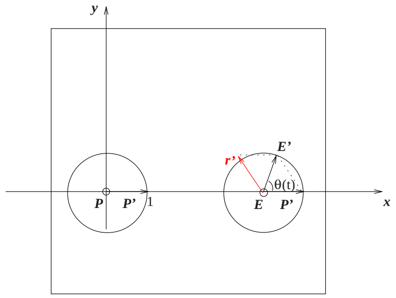

Figure 3: Vectors P, P', E, E', and r' as in Proposition 4.1.

that $V_P = V_E = 1$ (see Fig. 3).
Then

$$P'(t) = (1, 0) \quad \text{and} \quad \frac{r(t)}{|r(t)|} = (1, 0).$$

Moreover, by construction we have

$$E'(t) = (\cos\theta(t), \sin\theta(t)) \quad \text{and} \quad r'(t) = E'(t) - P'(t).$$

It follows that $r'(t) = (\cos\theta(t) - 1, \sin\theta(t))$ and

$$d'_{PE}(t) = \frac{r(t)}{|r(t)|} \cdot r'(t) = \cos\theta(t) - 1 \tag{19}$$

so that (18) holds. □

5 Some hints for the algorithm

In this section we give some hints for an efficient implementation of the algorithm for the solution of differential games. The main goal is to reduce the computational cost since this is a crucial step toward applications. The first hint deals with a fast way to compute the term $v_h^k(x_i + hf(x_i, a, b))$ in high-dimensional spaces.

In fact, the linear interpolation used in the definition of the fixed point iteration is appealing from the theoretical point of view but not very efficient since it would require the solution of a linear system of size $2n+1$ for every x_i, a, and b. The procedure we suggest solves this problem by a sequence of linear interpolations in 1-dimension. The second hint exploits the natural symmetry in the game problem whenever it is played in a square domain in order to reduce (by a factor 2 in two dimensions and 4 in four dimensions) the domain where the solution is actually computed. Both procedures have shown to be very efficient and have contributed to a dramatic reduction of the CPU time.

5.1 Interpolation in high-dimensions

It is important to note that the semi-Lagrangian scheme (HJI_h^k–Ω) requires that at every iteration, at every node and every a and b, the value $v_h^k(x_i + hf(x_i, a, b))$ is computed and this computation needs an interpolation of the values of v_h^k at the nodes. [14] extensively analyzed a fast and efficient interpolation method in high-dimension suitable to our purposes. We recall briefly this method giving a precise error estimate.

Consider a point $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ and the cell of the grid which contains it (see Fig. 4 for an example in 3D). Suppose that a function f is known in the 2^n vertexes of the cell and we want to compute the value $f(x)$ by linear interpolation. The basic idea is to project the point x onto lower and lower dimensional subspaces until dimension 1 is reached. More precisely, choose a dimension (in Fig. 4 we chose x_1) and project the point x in that dimension on both sides of the cell finding the points P_1^1 and P_2^1. Then, choose a direction different from the first one (we chose x_2) and project the points P_1^1 and P_2^1 on the sides of the cell finding the points P_1^2, P_2^2, P_3^2, and P_4^2. Iterate the projection procedure $2^{n+1}-2$ times in the same way until all vertexes of the cell are reached. At this stage a tree structure containing all points $P_i^j\,, i = 1, \dots, 2^n\,,\; j = 1, \dots, n$ is computed from top to bottom. Now evaluate by *unidimensional* linear interpolations the values of f at the points $P_i^j\,, i = 1, \dots, 2^n\,, \quad j = 1, \dots, n$ in the reverse order with respect to that used to find them (from bottom to top). This procedure leads to an approximate value of $f(x)$ obtained by $2^n - 1$ unidimensional linear interpolations. It is interesting to give a precise error estimates of this first-order interpolation method.

Theorem 5.1. Let $\mathbb{R}^n \supset Q := [a_1, b_1] \times \dots \times [a_n, b_n]$ and $x = (x_1, \dots, x_n)$. Assume $f \in C^2(Q; \mathbb{R})$ and let $q(x)$, $x \in Q$ be the approximate value of $f(x)$ obtained by the n-dimensional linear interpolation described above.
Then, the error $E(x) := f(x) - q(x)$ is bounded by

$$|E(x)| \leq \sum_{i=1}^{n} \frac{\Delta_i^2}{8} M_i\,, \qquad \text{for all } x \in Q\,,$$

where $M_i = \max\limits_{x \in Q} |\frac{\partial^2 f(x)}{\partial x_i^2}|$ and $\Delta_i = b_i - a_i$.

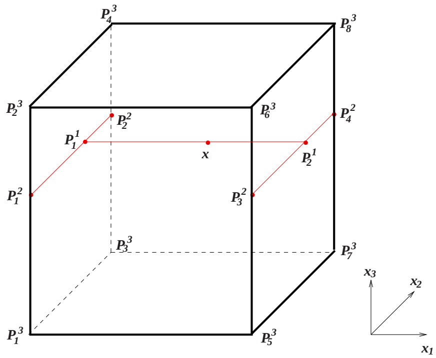

Figure 4: Example in 3D.

Proof. The proof is easily obtained by induction using the basic theory of linear interpolation. The interested reader can find the complete proof in [15]. □

5.2 Reducing the size of the problem

If the dimension of the space is greater than 4, the algorithm has a high computational cost. As already noted in [2,22], due to the state constraints it is not possible in general to use reduced coordinates $\tilde{x} = x_P - x_E$ unless the problem has a special structure. In fact, using reduced coordinates we loose every information about the real positions of the two players, so that we cannot detect when they touch the boundary of the domain (and then change the dynamics consequently). Note that if we consider the 2-player Tag-Chase game constrained in a circle (see for example [9]), the problem can be described by three coordinates instead of four since the problem is invariant with respect to the rotation of the domain. Obviously, this is not true if the game's field is a square as in the case of the numerical tests we carry on in this paper.

Although it is not always possible to describe the game in a reduced space due to the state constraints, we can simplify the computation taking into account the symmetries of the problem, if any. We explain our technique first in the 1-dimensional case and then in the 2-dimensional case. From now on we denote by n the number of grid nodes in each dimension.

Unidimensional Tag-Chase game

Assume that each player can move along a line in the interval $[-x_0, x_0]^2$, then the game is set in the square $[-x_0, x_0]^2$.

In Fig. 5 we show the level sets of the solution $T = -\log(1 - v)$ in the case $V_P = 2$, $V_E = 1$, $x_0 = 2$ and an optimal trajectory starting from $(-1.5, 0)$. It is

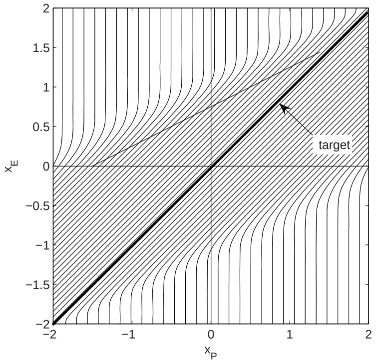

Figure 5: Level sets of the solution $T = -\log(1 - v)$ for $V_P = 2$, $V_E = 1$.

easy to see that

$$v(x_P, x_E) = v(-x_P, -x_E) \quad \text{for all } x_P, x_E \in [-x_0, x_0]$$

so that we can recover the entire solution either from the triangular sector $S_{NW} = \{(x_P, x_E) : x_P \leq x_E\}$ or from the rectangular sector $S_W = \{(x_P, x_E) : x_P \leq 0\}$. This corresponds to the fact that it is sufficient to compute the solution for all the initial positions of P and E in which P is on the left of E or P is in the left side of the domain $[-x_0, x_0]$ (see Fig. 6). There is an important difference

Figure 6: Two initial positions which correspond to the same value for v.

between the two approaches. In fact, the target $\mathcal{T} = \{(x_P, x_E) : x_P = x_E\}$ is entirely contained in S_{NW} but not in S_W. Moreover, since the target divides the domain $\overline{\Omega} = [-x_0, x_0]^2$ in two parts and no characteristics can pass from one part to the other, all the optimal trajectories starting from S_{NW} remains in S_{NW}. This is clearly not true for S_W. As a consequence, if we compute the solution only in S_W this will be not correct because not all the usable part of the target is visible from the domain.

Unfortunately, the domain S_{NW} has not a correspondence in the two-dimensional Tag-Chase game. In fact, the target $\mathcal{T}$ does not divide the entire space $\overline{\Omega} = ([-x_0, x_0] \times [-x_0, x_0])^2$ into two parts since the co-dimension of the target is strictly greater than 1. On the contrary, we will see that the domain S_W has a natural generalization in the two-dimensional case.

For this reason it is preferable to localize the computation only in S_W. In order to do this we adopt the following idea. First of all we choose n even. Then we compute the approximation of v at the node corresponding to the indices (i,j), for $i = 1,\dots,n/2, j = 1,\dots,n$ via the numerical scheme (HJI$_h^k$–Ω) (note that now i is the index corresponding to the position of the player P so is a column index whereas j is a row index). After every iteration we copy the line $(i = n/2,\ j = 1 : n)$ in $(i = n/2+1,\ j = n : 1)$ as a sort of "periodic boundary condition" for S_W. In this way the information coming from the south-western part of the target can substitute the missing information needed by the north-western part of the domain.

When the algorithm reached the convergence we can easily recover the solution on all over the domain Ω.

Two-dimensional Tag-Chase game

As we did in the unidimensional case, we want to use the symmetries of the problem to avoid useless computation.

We assume that each player can move in a square so that the game is set in a four-dimensional hypercube. The positions of P and E will be denoted, respectively, by (x_P, y_P) and (x_E, y_E). In this case we have more than one symmetry. In fact, it easy to check that the following three inequalities hold (see Fig. 7):

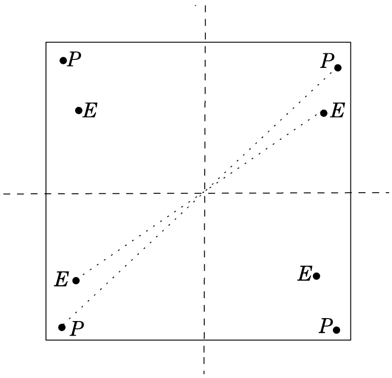

Figure 7: Four initial positions which correspond to the same value for v.

$$v(x_P, y_P, x_E, y_E) = v(-x_P, -y_P, -x_E, -y_E) \tag{20}$$

$$v(x_P, y_P, x_E, y_E) = v(-x_P, y_P, -x_E, y_E) \tag{21}$$

$$v(x_P, y_P, x_E, y_E) = v(x_P, -y_P, x_E, -y_E). \tag{22}$$

We note that once we take into account the symmetry (20) we took into account automatically the symmetry in (21).

The following nested `for`'s take into account only the symmetry (22) and they allow to compute correctly the whole 4D matrix containing the grid nodes.

```
for i=1:n
  for j=1:n/2
    for k=1:n
      for l=1:n
        {v(i,j,k,l)=SLscheme(...);
        v(i,n-j+1,k,n-l+1)=v(i,j,k,l);}
```

Now we are ready to make use of symmetry (20) by means of the technique introduce for the unidimensional Tag-Chase game. We compute just half matrix corresponding to the indexes $i = 1, \dots, n/2$ and after every iteration we copy the submatrix $(i = n/2,\ j = 1 : n,\ k = 1 : n,\ l = 1 : n)$ in $(i = n/2 + 1,\ j = n : 1,\ k = n : 1,\ l = n : 1)$ as a boundary condition.

At the end of computation we easily recover the solution in the whole domain.

Remark 5.2. We ran a Fast Sweeping [23] version of the one-dimensional Tag-Chase game. We noticed that no improvements about the number of iterations is achieved. This is probably due to the presence of state constraints so that the information first propagates from the target and then it comes back after hitting the boundary. A Fast Marching scheme for the unconstrained game in reduced coordinates has been presented in [16].

6 Numerical experiments

In this section we present some numerical experiments for two-dimensional constrained Tag-Chase game. We consider the case $V_P > V_E$ as well as $V_P = V_E$ and $V_P < V_E$. To our knowledge, these two last cases appear for the first time in a numerical test. The code is written in C++ and its parallel version has been obtained by means of OpenMP directives. The algorithm ran on a PC equipped with a processor Intel Pentium dual core 2 $\times$ 2.80 GHz, 1 GB RAM and on an IBM system p5 575 equipped with 8 processors Power5 at 1.9 GHz and 32 GB RAM located at CASPUR[1].

Notations and choice of parameters

We denote by n the number of nodes in each dimension. We denote by n_c the number of admissible directions/controls for each player, *i.e.*, we discretize the unit ball $B(0, 1)$ with n_c points. We restrict the discretization to the boundary $\partial B(0, 1)$ and in some cases we add the central point (in this case we denote the number of directions by n_c^-+1 where $n_c^- = n_c - 1$).

[1]Consorzio interuniversitario per le Applicazioni di Supercalcolo per Università e Ricerca, `www.caspur.it`.

We always use a uniform structured grid with four-dimensional cells of volume Δx^4 and we choose the (fictitious) time step h such that $\|hf(x,a,b)\| \leq \Delta x$ for all x, a, b (so that the interpolation is made in the neighboring cells of the considered point).

We introduce the following stopping criterion for the fixed point iteration $V^{p+1} = F(V^p)$ (where $V_i = v_h^k(x_i)$)

$$\|V^{(p+1)} - V^{(p)}\|_\infty \leq \varepsilon \,, \qquad \varepsilon > 0.$$

We remark that the quality of the approximate solution depends on h, k and also (strictly) on the ratio h/k (see [4])

The real game is played in a square $[-2,2]^2$ so the problem is set in $\overline{\Omega} = [-2,2]^4$. The numerical target is $\mathcal{T} = \{(i,j,k,l) \in \{1,\ldots,n\}^4 : |i-k| \leq 1 \text{ and } |j-l| \leq 1\}$.

Once we computed the approximate solution, we recover the optimal trajectories. At this stage we have to choose the time step Δt in order to discretize the dynamical system by Euler scheme. It should be noted that this parameter can, in general, be different from the (fictitious) time step h chosen for the computation of the value function (our choice is $\Delta t = \Delta x/2$) and this is true also for the number of controls n_c. Moreover, computing optimal trajectory requires the evaluation of the argminmax which is done again choosing a value for h, and this value can be in principle different from that used in the first computation.

We plot some flags (circles for the Pursuer, squares for the Evader) on the approximate optimal trajectories every s time steps where s varies from 5 to 20 depending on the test. This allows to follow the position of one player with respect to the other during the game.

We denote by $v(x_P, y_P, x_E, y_E)$ the approximate value function and by $T(x_P, y_P, x_E, y_E) = -\ln(1 - v(x_P, y_P, x_E, y_E))$ the time of capture.

6.1 Case $V_P > V_E$

The case $V_P > V_E$ is the classical one and it was already studied by Alziary de Roquefort [2]. In this case, the value function v is continuous and all theoretical results we presented in this paper hold true. In the following we name "CPU time" the sum of the times taken by the CPUs and by "wallclock time" the elapsed time.

Test 1
We choose $\varepsilon = 10^{-3}$, $V_P = 2$, $V_E = 1$, $n = 50$, $n_c = 48 + 1$. Convergence was reached in 85 iterations. The CPU time (IBM - 8 procs) was 17h 36m 16s, the wallclock time was 2h 47m 37s. Figure 8 shows the value function $T(0, 0, x_E, y_E)$ and its level sets (we fix the Pursuer's position at the origin). It is immediately seen that if the distance between P and E is greater than $V_P - V_E = 1$ then the state constraints have a great influence on the solution. Moreover, it is clear that

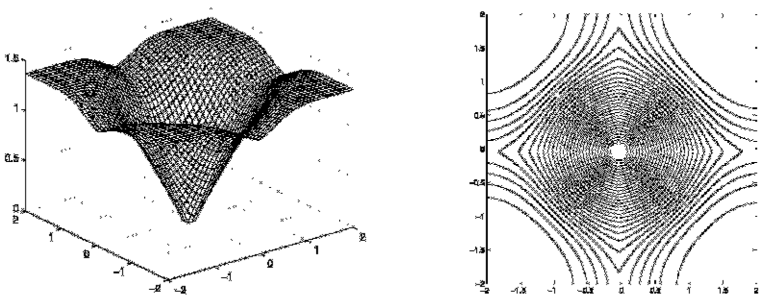

Figure 8: Test 1. Value function $T(0, 0, x_E, y_E)$ (left) and its level sets (right).

the presence of state constraints gives an advantage to the Pursuer.
Figure 9 shows four optimal trajectories corresponding to the starting points:

$$\begin{cases} P = (-1, 0) \\ E = (0, 0) \end{cases} \begin{cases} P = (-2, -2) \\ E = (1, 0.7) \end{cases} \begin{cases} P = (-1.8, -1.8) \\ E = (0.5, -1.6) \end{cases} \begin{cases} P = (-1.8, -1.8) \\ E = (0.5, -1.8). \end{cases}$$

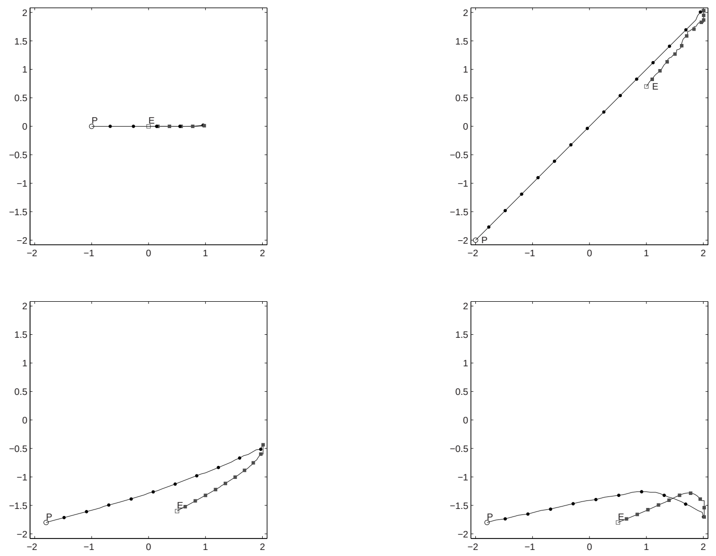

Figure 9: Optimal trajectories for Test 1.

Test 2
The second test is just to compare the CPU time corresponding to the two architectures mentioned above. It is interesting to test the new dual core processors in

order to understand how much they can be useful in parallel scientific computing. They are indeed conceived mainly to deal with distributed computing or simply multitasking. The performances of the parallel code are measured in terms of two well-known parameters, the *speed-up* and the *efficiency*. Let T_{ser} and T_{par} be the times corresponding, respectively, to the execution of the serial and parallel algorithms over n_p processors. We define

$$\textit{speed-up} := \frac{T_{ser}}{T_{par}} \qquad \text{and} \qquad \textit{efficiency} := \frac{\textit{speed-up}}{n_p}.$$

Note that an ideal parallel algorithm would have *speed-up* $= n_p$ and *efficiency* $= 1$. Table 1 shows the wallclock time, the *speed-up* and the *efficiency* for the following

Table 1: CPU time for Test 2

architecture	wallclock time	speed–up	efficiency
IBM serial	26m 47s	-	-
IBM 2 procs	14m 19s	1.87	0.93
IBM 4 procs	8m 09s	3.29	0.82
IBM 8 procs	4m 09s	6.45	0.81
PC dual core, serial	1h 08m 44s	-	-
PC dual core, parallel	34m 51s	1.97	0.99

choice of parameters: $\varepsilon = 10^{-5}$, $V_P = 2$, $V_E = 1$, $n = 26$, $n_c = 36 + 1$.

Test 3

In this test the domain has a square hole in the center. The side of the square is 1.06. We choose $\varepsilon = 10^{-4}$, $V_P = 2$, $V_E = 1$, $n = 50$, $n_c = 48 + 1$. Convergence was reached in 109 iterations. The CPU time (IBM - 8 procs) was 1d 00h 34m 18s, the wallclock time was 3h 54m 30s. Figure 10 shows the value function

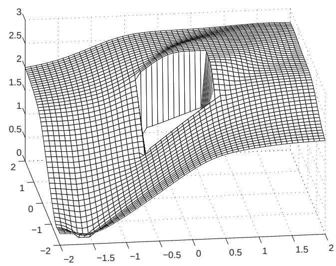

Figure 10: Test 3. Value function $T(-1.5, -1.5, x_E, y_E)$.

$T(-1.5, -1.5, x_E, y_E)$.
Figure 11 shows two optimal trajectories corresponding to the starting points:

$$\begin{cases} P = (-1.9, -1.9) \\ E = (1.9, 1.9) \end{cases} \qquad \begin{cases} P = (-1.9, 0) \\ E = (1, 0). \end{cases}$$

It is interesting to note that in both cases the Evader waits until the Pursuer decides

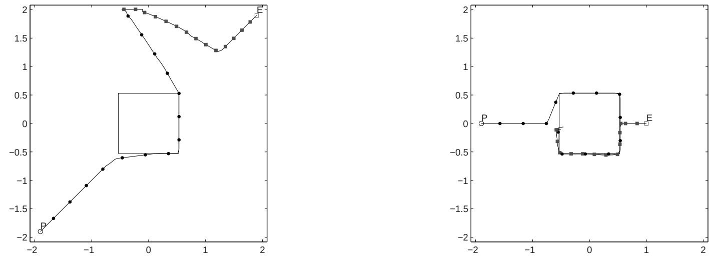

Figure 11: Optimal trajectories for Test 3.

if he wants to skirt around the obstacle clockwise or counterclockwise. After that, the Evader goes in the opposite direction. If both players touch the obstacle, they run around it until the capture occurs.

Test 4
In this test the domain has a circular hole in the center. The radius r of the circle is $7\Delta x$. We choose $\varepsilon = 10^{-4}$, $V_P = 2$, $V_E = 1$, $n = 50$, $n_c = 48 + 1$. Convergence was reached in 108 iterations. The CPU time (IBM - 8 procs) was 1d 17h 27m 43s, the wallclock time was 6h 39m 00s. Note that handling with a circular obstacle inside the domain of computation is not easy as in the previous test where the boundary of the obstacle matches with the lattice. We adopt the following procedure. First of all, we define the radius r of the circle as a multiple of the space step Δx. Then, at every node $(P{=}(i, j), E{=}(k, l))$, we compute the distance d_{PO} (resp., d_{EO}) between P (resp., E) and the center of the domain. Let us focus on P, E being treated in the same way. If $r \leq d_{PO} < r + \Delta x$, then we say that P in on the "numerical boundary" of the circle. The exterior normal vector $\eta(i, j)$ to the (numerical) boundary of the circle is simply given by the coordinates of the node (i, j), so that we can easily compute the scalar product $\eta \cdot a$ where a is the desired direction of P. If the scalar product is negative, we label the direction a as *not admissible*.

Figure 12 shows two optimal trajectories corresponding to the starting points:

$$\begin{cases} P = (-1.9, -1.9) \\ E = (1.9, 1.9) \end{cases} \qquad \begin{cases} P = (-0.6, 0) \\ E = (1, 0.4). \end{cases}$$

The behavior of the optimal trajectories is similar to the previous Test.

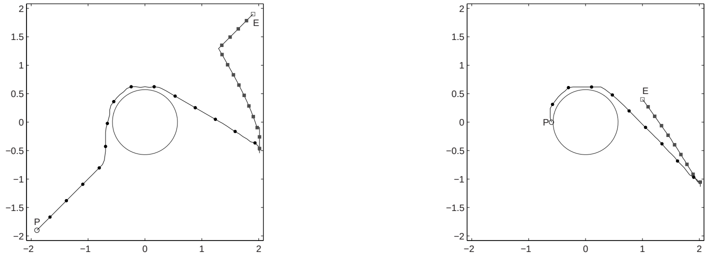

Figure 12: Optimal trajectories for Test 4.

6.2 Case $V_P = V_E$

When $V_P = V_E$ the value function v is discontinuous on $\partial\mathcal{T}$. In this case no convergence results are known, nevertheless the numerical scheme seems to work very well. We remember that results in Sec. 4 guarantee that $v < 1$ (the capture always occurs). This is confirmed by the following test.

Test 5

We choose $\varepsilon = 10^{-3}$, $V_P = 1$, $V_E = 1$, $n = 50$, $n_c = 36$. Convergence was reached in 66 iterations. Figure 13 shows the value function $T(0, 0, x_E, y_E)$ and its level sets. Figure 14 shows the value function $T(1.15, 1.15, x_E, y_E)$ and its

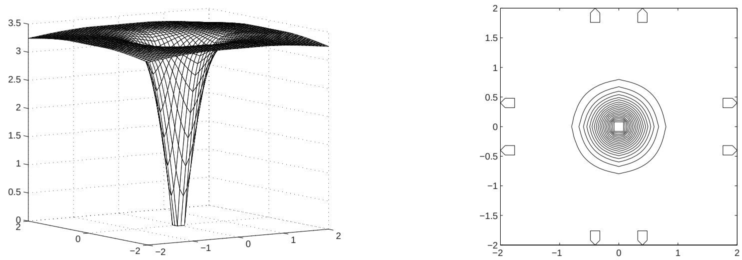

Figure 13: Test 5. Value function $T(0, 0, x_E, y_E)$ (left) and its level sets (right).

level sets. Figure 15 shows four optimal trajectories corresponding to the starting points:

$$\begin{cases} P = (0,1) \\ E = (0,0) \end{cases} \quad \begin{cases} P = (1,1.5) \\ E = (-0.5,0) \end{cases} \quad \begin{cases} P = (1.3,1.8) \\ E = (0,0) \end{cases} \quad \begin{cases} P = (-1.9,-1.9) \\ E = (-1.7,-1.9). \end{cases}$$

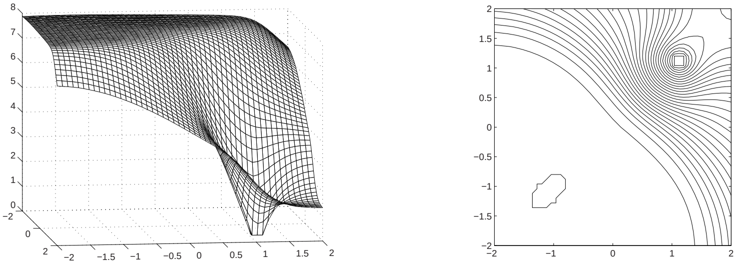

Figure 14: Test 5. Value function $T(1.15, 1.15, x_E, y_E)$ (left) and its level sets (right).

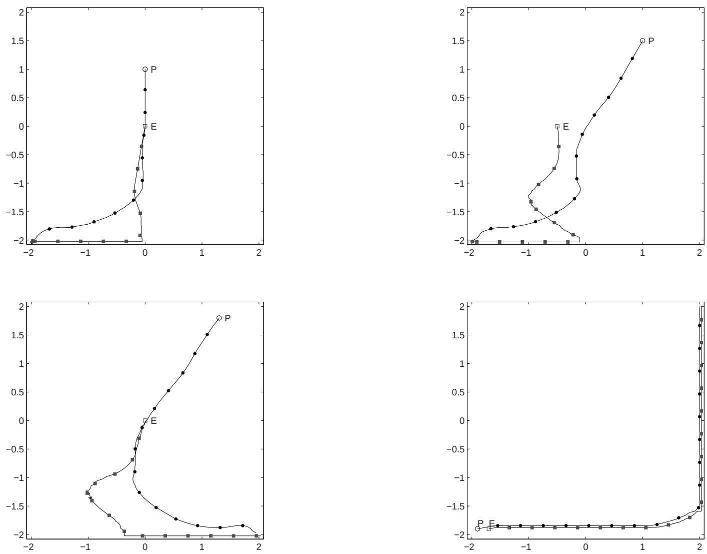

Figure 15: Optimal trajectories for Test 5.

Test 6

In this test the domain has a circular hole in the center. The radius of the circle is $7\Delta x$. Since the domain is no more convex, we have no guarantee that the time of capture is finite. Numerical results show that the value function v is equal to 1 in a large part of the domain.

It is well known that it is not possible to recover the optimal trajectories whenever $v = 1$ ($T = \infty$) since from that regions capture never happens. Indeed, if $V_P \leq V_E$ the approximate solution shows a strange behavior. Even if $v < 1$, in some cases the computed optimal trajectories tend to stable trajectories such that P never reaches E. Although this is due to some numerical error, these trajectories are extremely realistic so they give to us a guess about the optimal strategies of the players in the case E wins. In this Test (and others below) we show this behavior.

We choose $\varepsilon = 10^{-4}$, $V_P = 1$, $V_E = 1$, $n = 50$, $n_c = 48 + 1$. Convergence was reached in 94 iterations. The CPU time (IBM - 8 procs) was 1d 12h 05m 22s. Figure 16 shows one optimal trajectory corresponding to the starting point

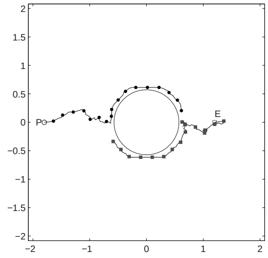

Figure 16: Optimal trajectories for Test 6.

$$\begin{cases} P = (-1.8, 0) \\ E = (1.2, 0). \end{cases}$$

In this example, there is no capture within 150 time steps. The asymptotic behavior of the trajectory is stable since once the two players reached the internal circle, they run around it forever. It should be noted that, at the beginning of the game, E leaves the time to go by in order to touch the boundary of the circle exactly when P touches it.

This strange behavior urges us to invent some method to compute rigorously the trajectories corresponding to the E's win, in order to confirm our guess. Maybe we can do it considering the time-dependent problem (so that we work in $\mathbb{R}^5$ as Alziary de Roquefort does [2]). This allows one to choose a time-dependent velocity $V_E(t)$ such that it is very fast for $0 \leq t < \bar{t}$ (capture impossible) and very slow for $t > \bar{t}$ (capture unavoidable). For such a velocity we have $v < 1$ so we can compute optimal trajectories but, for $0 \leq t < \bar{t}$, E will attempt to maintain a trajectory such that capture does not occur.

6.3 Case $V_P < V_E$

If $V_P < V_E$ the value function v is discontinuous on $\partial\mathcal{T}$. Moreover, we have no guarantee that the time of capture is finite. Numerical results show that the value function v is equal to 1 in a large part of the domain.

Test 7

We choose $\varepsilon = 10^{-3}$, $V_P = 1$, $V_E = 1.25$, $n = 50$, $n_c = 48 + 1$. Convergence was reached in 53 iterations. The CPU time (IBM - 8 procs) was 12h 43m 02s, the wallclock time was 2h 18h 06s.

Figure 17 shows the value function $T(-1, -1, x_E, y_E)$ and its level sets.

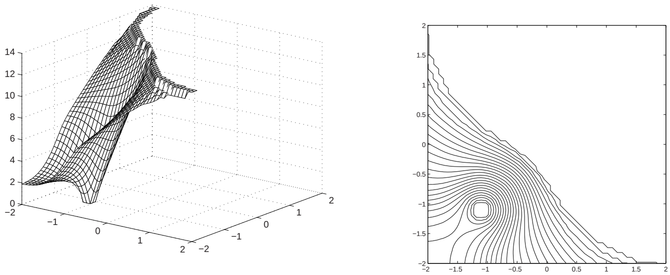

Figure 17: Test 7. Value function $T(-1, -1, x_E, y_E)$ (left) and its level sets (right).

Figure 18 shows two optimal trajectories corresponding to the starting points

$$\begin{cases} P = (-1, -1) \\ E = (-1, 1) \end{cases} \qquad \begin{cases} P = (-1, -1) \\ E = (-0.5, -0.5). \end{cases}$$

Note that the Pursuer approaches the corner in which capture occurs along the

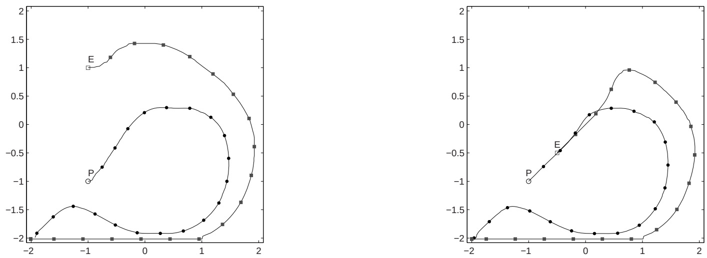

Figure 18: Optimal trajectories for Test 7.

diagonal of the square in order to block off the Evader's escape.

Test 8

We choose $\varepsilon = 10^{-4}$, $V_P = 1$, $V_E = 1.5$, $n = 50$, $n_c = 36$. Convergence was

reached in 65 iterations. The CPU time (IBM - 8 procs) was 15h 48m 46s, the wallclock time was 2h 30m 19s.
Figure 19 shows two optimal trajectories corresponding to the starting points:

$$\begin{cases} P = (0.5, 0.5) \\ E = (1.5, 1.5) \end{cases} \qquad \begin{cases} P = (0, -0.8) \\ E = (-0.3, -1.3). \end{cases}$$

In the example on the left, E makes believe he wants to be caught in the upper-left

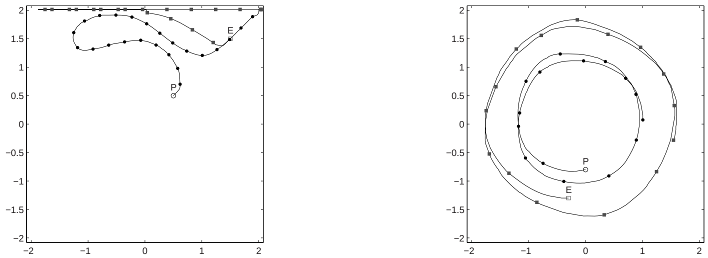

Figure 19: Optimal trajectories for Test 8.

corner but after a while he turns on the right toward the upper-right corner. In the example on the right, there is no capture within 2,000 time steps (see Test 6) and the asymptotic behavior of the trajectories is quite stable. Moreover, we note that the ratio between the two radii of the circles are about 1.5 as the ratio between the velocities of the two players (so that they complete a rotation in the same time).

Acknowledgements

The authors wish to thank D. Graziani, S. Cacace, and E. Cotognini for helpful discussions on the Tag-Chase game.

This work has been partially supported by funds from the MIUR Project 2005 “Modellistica Numerica per il Calcolo Scientifico ed Applicazioni Avanzate." We gratefully acknowledge the technical support given by CASPUR Consortium to the development of this research.

References

[1] Alziary de Roquefort B., *Jeux différentiels et approximation numérique de fonctions valeur,* 1^{re} *partie: étude théorique*, RAIRO Modél. Math. Anal. Numér., **25** (1991), 517–533.

[2] Alziary de Roquefort B., *Jeux différentiels et approximation numérique de fonctions valeur,* 2^e *partie: étude numérique,* RAIRO Modél. Math. Anal. Numér. **25** (1991), 535–560.

[3] Bardi M., *Some applications of viscosity solutions to optimal control and differential games,* in I. Capuzzo Dolcetta, P.-L. Lions (eds.), "Viscosity solutions and applications (Montecatini Terme, 1995)", Lecture Notes in Mathematics **1660**, 44–97, Springer, Berlin, 1997.

[4] Bardi M., Bottacin S., Falcone M., *Convergence of Discrete Schemes for Discontinuous Value Functions of Pursuit-Evasion Games,* in G. J. Olsder (ed.), "New trends in Dynamic Games and Application", Annals of the International Society of Dynamic Games **3**, 273–304, Birkhäuser, Boston, 1995.

[5] Bardi M., Capuzzo Dolcetta I., *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, Birkhäuser, Boston, 1997.

[6] Bardi M., Falcone M., Soravia P., *Fully discrete schemes for the value function of pursuit-evasion games,* in T. Başar and A. Haurie (eds.), "Advances in Dynamic Games and Application", Annals of the International Society of Dynamic Games **1**, 89–105, Birkhäuser, Boston, 1994.

[7] Bardi M., Falcone M., Soravia P., *Numerical methods for pursuit-evasion games via viscosity solutions*, in M. Bardi, T. Parthasarathy e T.E.S. Raghavan (eds.), "Stochastic and differential games: theory and numerical methods", Annals of the International Society of Dynamic Games **4**, 289–303, Birkhäuser, Boston, 1999.

[8] Bardi M., Koike S., Soravia P., *Pursuit-evasion games with state constraints: dynamic programming and discrete-time approximations,* Discrete and Continuous Dynamical Systems **6** (2000), 361–380.

[9] Breakwell J. V., *Time-optimal pursuit inside a circle*, Differential Games and Applications (Sophia-Antipolis, 1988), Lecture Notes in Control and Information Sciences **119**, 72–85, Springer, Berlin, 1989.

[10] Camilli F., Falcone M., *Approximation of optimal control problems with state constraints: estimates and applications*, in B. S. Mordukhovic, H. J. Sussman (eds.), "Nonsmooth analysis and geometric methods in deterministic optimal control", IMA Volumes in Applied Mathematics **78**, 23–57, Springer Verlag, Berlin, 1996.

[11] Cardaliaguet P., Quincampoix M., Saint-Pierre P. *Pursuit differential games with state constraints*, SIAM Journal of Control and Optimization **39** (2001), 1615–1632.

[12] Cardaliaguet P., Quincampoix M., Saint-Pierre P. *Differential games with state-constraints,* in the Proceedings of the Conference of the International

Society of Dynamics Games, St. Petersburg 2002.

[13] Cardaliaguet P., Quincampoix M., Saint-Pierre P., *Set valued numerical analysis for optimal control and differential games*, in M. Bardi, T. Parthasarathy and T. E. S. Raghavan (eds.), ”Stochastic and differential games: theory and numerical methods”, Annals of the International Society of Dynamic Games **4**, 177–247, Birkhäuser, Boston, 1999.

[14] Carlini E., Falcone M., Ferretti R., *An efficient algorithm for Hamilton-Jacobi equations in high-dimension*, Computer and Visual Sciences **7** (2004), 15–29.

[15] Cristiani E., *Fast Marching and semi-Lagrangian methods for Hamilton-Jacobi equations with applications*, Ph.D. Thesis, Dipartimento di Metodi e Modelli Matematici per le Scienze Applicate, Università ”La Sapienza”, Rome, Italy, 2006.

[16] Cristiani E., Falcone M., *A Fast Marching Method for Pursuit-Evasion Games*, published electronically in “Communications to SIMAI Congress", SIMAI 2006 (Baia Samuele, Ragusa, Italy, May 22-26, 2006), 1 (2006). http://cab.unime.it/journals/index.php/congress/article/view/136 /136

[17] Falcone M., *Numerical methods for differential games based on partial differential equations*, International Game Theory Review **8** (2006), 231–272.

[18] Falcone M., *The minimum time problem and its applications to front propagation*, in A. Visintin e G. Buttazzo (eds.), ”Motion by mean curvature and related topics”, De Gruyter Verlag, Berlin, 1994.

[19] Isaacs R., *Differential Games*, John Wiley and Sons, New York, 1965.

[20] Ishii H., Koike S., *A new formulation of state constraint problems for first-order PDEs*, SIAM Journal of Control and Optimization **34** (1996), 554–571.

[21] Koike S., *On the state constraint problem for differential games*, Indiana University Mathematics Journal **44** (1995), 467–487.

[22] Pesch H. J., Gabler I., Miesbach S., Breitner M. H., *Synthesis of optimal strategies for differential games by neural networks*, in G. J. Olsder (ed.), ”New trends in Dynamic Games and Application”, Annals of the International Society of Dynamic Games **3**, 111–141, Birkhäuser, Boston, 1995.

[23] Zhao H., *A fast sweeping method for eikonal equations*, Mathematics of Computation **74** (2005), 603–627.

Numerical Solution of the Game of Two Cars with a Neurosimulator and Grid Computing

Hans-Jörg von Mettenheim
Institut für Wirtschaftsinformatik
Gottfried Wilhelm Leibniz Universität Hannover
Königsworther Platz 1, D-30167 Hannover, Germany
mettenheim@iwi.uni-hannover.de

Michael H. Breitner
Institut für Wirtschaftsinformatik
Gottfried Wilhelm Leibniz Universität Hannover
Königsworther Platz 1, D-30167 Hannover, Germany
http://www.iwi.uni-hannover.de
breitner@iwi.uni-hannover.de

Abstract

The famous game of two cars is a pursuit-evasion dynamic game. In the extended version presented here, a correct driver (evader) on a freeway detects a wrong-way driver (pursuer in a worst case scenario), i.e., a car driving on the wrong lanes of the road or in the wrong direction. The correct driver must try to avoid collision against all possible maneuvers of the wrong-way driver. Additionally, he must try to stay on the freeway lanes. Analytically, the game is not fully solvable. The state-space is cut by various singular manifolds, e.g., barriers, universal, and dispersal manifolds. Here, discretized Stackelberg games are solved numerically for many positions in the state-space. The resulting trajectories and their adherent information are used to synthesize optimal strategies with artificial neural networks. These networks learn the optimal turn rates and optimal velocity change rates. The networks are trained with the high-end neurosimulator FAUN (Fast Approximation with Universal Neural Networks). A grid computing implementation is used which allows significantly shorter computing times. This implementation runs on low-budget, idle PC clusters and moreover power saving allows to wake up and shut down computers automatically. Parallelization on cheap hardware is one of the key benefits of the presented approach as it leads to fast but nonetheless good results. The computed artificial neural networks approximate the Stackelberg strategies accurately. The approach presented here is applicable to many other complex dynamic games which are not (fully) solvable analytically.

Key words. Dynamic game, Stackelberg game, synthesis of optimal strategies, artificial neural networks, parallel computation, grid computing, game of two cars.

AMS Subject Classifications. Primary 91A25; Secondary 68T10, 68M14.

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_12,

1 The Game of Two Cars Revisited

Even today, several thousand cases of people driving on the wrong side of the road are officially registered on German freeways every year. The actual number is estimated to be several times higher. Often, people are injured or even killed. A collision avoidance device can help to minimize the dangers resulting from wrong-way drivers. This can be modeled as a pursuit-evasion dynamic game. In the present paper, an enhanced version of the game of two cars is investigated. The idea is that the maneuvers of wrong-way drivers are generally unpredictable, e.g., because of drunkenness, fatigue, or panic. It is therefore suitable for the correct driver to assume the worst and to act, as if the wrong-way driver was trying to capture him. The correct driver chooses the optimal strategy against all possible decisions of the wrong-way driver. The resulting dynamic game is modeled as Stackelberg game and the neurosimulator FAUN is used to synthesize strategies.

To simplify our real life problem, we restrict ourselves to a collision avoidance problem for two cars: one is driven by the correct driver E, the other one by the wrong-way driver P. For this collision avoidance problem the kinematic equations can be modeled by:

$$\dot{x}_P = v_P \sin\phi_P, \tag{1}$$
$$\dot{x}_E = v_E \sin\phi_E, \tag{2}$$
$$\dot{y} = v_P \cos\phi_P - v_E \cos\phi_E, \tag{3}$$
$$\dot{v}_E = b_E \eta_E, \tag{4}$$
$$\dot{\phi}_P = w_P u_P, \tag{5}$$
$$\dot{\phi}_E = w_E(v_E) u_E; \tag{6}$$

see Fig. 1. The subscripts P and E of the notations refer to the wrong-way driver (the pursuer) P and the correct driver (the evader) E, respectively. The independent variable t denotes time, the state variables x_P, x_E denote the distance of P and E from the left-hand side of the freeway, and y denotes the distance between P and E orthogonal to the x-direction. The state variables ϕ_P, ϕ_E denote the driving directions of P and E and v_P, v_E are the velocities of P and E. The control variables u_P, u_E denote the turn rates of P and E, and η_E the velocity change rate. Without oversimplification, v_P is taken as constant and the maximum angular velocities $w_P(v_P)$ and $w_E(v_E)$ are prescribed depending on the type of car. The kinematic constraints of bounded radii of curvature are taken into account by the control variable inequality constraints

$$-1 \leq u_P \leq +1, \tag{7}$$
$$-1 \leq u_E \leq +1. \tag{8}$$

Steering $u_P = \pm 1$ and $u_C = \pm 1$ means executing an extreme right/left turn for P and E, respectively. The kinematic constraint of acceleration and deceleration for

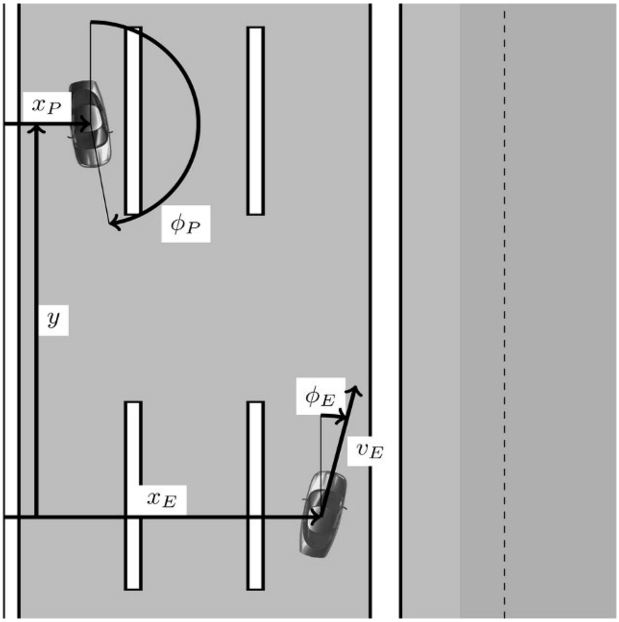

Figure 1: State variables of the collision avoidance problem for two cars. The evader is allowed to be on the edge strip with two wheels. The corresponding state variables for evader and pursuer are denoted with subscripts E and P, respectively.

E is taken into account by the control variable inequality constraint

$$-1 \leq \eta_E \leq +0.48. \tag{9}$$

The constraint

$$0\text{m} \leq x_E \leq 15\text{m} \tag{10}$$

considers the fact that E wants to stay on the freeway. A similar constraint for P is explicitly rejected, as a worst-case situation for E is analyzed. It is indeed imaginable that the wrong-way driver P doesn't care if his car is damaged or he is injured when leaving the freeway.

The maximum deceleration rate is given by $b_E = 10.72\text{m}/\text{s}^2$. This corresponds to a modern car. The maximum angular velocities are given by:

$$w_P = \frac{\mu_0 g}{v_P}, \tag{11}$$

$$w_E = \frac{\mu_0 g}{v_E}. \tag{12}$$

The constant $\mu_0 = 0.55$ corresponds to the adhesion coefficient on dry asphalt and $g = 9.81\text{m}/\text{s}^2$ is the acceleration of gravity.

The goal of the correct driver E is to avoid collision against all possible maneuvers of P. It can be formulated as

$$\min_{\gamma_P \in \Gamma_P} \max_{\gamma_E \in \Gamma_E} \min_{t_0 \le t < \infty} d(z(t)) \tag{13}$$

with the vector $z := (x_P, x_E, y, v_E, \phi_P, \phi_E)$ of state variables, with (feedback) strategies $\gamma_P(z)$ and $\gamma_E(z)$, with sets Γ_P and Γ_E of all admissible strategies, and with Euclidean distance $d(z)$ between P and E. The control variables are determined by $u_P(z) := \gamma_P(z)$ and $(u_E(z), \eta_E(z)) := \gamma_E(z)$. The strategy γ_P is admissible, if the constraint (7) is fulfilled for all $t \in [t_0, t_f]$, and the strategy γ_E is admissible, if the constraints (8), (9), and (10) are fulfilled for all $t \in [t_0, t_f]$. The collision avoidance maneuver starts at the time t_0 at $z(t_0) = z_0$, when the correct driver E notices P. The collision avoidance maneuver ends at the terminal time $t_f < \infty$ when the minimum distance between P and E is reached. Theoretically, it is imaginable that the game continues if P turns on the street. This case is deliberately ignored.

The present collision avoidance game is an enhancement of the game of two cars as introduced, e.g., in [23]. It can be embedded in the theory of pursuit and evasion. If the state constraint (10) is omitted and constant velocity v_E is assumed, the present game corresponds to the game of two cars. If additionally the correct driver E needs not obey the constraint of bounded curvature, i.e., $u_E \in]-\infty, \infty[$, the homicidal chauffeur game arises. In [33] a neural network solution for this relatively simple game is presented. Nevertheless, the full solutions of the game of two cars and even the homicidal chauffeur game are extraordinarily baffling. The state-space is cut by diverse singular manifolds. Neither the game of two cars nor the homicidal chauffeur game can be solved fully, i.e., optimal strategies $\gamma_P^*(z)$ and $\gamma_E^*(z)$ cannot be calculated explicitly for all z, for all velocities v_P and v_E, and for all maximum angular velocities w_P and w_E. Nevertheless various collision avoidance problems have been investigated; see, e.g., [3, 4, 13–15, 26–30, 34, 37–39, 42, 44, 46, 47].

As the goal of this paper is not to give an analytical solution for the presented enhanced game of two cars, but to show that even complex dynamic games can be solved with neural networks, a more basic approach is chosen; see also [7]. Therefore, the computation of strategies is realized with a Stackelberg game. First, E optimizes his strategy against all possible maneuvers of P. Then, P optimizes his strategy against the well-known strategy of E. This very conservative assumption is unfair for E but is part of our worst-case analysis.

It has to be noted, that the need to compute a solution first, albeit numerical, presents a potential limitation. The required numerical computation is time consuming and can only be difficultly realized in real-time. This is, of course, one of the reasons, why artificial neural networks are used.

2 The Neurosimulator FAUN

The following excursus explains the basics of a neurosimulator and the advantages of parallelization. Generally, a neurosimulator can be useful when collected data has to be connected even if classical methods like linear or nonlinear regression don't work. At the Institut für Wirtschaftsinformatik in Hanover, the neurosimulator FAUN (Fast Approximation with Universal Neural Networks) has been under active development since 1996.

The process of finding suitable artificial neural networks is called training. This training can take CPU hours or even days on a single personal computer. This is explained by the fact that several networks are tried, normally between 100 and 10,000. The trained neural networks have no mutual dependencies. It is possible to split the training task on several threads. This is meant by coarse-grained parallelization in contrast to fine grained parallelization. The advantage of coarse-grained parallelization is that the communication network can be quite slow, standard fast Ethernet with 100 Mbit/s is sufficient. The parallel version of FAUN runs on homogeneous computer clusters as well as on heterogeneous personal computer networks using a parallel programming package which has to be previously installed and configured. This is a disadvantage and inspired the authors to the development of the grid computing client briefly described in this paper.

The Hanover School of Economics owns a student and staff computer cluster with — amongst others — 38 Pentium IV computers at 2.66 GHz running the operating system Windows XP professional. As the cluster is rarely used and almost only for special lectures the combined power of these computers is free most of the time. This leads to the idea of using the unused capacities for calculations and therefore reducing the need of an external computation-only cluster or an expensive supercomputer. The condition is that other occasional users are not disturbed and that the configuration necessities are low. To make the configuration more ecological, a mechanism is needed that allows to wake up the computers from the standby state and to shut them down again.

The currently available parallel programming packages cannot fulfill all these wishes. They require configuration and wake up or shutdown is not implemented. This means that an alternative concept has to be developed. Our new grid computing client is installed by simply copying the files onto the computer and is further on managed via a web interface. As modularity is a main design goal, the client cannot only cooperate with the neurosimulator but offers functionalities that make it possible to hook up other programs. For the basic programming paradigms used, see [1, 17, 18, 22, 24, 35, 40, 41].

In the following the concepts of low-budget grid computing and the FAUN grid computing client are briefly discussed to give an idea of the software used for the strategy synthesis. Reachable economies at the School of Economics are highlighted. Then, the Stackelberg game used to generate trajectories for the neural network training is presented. Finally, the results of neural network guidance compared to the equivalent Stackelberg game are discussed.

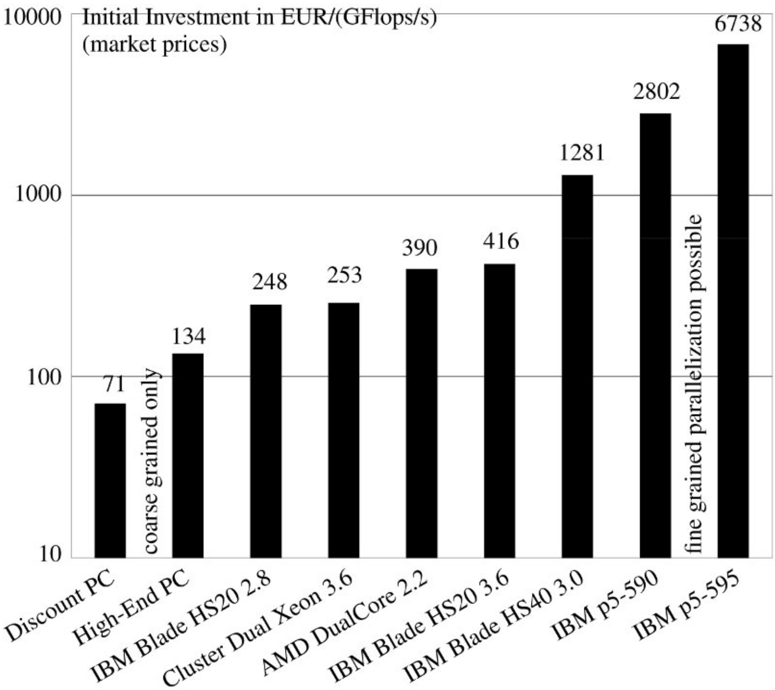

Figure 2: Price comparison of different computer types for coarse-grained parallelization. Note that costly computers offer more advantages than simple CPU speed and that therefore it is not advisable to use a cluster for massively parallel tasks.

3 The FAUN Grid Computing Client

Grid computing is normally used when high performance in a broad sense is needed. This can mean high availability or HPC in its actual sense, i.e., raw computing power is required. It is worth looking at the significant differences in the initial investment for obtaining roughly the same computing power; see Fig. 2. Ranging from affordable 71 Euros per GFlops/s with a standard personal computer the range goes up to more than 6,000 Euros per GFlops/s with a massively parallel IBM supercomputer. Yet, these two extremes are not identical. The standard personal computer comes without enhanced manufacturer support, doesn't offer remote administration features and — perhaps most important of all — has a comparatively slow interprocessor communication. Indeed the maximum available speed is 1 GBit/s with a standard network card. On the other hand, the supercomputer ships with 32 CPUs which are mutually connected at a 100times that speed. The 32 CPUs have two cores and correspond to $(2 \times 32 =)$ 64 conventional CPUs.

Therefore, Fig. 2 has to be read carefully and the potential customer has to decide, whether high interprocessor communication speeds are important or not so relevant for the given task. In several cases and also with a neurosimulator the answer is clear: interprocessor communication speed isn't the bottleneck. It is advisable to use standard hardware. However, it is not enough to simply buy the hardware. The computers also have to communicate with one another in a sensible way so that the required task is achieved. That is the reason why certain manufacturers offer so-called blades. These are made of mostly standard CPUs but have a very compact design and are easy to manage remotely. Although not offering the advantages of massively parallel systems they cost at least approximately four times as much as a standard computer even for a small blade. Most institutions have unused computing capacities like clusters or staff computers. This leads to the wish to use these resources comfortably. A middleware is needed that interfaces the actual computational application, e.g., the neurosimulator, with basic management functions being file transfer, message passing and power management.

The importance of the latter point is shown in Fig. 3. Actually, the cluster of the School of Economics for students and staff with its 38 Pentium IV computers is powered on only on the five working days from 10 am to 4 pm, i.e., 6 hours per day on roughly 40 weeks per year. The average power consumption is shown in Tab. 3. The standby state means that the computers can be woken up with a so-called magic packet.

It is assumed that the computers are powered on, but mostly without load as is the case when doing simple typewriting or Internet surfing. Using a standard rate of 18.13 ct/kWh the energy costs per year are calculated. The second considered case (full on) happens when the cluster is used for computation *without* power management, i.e., the cluster would be used to full capacity four hours per working day on 40 weeks for computation additionally. If the cluster wasn't powered off during the computation pauses, 3134 Euros would be wasted in idle time every year. The same case *with* power management (upper graph in the figure) would reduce the energy costs considerably.

Another concept — distributed computing — can also provide additional computing capacity nearly for free. Distributed computing describes the fact that the computation is done at various locations and on various types of computers. This can be, e.g., staff computers or laptops all over the university or even at home. The computers only need to be at least temporarily connected via a network. This network is normally an Ethernet or the Internet. The challenge for distributed com-

Table 1: Measured average power consumption (Pentium IV, 2.66 GHz)

state	consumption in Wh
standby	3.4
idle on	67.5
full load	113.3

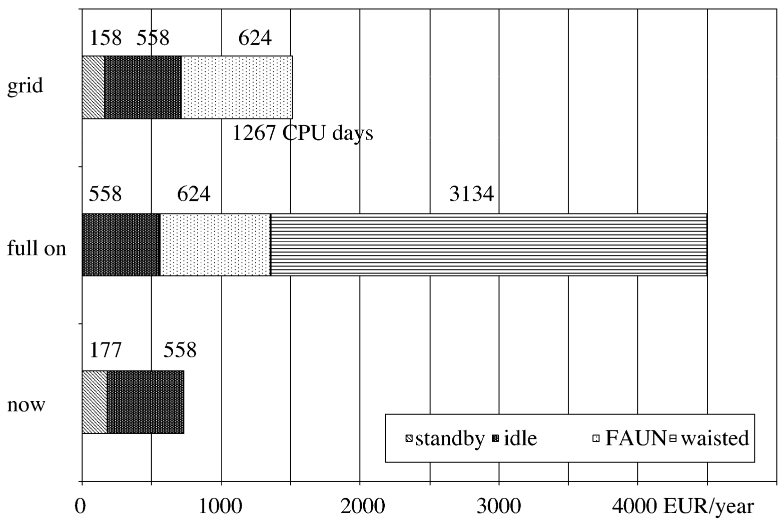

Figure 3: Power consumption (equivalent value in Euro) of the cluster with grid functionality and in full on status. *Standby* describes the power consumed by the computers simply being plugged. *Idle* means that the computers are on but do nothing. *FAUN* is the power used by FAUN computations. *Wasted* finally stands for the value of the current that would be drained if the computers were on all the year.

puting programs is to unify the computers to work together. See [2] for advanced links. Yet, a client providing grid computing functionalities as middleware is also suitable for distributed computing. The combination of distributed and grid computing used with parallel programming techniques is the target of the client. Distributed computing offers the opportunity to use computers all over the world with an easy to install client. Grid computing leads to error tolerance and an easy configurable system with advanced features. Up until now, usually only automatic wakeup and shutdown is concretely implemented.

For simple parallelization purposes specialized programming packages can be used. For general information, [22] and [10] are a good start. The most common packages are the Message Passing Interface (MPI) and the Parallel Virtual Machine (PVM); see [12, 16] for information. A first parallel version of the neurosimulator FAUN indeed uses PVM; see [9] and [8] for examples. But these packages are not suitable for programming a modular grid computing client.

The testbed for the grid computing client is shown in Fig. 4. It is decided that an individual protocol is developed for the grid computing client, which will connect to a central computing server. Known criteria for quality software are taken into account.

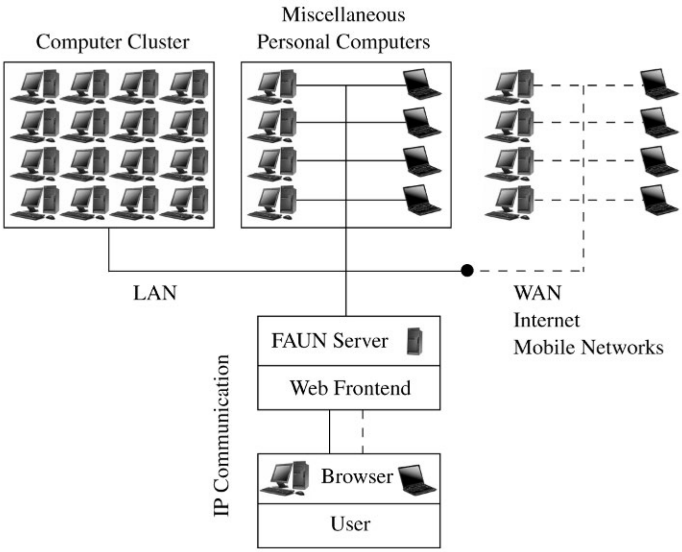

Figure 4: FAUN grid computing configuration.

A detailed description of the neurosimulator FAUN can be found in [6]; see also [5] and [25] for real life applications and [31, 32] for the last development steps. For information on differences between a coarse-grained and fine-grained parallelization see [8]. General introductions and more details are offered by [11, 16, 17, 19–21, 36, 43, 45]. The present parallelization of FAUN uses the master slave principle. The master program `faun_m` starts several instances of the slave program `faun_s`.

To get a good quality neural network it is necessary to compute a certain amount of neural networks. This amount varies with the difficulty of the problem. The quality of the neural network is primarily measured by the difference between the output as it is given by the neural network formula and the real output given by the training and validation data.

The specific amount of neural networks that will be calculated has to be determined heuristically. This amount also depends on the “difficulty” of the problem. After the computation the best network is chosen. This is the network that approximates the data better than all others. The neurosimulator uses a global optimization procedure with SQP optimization. Yet, it is not uncommon that difficult problems need long calculation times: One minute or more for every successfully trained neural network on a single personal computer, Pentium IV at 3.4 GHz.

The calculation time increases linearly with the amount of training and validation data. This is the initial reason why a parallel and now a grid computing version of

FAUN is developed. Computing times of one day or more are simply not acceptable especially for real time applications.

The parallelized version of FAUN based on the grid computing client runs without preliminary installation and offers additional features in comparison to the first version using PVM. The advantage of coarse-grained parallelization is common to all versions: a parallelization degree between 0.9 and 0.97. The meaning of this number is that if 100 computers work together, the computation will not be exactly 100 times as fast as with a single computer. But it will still be 90 – 97 times as fast as with a single computer. The loss of computing power can be explained with the additional administration and communication overhead[1].

A design goal coming together with the development of the grid computing client is to interface the parallel HPC version of FAUN to the already existing web interface for single processor computation developed by S. König. Now the web interface allows to manage the additional grid computing features, i.e., wake up and shutdown of the clients. It is then possible to decide which computers will be used for the computation. This offers the possibility that several users can share the resources simultaneously.

Another specialty has to be taken into account. The cluster of the School of Economics which is mainly used as testbed is protected with a firewall. That means that no access is possible from *outside* although the clients can connect from *inside* to arbitrary destinations. This is a problem when “sleeping” computers have to be woken up, because no inner access is available. A solution consists in leaving an intermediary computer with low power consumption on in the cluster. It will receive the wake up requests and send a magic packet to the desired computer.

4 Stackelberg Game and Neural Network Solution

In order to be able to compute a Stackelberg game, a sensible discretization has to be chosen. For the three control variables u_P, u_E, and η_E an appropriate set of controls has to be found. If $|U_P|$, $|U_E|$, and $|H_E|$ denote the number of elements in the respective set and s denotes the number of steps in the discretization, we have:

$$p = (|U_P| \cdot |U_E| \cdot |H_E|)^s \tag{14}$$

different strategies to choose from. As these strategies have to be computed for many different points to get good training data, it is wise to take small sets. There-

[1] Imagine a bunch of potatoes that have to be peeled (famous “potato peeling problem”). A single person will work for a long time. If a second person helps the peeling will probably be twice as fast. But if more and more persons join in the peeling time will be spent on handing the potatoes from one to another and the entire process will be more and more inefficient. Finally, the largest potato determines the minimum time achievable.

fore, the smallest sets are chosen, which gives:

$$U_P = \{-1, 1\}, \tag{15}$$
$$U_E = \{-1, 1\}, \tag{16}$$
$$H_E = \{-1, 0.48\}. \tag{17}$$

Rightly, the reader may argue that this is too small a set for getting exact results. But two facts, besides the required computation time, have to be considered. First, sensible strategies seem to indicate that only extreme left/right turns or full deceleration/acceleration give good results. Second, the question has to be answered, which *exact* numbers should then be chosen to complete the sets. As the focus is on showing the opportunities of neural networks related to dynamic game, it seems acceptable to go with this simple approach. The choice of s is also dictated by the need for keeping the computation time small. Even with parallelization readily available an exponential increase can't be ignored. A viable approach is given with

$$s = 10, \tag{18}$$

although it is clear, that a higher value would have yielded a finer discretization. Stackelberg strategies are computed for 69800 points in the six-dimensional state space. For the computation parallelization is used, too. The strategies lead to 690000 points for which strategies exist because of (18). Every point in the state space belonging to a strategy can in fact be considered as being the initial set of another strategy.

Additionally, the strategies can be “mirrored” for P at the middle of the road, when the corresponding state variables, especially the angles, are also mirrored. This doubles the number of points so that approximately 1.4 million exist. The resulting maneuvers for five chosen start points can be seen in the upper row of Fig. 5. Even without a detailed analysis, it can be seen that the reactions are sensible and that the street boundary conditions are met. As positions are calculated by the display program from the center of the rectangle representing the car, it is acceptable for E to be off the road with *two* wheels. But the reader will notice hat “half” of E is always on the road and that means that the street boundaries are honored.

It is now necessary to cut out training and validation data for the neurosimulator FAUN. As 1.4 million points would be too much, $50{,}000$ points are taken. For this, the state space is equally divided in 100 hypercubes. These hypercubes are filled with 500 randomly chosen points of the appropriate region. The purpose of this procedure is to prevent the neurosimulator from overtraining one region and neglecting another. Finally, three hyperspheres are used as validation data. The ratio of training and validation data is $4:1$ giving approximately $42{,}000$ training and $8{,}000$ validation inputs.

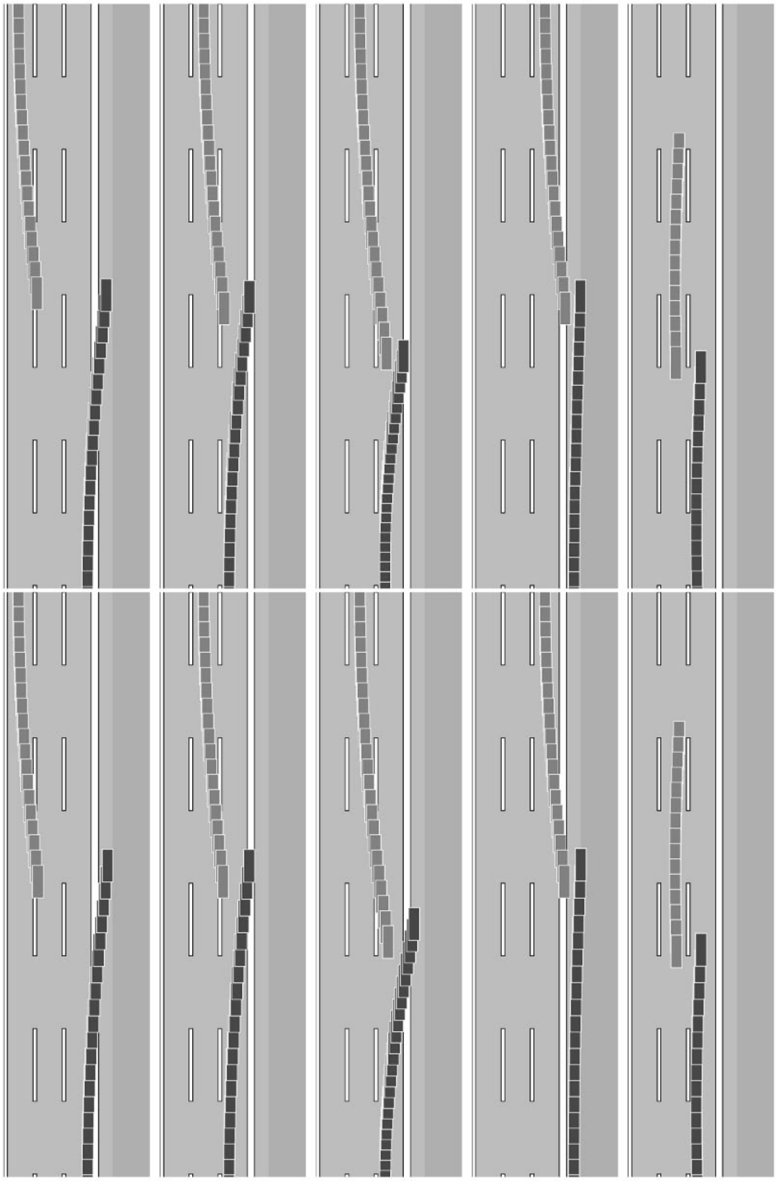

Figure 5: Trajectories for different start configurations of x_P and x_E. The first row shows the result from the Stackelberg simulation; in the second row neural network guidance is used. The results hardly differ. Pursuer coming from above (grey), evader coming from below (black).

In the case of neural networks parallelization can achieve a high speed up and shortens the computation time to *only* several hours in contrast to several days. For different parameter settings of the neurosimulator $1{,}000$ neural networks are computed for each control variable and the best is taken. For instance, the resulting neural network for control variable u_E is given by the following lengthy result. It has to be kept in mind, though, that this formula can be computed in *real time* compared to the Stackelberg solution:

$$\begin{aligned}
u_E =&((-1) + (2 * ((\tanh(((\tanh((3.36253335\\
&+ (((2 * ((x_P - 0.01061)/14.97877)) - 1) * 0.78373677)\\
&+ (((2 * ((x_E - 0.01296)/14.97407)) - 1) * (-3.2545598))\\
&+ (((2 * ((y - 0.76436)/65.01923)) - 1) * (-0.2829929))\\
&+ (((2 * ((v_E - 14.04071)/40.37917)) - 1) * 0.02903405)\\
&+ (((2 * ((\phi_P - 2.92346)/0.43626)) - 1) * (-0.14377207))\\
&+ (((2 * ((\phi_E - (-0.27143))/0.60569)) - 1) * (-2.4273138))))\\
&* 7.3417816) + (\tanh((4.32403995\\
&+ (((2 * ((x_P - 0.01061)/14.97877)) - 1) * (-0.89812738))\\
&+ (((2 * ((x_E - 0.01296)/14.97407)) - 1) * 3.813688)\\
&+ (((2 * ((y - 0.76436)/65.01923)) - 1) * (-0.3494907))\\
&+ (((2 * ((v_E - 14.04071)/40.37917)) - 1) * 0.065534994)\\
&+ (((2 * ((\phi_P - 2.92346)/0.43626)) - 1) * 0.21560794)\\
&+ (((2 * ((\phi_E - (-0.27143))/0.60569)) - 1) * 2.8371935)))\\
&* (-6.4195971)) + (-0.42285788)\\
&+ (((2 * ((x_P - 0.01061)/14.97877)) - 1) * (-5.0039272))\\
&+ (((2 * ((x_E - 0.01296)/14.97407)) - 1) * 5.531147)\\
&+ (((2 * ((y - 0.76436)/65.01923)) - 1) * (-0.024304409))\\
&+ (((2 * ((v_E - 14.04071)/40.37917)) - 1) * 0.028052)\\
&+ (((2 * ((\phi_P - 2.92346)/0.43626)) - 1) * 1.5422002)\\
&+ (((2 * ((\phi_E - (-0.27143))/0.60569)) - 1) * 4.8255601)))\\
&+ 0.95)/1.9))).
\end{aligned} \tag{19}$$

In the context of this paper we don't want to explain the above formula in detail as an in-depth analysis of neural network formulas is furnished in [6]. For the reader it is more important to grasp its general significance. The formula is *simple* in the sense that besides the $\tanh$ function only elementary operators are used on the state variables: The first factor of the outer product on each line is only responsible for appropriate scaling. The second factor is the corresponding weight. The formula is thus *fast* to compute.

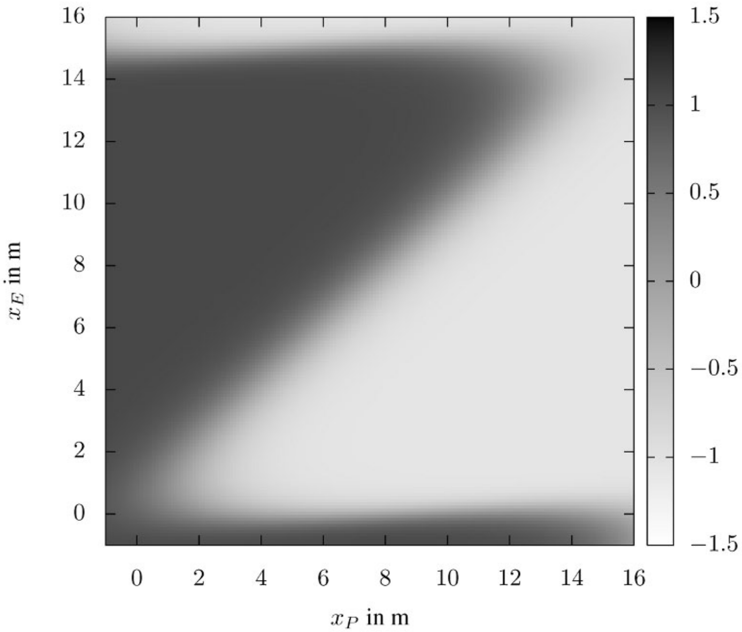

Figure 6: Resulting guidance u_E for different positions of x_P and x_E using neural networks ($y = 42$m, $v_E = 42\frac{m}{s}$, $\phi_P = \pi$, $\phi_E = 0$). The Z-shape marks the dispersal surfaces of this game.

As can be seen on Fig. 5 which compares Stackelberg game and neural network guidance, the results hardly differ “optically”. And, indeed, a more detailed analysis at some exemplary points yields good results. Figure 6 shows the result of neural network guidance for the control variable u_E. It is clear that E has to steer left when he is left of P and right when he is right of P. Therefore, a dispersal surface at the diagonal for $x_P = x_E$ can be expected and is indeed reproduced by the neural network. Continuous colors are added for better legibility of the plot. Normally the output of the neural network would be mapped to ± 1 as this is, what is trained with the Stackelberg game. Figure 6 also shows that the boundary constraints are observed, leading to the typical Z-shape of the figure. Indeed, for $x_E \approx 0$m and $x_E \approx 15$m E has to steer in the opposite direction to stay on the freeway. Of course, this brings him nearer to P and thus to capture.

The occurring error is analyzed in more detail in Fig. 7. A point is set if the sign of the neural network output is opposite to the Stackelberg game. Errors only happen at the dispersal surfaces and don’t reach far. This means, that a little away from the dispersal surface the sign is correct. A small error on the diagonal doesn’t

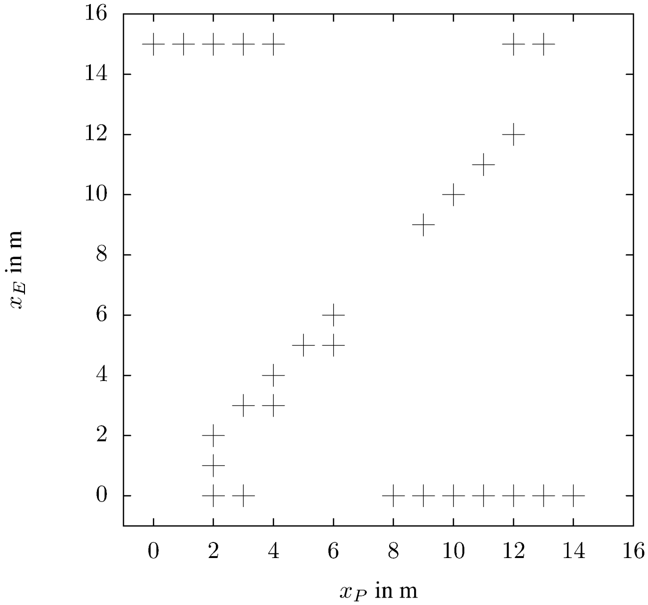

Figure 7: Occurring errors for u_E. A point is set, if the neural network would steer in the opposite direction of the Stackelberg solution. This happens only on the diagonals, where it doesn't matter much and on the limits of the road for x_E where the boundary conditions have to be met ($y = 42$m, $v_E = 42\frac{m}{s}$, $\phi_P = \pi$, $\phi_E = 0$). Figure 10 complements this qualitative error plot with absolute values.

harm much as this constellation is very bad for E anyway. Other small errors can be observed at the street boundaries, where the neural network guidance sometimes reacts too early and sometimes too late as will be explained in more detail on Fig. 8. Note, that in fact errors at the street boundaries are of marginal importance as this could also be controlled by other means. Indeed, today's trucks are already often equipped with special devices that warn the driver, if the truck leaves its lane.

Figure 9 shows the value of the game. In this case, the time until capture is not taken, but merely the minimal distance between P and E, because it is the payoff which divides between accident (capture from P) and unharmed drive (evasion of E). The actual minimum distance for evasion can be set adequately but $d = 1.5$m seems a good approximation for a normal car. Figure 9 has to be read as follows: black areas signalize capture in other cases E can evade. The upper plot makes the whole game look satisfactorily for E as capture only occurs when P and E are almost face to face on the freeway. But this doesn't have to always be like that. The good results for E can only be achieved because of the relatively small $y = 42$m at the beginning of the game in the upper plot.

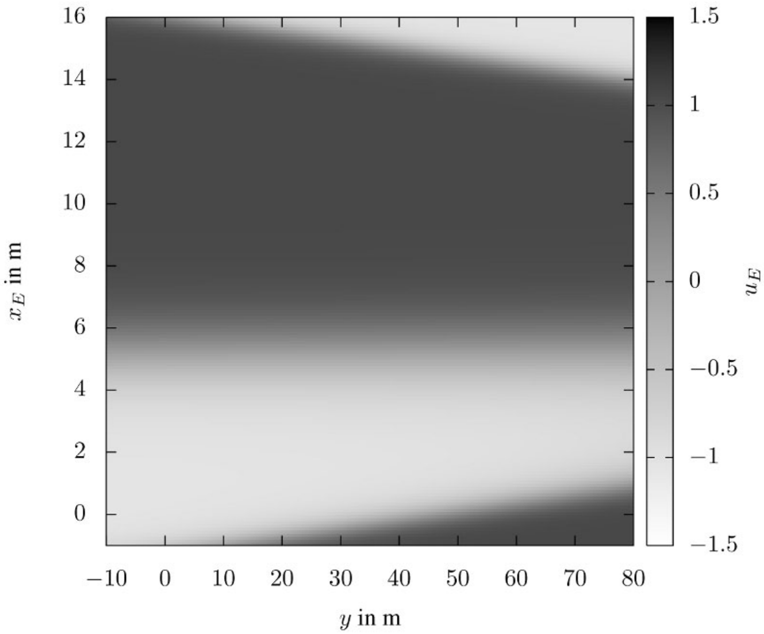

Figure 8: Neural network guidance for different values of y and x_E, while $x_P = 5$m fixed ($v_E = 42\frac{m}{s}$, $\phi_P = \pi$, $\phi_E = 0$). The guidance is (almost) error free on the road. The (hard) boundary conditions are not always honored for small distances of y, but this doesn't change the outcome of the game much.

If this distance is significantly greater P *can* achieve capture. We have to consider that at least E is bounded by the width of the road. The Game of Two Cars therefore offers disadvantages for E compared to a game where, for instance, the whole xy-plane can be played on. Small distances are taken, because on a normal freeway the distance drivers and possibly collision avoidance gear can overview are not too long. Only for small distances are collision avoidance devices of interest, because they close the gap which results from the human capacity of reaction. A gloomier outlook for E can be seen in the lower part of Fig. 9. The initial distance is approximately twice as large as in the upper plot and the speed at the begin of the game is also reduced. When both actors start in the middle of the road nothing changes much. But when P and E are both near the border of the same side of the freeway the situation worsens for E as P has more time (and more space for maneuvers) to close in on E. If the initial distance between P and E is larger the situation degrades even more. A similar worsening can also be noticed for smaller starting velocities v_E of E, although a slight compensation happens due to the increased possibility for maneuvers. Indeed, the angular velocity w_E depends on E as shown in (12). In certain situations a speed reduction can therefore be interesting, see below.

Figure 10 analyzes the case in more detail. The difference between the Stackelberg and the neural network result is shown. As can already be expected from the qualitative Fig. 7, the biggest errors happen on the dispersal surfaces. On the other hand, it has to be noticed that the errors are small in comparison to the width of the street. For the biggest error the inequality:

$$\Delta_{\text{max}} < \frac{0.7\text{m}}{15\text{m}} \approx 0.047$$

holds. The average error is significantly smaller:

$$\Delta_{\text{avg}} < \frac{0.2\text{m}}{15\text{m}} \approx 0.013.$$

Most parts of the diagonal are within Δ_{avg} and therefore the outcome for E is not much worse using neural network guidance compared to the Stackelberg game.

Finally, the resulting dependencies of u_E using neural network guidance for various distances y is plotted in Fig. 8. Deliberately the initial position of P is set to $x_P = 5\text{m}$. As might be expected, a dispersal surface is clearly seen at $x_E = 5\text{m}$. However, the behavior at the boundaries of the freeway is surprising. The "rough direction" is right, but the neural network seems to react too early for large distances y and too late for small distances. This is a disadvantage but can be explained by the distribution of the training data. Although the distribution is smoothed by cutting out hypercubes it is nevertheless the case, that much more points have been calculated for small distances y as these are the most interesting. For larger distances ($y \approx 80\text{m}$) fewer points have been calculated and very few are on the border. The frequency of boundary points is low for large distances y and high for small distances. As neural networks are continuous functions the results are exaggerated for small and large y. A better fine tuning of the training parameters for the neurosimulator might improve the behavior and will be explored.

Up to the emphasis of the discussion has been on the control variables u_E and u_P — the steering directions of E and P. However, the game comprises a third control η_E being the velocity change rate. Figure 11 shows that acceleration or deceleration is sensible for different initial ratios of $v_E(t_0) : v_P$. First, a dispersal surface can be found at $v_E(t_0) = v_P(= \text{const.})$. In the neighborhood of this boundary E will accelerate if $v_E(t_0) > v_P$ and decelerate if $v_E(t_0) < v_P$, the regions marked *B* and *C*, respectively. (The singular case $v_E(t_0) = v_P$ is not analyzed here.) This can be explained by the fact that in both cases an advantage can be gained. Accelerating reduces the time until $y = 0$ and thus the time P has for maneuvering. In contrast, by decelerating E increases his maneuverability but also the total time of the game. If, e.g., E accelerates with $v_E(t_0) < v_P$ the disadvantage of the reduced maneuverability doesn't outweigh the advantage of the total game time reduction and vice versa.

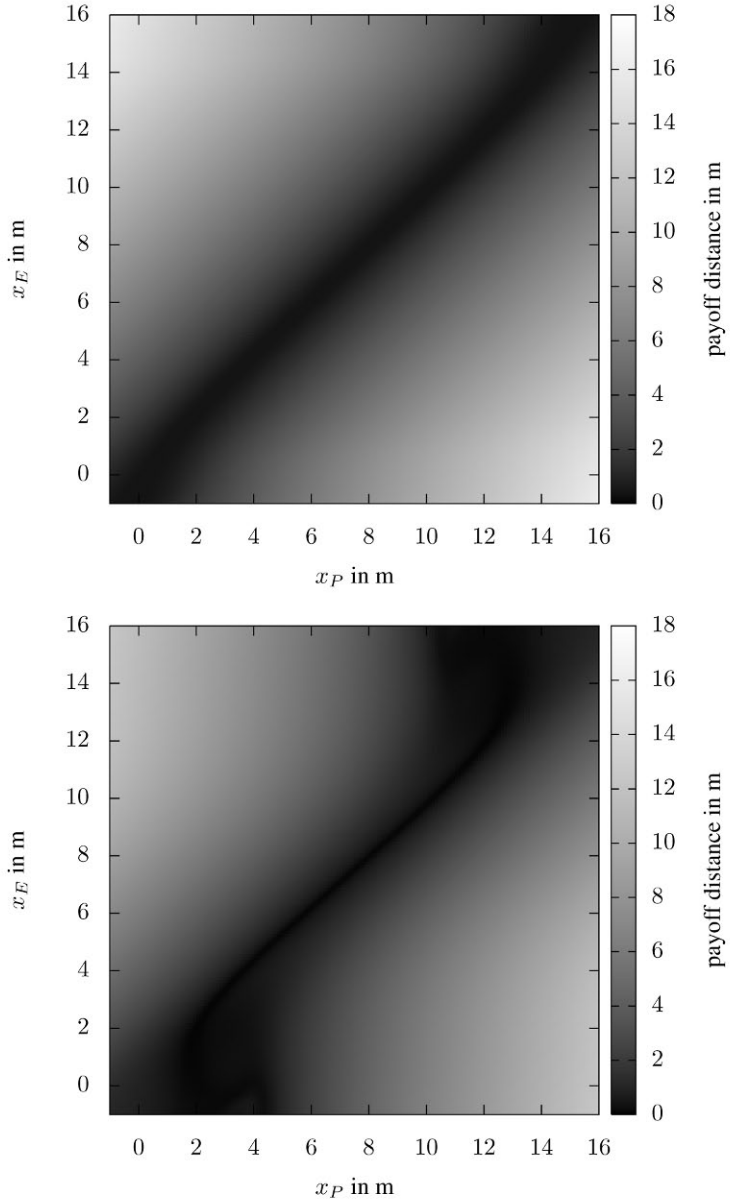

Figure 9: Payoff distance. Upper plot: $y = 42$m, $v_E = 42\frac{m}{s}$, $\phi_P = \pi$, $\phi_E = 0$. Lower plot: $y = 80$m, $v_E = 27\frac{m}{s}$, $\phi_P = \pi$, $\phi_E = 0$. Black color signalizes that a collision cannot be avoided.

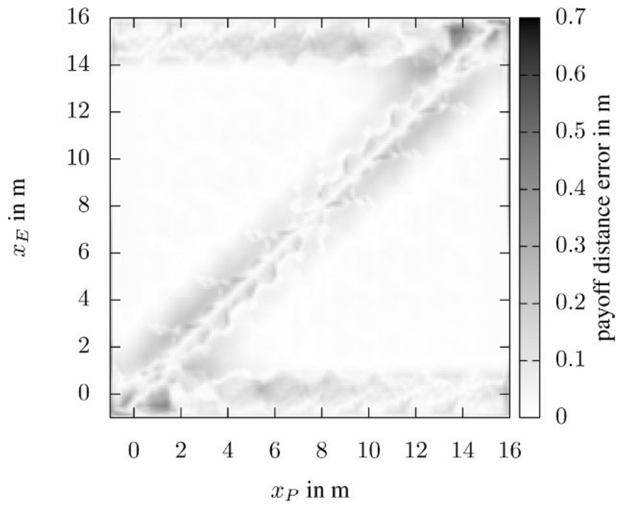

Figure 10: Payoff distance error: $y = 42$m, $v_E = 42\frac{m}{s}$, $\phi_P = \pi$, $\phi_E = 0$. The typical Z-shape of the dispersal surface appears. The error stays mostly below 0.2m. The pattern which can be seen in the low error regions is due to the interpolation through neural networks, which will deviate from the Stackelberg solution in the areas where no training data are available.

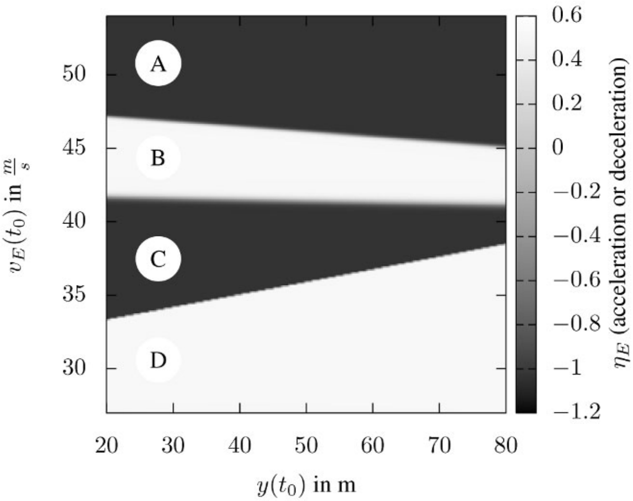

Figure 11: η_E for fixed $x_E(t_0) = 11$m, $x_P(t_0) = 4$m. Negative values signify that E is decelerating while positive values show an acceleration. The dispersal surface at $v_E(t_0) = v_P$ is noticeable.

There is, however, a sharply delimited dispersal surface where it again makes sense for E to decelerate (even if $v_E(t_0) > v_P$) or accelerate (even if $v_E(t_0) < v_P$), regions *A* and *D*. The greater the initial distance $y(t_0)$, the sooner this inversion happens. Qualitatively, both phenomena are understandable. On the one hand, the disadvantages of reduced maneuverability resp. velocity cannot be negated. On the other hand, for short distances and hence short game duration, the choice of η_E is more persistent and an inverse strategy only occurs for very high or very low velocities. It is clear that this qualitative analysis is unsatisfying and that further research will be dedicated to a solid explanation of Fig. 11.

The question remains whether the resulting neural network can be implemented as onboard system. The problems of positioning in the state space apart: Eq. (19) consists of 3 hyperbolic tangents, 39 multiplications, 19 divisions, 22 additions, and 50 subtractions. The hyperbolic tangents are expensive in terms of floating point operations and count for 10. This gives 160 floating point operations for one control. As E has two controls a total of 320 operations follows. Even if 1000 computations occur per second, i.e., $320*10^3$ operations, modern processors easily manage $3 * 10^9$ operations per second.

5 Conclusion and Outlook

Artificial neural networks are successfully employed to synthesize optimal strategies for the enhanced game of two cars. This game is much more realistic than the original game of two cars because constraints keep the correct driver on the freeway lanes always.

The usage of grid/distributed computing offers high-performance, low-budget computing (often at no costs for idle computers). Idle PCs are typical for public or company PC/computer clusters, e.g., at the Hanover School of Economics 38 modern PCs are idle during every night and on weekends. Software quality requirements increase: Issues like security and integrity become more important and moreover user friendliness, maintainability, and secrecy (user shouldn't notice computations on their office PCs). Grid/distributed computing is of particular relevance for all tasks of coarse-grained parallelization where interprocessor communication speed is of minor importance for the overall performance.

The neurosimulator FAUN repeats analogous tasks many times in order to train many artificial neural networks. Thus FAUN is well suited for distributed computing and a coarse-grained parallelization cuts down computing times significantly. Problems like the enhanced game of two cars which need computing days or even weeks on a standard PC can be solved in few hours. FAUN can also be used, e.g., for the solution of other (pursuit-evasion) dynamic games, aerospace guidance problems or forecasting problems of exchange or interest rates.

Nevertheless the usage of artificial neural networks is no easy to use black box method as often assumed. Problem analysis and formulation and also training necessitate deep knowledge. Usually knowledge about the special application is important, too. With the distributed computing version of FAUN the user friendliness increases as usage is possible via a graphical web frontend. No manual editing of configuration files is necessary.

The developed grid computing client is not limited to the neurosimulator FAUN. Modularity enables the usage with other programs which can therefore benefit from automatic client updates and remote power management. Future development aims at improved security, reduction of bandwidth requirements, and better automation of installation and management.

The application of artificial neural networks to general dynamic games is very promising. It cannot replace a thorough mathematical analysis (and a solution where possible) but for many problems not (fully) solvable a fast and reliable numerical approximation of the solution often becomes possible. Six state and three control variables are involved here but even higher-dimensional dynamic games can be solved numerically using artificial neural networks and a high-end neurosimulator.

References

[1] A. Abbas. *Grid Computing: A Practical Guide to Technology and Applications*. Charles River Media, Boston, 2004.

[2] H. Attiya and J. L. Welch. *Distributed Computing: Fundamentals, Simulations, and Advanced Topics*. Wiley-Interscience, New York, 2004.

[3] T. Ba^{3}ar and G. J. Olsder. *Dynamic Noncooperative Game Theory*. Academic Press, London, 1995.

[4] P. Borowko and W. Rzymowski. On the game of two cars. *Journal of Optimization Theory and Applications*, 44(3):381–396, November 1984.

[5] M. H. Breitner. Robust optimal onboard reentry guidance of a space shuttle: Dynamic game approach and guidance synthesis via neural networks. *Journal of Optimization Theory and Applications*, 107:484–505, 2000.

[6] M. H. Breitner. *Nichtlineare, multivariate Approximation mit Perzeptrons und anderen Funktionen auf verschiedenen Hochleistungsrechnern (in German)*. Akademische Verlagsgesellschaft, Berlin, 2003.

[7] M. H. Breitner. Usage of artificial neural networks for the numerical solution of dynamic games. In T. L. Vincent, editor, *Proceedings of the Eleventh International Symposium on Dynamic Games and Applications, Tucson, Arizona*, volume 1, pages 62–79. University of Arizona Press, 2004.

[8] M. H. Breitner, P. Mehmert, and S. Schnitter. Coarse- and fine-grained parallel computation of optimal strategies and feedback controls with multilayered feedforward neural networks. In A. Nowak, editor, *Proceedings of the Ninth*

International Symposium on Dynamic Games and Applications, Adelaide, Australia, 2000.

[9] M. H. Breitner and H.-J. v. Mettenheim. Coarse-grained parallelization of the advanced neurosimulator FAUN 1.0 with PVM and the enhanced cornered rat game revisited. *International Game Theory Review*, 7:347–365, 2005.

[10] J. Dongarra, I. Foster, G. C. Fox, W. Gropp, K. Kennedy, L. Tonczon, and A. Whithe. *Sourcebook of Parallel Computing*. Morgan Kaufmann, Amsterdam, 2003.

[11] R. Fletcher. *Practical Methods of Optimization*. John Wiley & Sons, New York, 2000.

[12] G. A. Geist, A. L. Beguelin, J. Dongarra, W. C. Jiang, R. J. Manchek, and V. S. Sunderam. *PVM: Parallel Virtual Machine – A Users' Guide and Tutorial for Networked Parallel Computing*. MIT-Press, Cambridge, 1994.

[13] W. M. Getz. Capturability in a two target 'game of two cars'. *Journal of Guidance, Control and Dynamics*, 4(1):15–21, 1981.

[14] W. M. Getz and M. Pachter. Two-target pursuit-evasion differential games in the plane. *Journal of Optimization Theory and Applications*, 34(3):383–403, July 1981.

[15] I. Greenfeld. A differential game of surveillance evasion of two identical cars. *Journal of Optimization Theory and Applications*, 52(1):53–79, January 1987.

[16] W. Gropp, E. Lusk, and A. Skjellum. *Using MPI*. MIT-Press, Cambridge, 1999.

[17] X. Gui, Q. Wang, and D. Quian. A grid middleware for aggregative scientific computing libraries and parallel programming environments. In *Proceedings of the 6^{th} Asia-Pacific Web Conference, Hangzhou*, pp. 671–676. Springer, Heidelberg, 2004.

[18] A. J. G. Hey, G. Fox, and F. Berman. *Grid Computing*. John Wiley & Sons, Chichester, 2003.

[19] K. Hornik. Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 4:251–257, 1991.

[20] K. Hornik, M. Stinchcombe, and H. Whithe. Multi-layer feedforward networks are universal approximators. *Neural Networks*, 2:359–366, 1989.

[21] K. Hornik, M. Stinchcombe, and H. Whithe. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Networks*, 3:551–560, 1990.

[22] C. Hughes and T. Hughes. *Parallel and Distributed Programming Using C++*. Addison-Wesley, Boston, 2003.

[23] R. Isaacs. *Differential Games – A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization*. John Wiley & Sons, New York, 1965 – reprint.

[24] J. Joshy and C. Fellenstein. *Grid Computing*. Prentice Hall PTR, London, 2003.

[25] F. Köller and M. H. Breitner. Efficient synthesis of strategies with the advanced neurosimulator FAUN 1.0. In M. H. Breitner, editor, *Proceedings of the Fourth International ISDG Workshop, Goslar*, 2003.

[26] R. Lachner. *Echtzeitsynthese optimaler Strategien fÃ¼r Differentialspiele schneller Dynamik mit Anwendungen bei der Kollisionsvermeidung (in German)*. PhD thesis, Technische UniversitÃ¤t Clausthal, Clausthal-Zellerfeld, 1997.

[27] R. Lachner, M. H. Breitner, and H. J. Pesch. Real-time collision avoidance: Differential game, numerical solution, and synthesis of strategies. In J. A. Filar, V. Gaitsgory, and K. Mizukami, editors, *Advances in Dynamic Games and Applications*. Birkhäuser, Boston, 2000.

[28] J. P. Marec and N. V. Nhan. Two-dimensional pursuit-evasion game with penalty on turning rates. *Journal of Optimization Theory and Applications*, 23(2):305–345, October 1977.

[29] A. W. Merz. *The Homicidal Chauffeur – A Differential Game*. PhD thesis, Department of Aeronautics and Astronautics, Stanford University, Stanford, 1971.

[30] A. W. Merz. The game of two identical cars. *Journal of Optimization Theory and Applications*, 9(5):324–343, May 1972.

[31] H.-J. v. Mettenheim and M. H. Breitner. Neural network forecasting with high performance computers. In E. P. Hofer and E. Reithmeier, editors, *Proceedings of the Thirteenth International Workshop on Dynamics and Control*, pp. 33–40. Shaker, Aachen, 2005.

[32] H.-J. v. Mettenheim and M. H. Breitner. Distributed neurosimulation. In H.-D. Haasis, H. Kopfer, and J. SchÃ¶nberger, editors, *Operations Research Proceedings*. Springer, Heidelberg, 2006.

[33] A. Meyer, M. H. Breitner, and M. Kriesell. A pictured memorandum on synthesis phenomena occurring in the homicidal chauffeur game. In G. Martin-Herran and G. Zaccour, editors, *Proceedings of the Fifth International ISDG Workshop, Segovia*, pp. 17–32, 2005.

[34] T. Miloh. The game of two elliptical ships. *Optimal Control Applications & Methods*, 4, 1983.

[35] S. Mukherjee, J. Mustafi, and A. Chaudhuri. Grid computing: The future of distributed computing for high performance scientific and business applications. In *Distributed Computing. Mobile and Wireless Computing. 4th International Workshop, IWDC*. Springer, Heidelberg, 2002.

[36] B. Müller, J. Reinhardt, and M. T. Strickland. *Neural Networks*. Springer, Berlin, 1995.

[37] M. Pachter and W. M. Getz. The geometry of the barrier in the game of two cars. *Optimal Control Applications and Methods*, 1:103–118, 1980.

[38] J. Shinar and A. Davidovitz. New results in the game of two identical cars. In *Analysis and Optimization of Systems*, volume Volume 83/1986, pp. 759–780, 1986.

[39] J. Shinar and R. Tabak. New results in optimal missile avoidance analysis. *Journal of Guidance, Control & Dynamics*, 17, 1994.

[40] W. R. Stevens, B. Fenner, and A. M. Rudoff. *UNIX Network Programming - The Sockets Networking API*, volume 1. Addison-Wesley, Boston, 2004.

[41] W. R. Stevens and S. A. Rago. *Advanced Programming in the UNIX Environment*. Addison-Wesley, Upper Saddle River, NJ, 2005.

[42] A. E. Utemov. Numerical control optimization methods in one pursuit-evasion problem. *Journal of Computer and Systems Sciences International*, 45(3):395–412, May 2006.

[43] V. R. Vemuri. *Artificial Neural Networks*. IEEE Computer Society Press, Los Angeles, 1994.

[44] T. L. Vincent, E. M. Cliff, W. J. Grantham, and W. Y. Peng. Some aspects of collision avoidance. *AIAA Journal*, 12, 1974.

[45] H. Whithe, A. R. Gallant, K. Hornik, M. Stinchcombe, and J. Wooldridge. *Artificial Neural Networks – Approximation and Learning Theory*. Blackwell, Oxford, 1992.

[46] Y. Yavin and R. de Villiers. Game of two cars: Case of variable speed. *Journal of Optimization Theory and Applications*, 60(2):327–339, February 1989.

[47] Y. Yavin and R. de Villiers. Proportional navigation and the game of two cars. *Journal of Optimization Theory and Applications*, 62(3):351–369, September 1989.

Numerical Investigation of the Value Function for the Homicidal Chauffeur Problem with a More Agile Pursuer

Valerii S. Patsko
Institute of Mathematics and Mechanics
S. Kovalevskaya str. 16, 620219
Ekaterinburg, Russia
patsko@imm.uran.ru

Varvara L. Turova
Technical University of Munich
Boltzmannstr. 3, 85747
Garching, Germany
turova@ma.tum.de

Abstract

The paper is devoted to the investigation of a time-optimal differential game in which the pursuer possesses increased control capabilities comparing to the classical homicidal chauffeur problem. Namely, the pursuer steers not only the angular velocity of turn but can additionally change the magnitude of his linear velocity. For such a new variant of the dynamics with non-scalar control of the pursuer, a complete description of families of semipermeable curves is given and the dependence of the structure of level sets of the value function on a parameter that defines the bound on the magnitude of the pursuer's velocity is explored by numerical methods.

Key words. Time-optimal control, pursuit-evasion differential game, value function, semipermeable curves, homicidal chauffeur game, numerical construction.

AMS Subject Classifications. Primary 49N70, 49N75; Secondary 93B40.

1 Introduction

One of the most well-known model problems in the theory of differential games is the homicidal chauffeur problem proposed by R. Isaacs in 1951 (see [9]; [10]). The pursuer P (a car) and the evader E (a pedestrian) are moving in the plane. The dynamics are:

$$P: \quad \begin{aligned} \dot{x}_p &= w\sin\theta \\ \dot{y}_p &= w\cos\theta \\ \dot{\theta} &= w\,u/R, \quad |u| \le 1; \end{aligned} \qquad\qquad E: \quad \begin{aligned} \dot{x}_e &= v_1 \\ \dot{y}_e &= v_2 \\ |v| &\le \rho. \end{aligned} \tag{1}$$

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_13,

The linear velocity w of the car is constant. The angular velocity of rotation of the linear velocity vector is bounded, which means that the radius of turn of the car is bounded from below. The minimal turning radius is denoted by R.

The pedestrian is a non-inertial object that can change the value and direction of his velocity $v = (v_1, v_2)'$ instantaneously. The maximal possible value of the velocity is specified.

By a given circular neighborhood of his current geometric position, player P tries to capture player E as soon as possible.

The book by R. Isaacs [10] contains some elements of solution of the homicidal chauffeur problem. A complete solution to the problem is given in works by J. V. Breakwell and A. W. Merz (see [4]; [15]). Some other variants of differential games with the homicidal chauffeur dynamics are investigated in [12]; [5]; [13]; [2]; [6]. In many papers (see, e.g., [1], [19]), the homicidal chauffeur game is used as a test problem for evaluating the efficiency of algorithms for the computation of the value function and optimal strategies.

The dynamics like the ones of player P have very long history beginning with the paper by A. A. Markov [14]. In particular, the model is utilized when considering control problems related to the aircraft motion in a horizontal plane (see, e.g., [18]).

In papers on theoretical robotics, this model of dynamics is often referred to as Dubins' car because the paper [8] contains a theorem on the number and type of switches of the open-loop control that brings the object from a given state with a specified direction of the velocity vector to a terminal state for which the direction of the velocity vector is also prescribed.

By normalizing the time and geometric coordinates, one can always achieve $w = 1$, $R = 1$ in system (1). Therefore, system (1) takes the form

$$P: \quad \begin{aligned} \dot{x}_p &= \sin\theta \\ \dot{y}_p &= \cos\theta \\ \dot{\theta} &= u, \quad |u| \le 1; \end{aligned} \qquad\qquad E: \quad \begin{aligned} \dot{x}_e &= v_1 \\ \dot{y}_e &= v_2 \\ |v| &\le \nu. \end{aligned} \tag{2}$$

Recent theoretical works on controlled cars use intensively the following dynamics (see [20]; [11]):

$$\begin{aligned} \dot{x}_p &= w\sin\theta \\ \dot{y}_p &= w\cos\theta \\ \dot{\theta} &= u \\ |u| &\le 1, \quad |w| \le 1, \end{aligned} \tag{3}$$

in which the car has two controls. The first control u steers the forward motion direction $h = (\sin\theta, \cos\theta)'$ of the car. The second control w changes the magnitude of the linear velocity instantaneously. The velocity vector is directed along h for $w > 0$, and opposite to h for $w < 0$. If $w = 0$, the object remains immovable. Of course, instantaneous change of the velocity is a mathematical idealization. But following [20], p. 373, "for slowly moving vehicles, such as carts, this seems like a reasonable compromise to achieve tractability".

In papers on robotics, model (3) is called Reeds and Shepp's car. Time-optimal problems with Reeds and Shepp's dynamics were considered in [20], [24], [3], [22], [25], [23].

It seems quite natural to consider differential games where the pursuer has additional control capabilities comparing with dynamics (2). Namely, the pursuer can instantaneously change the value of his linear velocity within some bounds. In other words, Isaacs-Dubins' car turns into Reeds-Shepp's type car. The present paper is devoted to the investigation of such a differential game.

The purpose of this paper is to give a complete description of families of semipermeable curves for the problem considered and construct level sets of the value function by numerical methods. In the paper [16], a similar investigation is done by the authors for the classical homicidal chauffeur problem as well as for an "acoustic" variant of the dynamics and for a conic surveillance-evasion game. The present work continues the investigation in [16]. The difference is that, for the first time, non-scalar control of the pursuer in the homicidal chauffeur game is considered. This makes the families of semipermeable curves and the structure of level sets of the value function more complicated.

In papers on theoretical robotics, considerable attention is devoted to the analytical description of the boundary of reachable sets in the plane of geometric coordinates for Dubins' car model (i.e., $a = 1$) and for Reeds and Shepp's model (i.e., $a = -1$) (see [7], [21], [25], and references herein). Results on the construction of level sets of the value function presented in the paper can be interpreted as a description of time-limited game reachable sets for a car with a given range of the instantaneous change of the linear velocity magnitude. During the process of motion, the car is subjected to a dynamic non-inertia disturbance that can distort the velocity vector within prescribed bounds. If this disturbance becomes vanishing small, the game reachable sets are transformed into usual reachable sets for the control problem in the absence of disturbances. But even for $a = 1$ and $a = -1$ the analytical description of these sets is not simple. The authors do not know any papers that study reachable sets for $a \neq \pm 1$, and all the more so when dynamic disturbances are presented.

2 Statement of the homicidal chauffeur game with a more agile pursuer

Let the dynamics in original coordinates be:

$$
\begin{aligned}
P:\quad & \dot{x}_p = w \sin\theta & \qquad E:\quad & \dot{x}_e = v_1 \\
& \dot{y}_p = w \cos\theta & & \dot{y}_e = v_2 \\
& \dot{\theta} = u & & \\
& |u| \le 1,\ a \le w \le 1; & & |v| \le \nu.
\end{aligned}
\tag{4}
$$

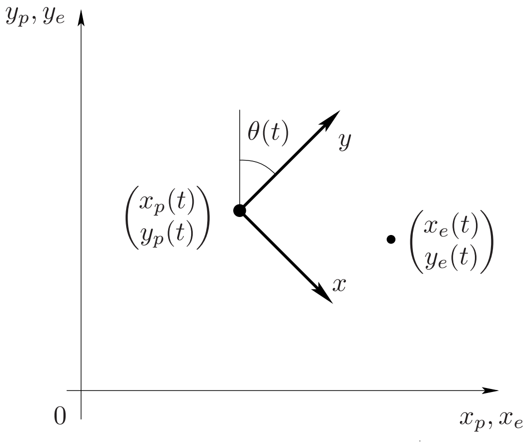

Figure 1: Movable reference system.

The constraint on the value of the angular velocity u does not depend on the magnitude of the linear velocity. The value w of the linear velocity can change instantaneously within the bounds $a \le w \le 1$, where a is a fixed parameter that fulfills the condition $a \in [-1, 1]$. The parameter ν in the constraint on the magnitude $|v|$ of the velocity $v = (v_1, v_2)'$ of player E is assumed to be from the interval $[0, 1)$.

The objective of player P is to capture player E in a closed circular neighborhood of radius r around his geometrical state as soon as possible.

Let us apply the reduced coordinates from ([10], pp. 29–30) which determine the relative position of player E in the rectangular coordinate system x, y with the origin at P and the axis y directed along the vector h, i.e., along the velocity vector of P computed for $w = 1$. The reduced coordinate system is explained in Fig. 1.

The dynamics (4) are transformed into:

$$\begin{aligned} &\dot{x} = -y\,u + v_x \\ &\dot{y} = \ \ x\,u - w + v_y \\ &|u| \le 1, \ \ a \le w \le 1, \ \ v = (v_x, v_y)', \ \ |v| \le \nu. \end{aligned} \tag{5}$$

In the reduced coordinates, the objective of player P is to bring the state (x, y) to a given terminal set M being a circle of the radius r with the center at the origin.

If $a = 1$, the dynamics (5) coincides with the homicidal chauffeur ones.

In the paper, numerically obtained level sets $W_\tau = \{(x, y) : \ V(x, y) \le \tau\}$ of the value function $V(x, y)$ of the time-optimal differential game with dynamics (5) will be presented. The structure of these sets depending on the parameter a will be explored.

The level sets are computed using an algorithm for solving time-optimal differential games developed by the authors [16].

3 Families of semipermeable curves

When solving differential games in the plane, it is useful to carry out a preliminary study of families of smooth semipermeable curves that are determined by the dynamics of the controlled system. The knowledge of these families for time-optimal problems allows to verify the correctness of the construction of barrier lines on which the value function is discontinuous.

A smooth semipermeable curve is a line with the following preventing property: one of the players can prevent crossing the curve from the positive side to the negative one, the other player can prevent crossing the curve from the negative side to the positive one.

Below, the notations

$$\mathcal{P} = \{(u,w):\ |u| \le 1,\, w \in [a,1]\}, \qquad \mathcal{Q} = \{v : |v| \le \nu\}$$

are used for the constraints on the controls of players P and E.

A. Consider the Hamiltonian

$$H(\ell, z) = \min_{\xi\in\mathcal{P}} \max_{v\in\mathcal{Q}} \ell' f(z,\xi,v), \quad z,\, \ell \in R^2. \tag{6}$$

Here,

$$f(z,\xi,v) = p(z)u + gw + v,$$

$$z = (x,y)', \quad \xi = (u,w)', \quad p(z) = (-y,x)', \quad g = (0,-1)'.$$

We study nonzero roots of the equation $H(\ell,z) = 0$, where $z \in R^2$ is fixed. Since the function $\ell \to H(\ell,z)$ is positively homogeneous, it is convenient to assume that $\ell \in \mathcal{E}$, where $\mathcal{E}$ is the circumference of unit radius centered at the origin.

Let $\ell,\, \ell_* \in \mathcal{E},\ \ell \ne \ell_*$. The notation $\ell \prec \ell_*$ ($\ell \succ \ell_*$) means that the vector ℓ can be obtained from the vector ℓ_* using a counterclockwise (clockwise) rotation through an angle smaller than π. In fact, this order relation will be used only for vectors that are sufficiently close to each other.

Fix $z \in R^2$ and consider roots of the equation $H(\ell,z) = 0,\ \ell \in \mathcal{E}$. A vector ℓ_* is called the strict root "$-$" to "$+$" if there exist a vector $\kappa \in R^2$ and a neighborhood $S \subset \mathcal{E}$ of the vector ℓ_* such that $H(\ell_*,z) = \ell_*'\kappa = 0$ and $H(\ell,z) \le \ell'\kappa < 0$ ($H(\ell,z) \ge \ell'\kappa > 0$) for vectors $\ell \in S$ satisfying the relation $\ell \prec \ell_*$ ($\ell \succ \ell_*$). Similarly, the strict root "$+$" to "$-$" is defined through replacing $\ell \prec \ell_*$ ($\ell \succ \ell_*$) by $\ell \succ \ell_*$ ($\ell \prec \ell_*$). The roots "$-$" to "$+$" and "$+$" to "$-$" are called roots of the first and second type, respectively. In the following, when utilizing the notation $\ell \prec \ell_*$ ($\ell \succ \ell_*$) we will keep in mind that ℓ is from a neighborhood S like that mentioned in the definition of the roots.

Denote by $\Xi(\ell,z)$ the collection of all $\xi \in \mathcal{P}$ that provide the minimum in (6), that is:

$$\Xi(l,z) = \operatorname{argmin}\{\ell'(p(z)u + gw) : \xi \in \mathcal{P}\}.$$

If ℓ_* is a strict root of the first (second) type, take $\xi^{(1)}(\ell_*, z)$ $(\xi^{(2)}(\ell_*, z))$ equal to argmin$\{\ell'(p(z)u + g\,w) : \xi \in \mathcal{P}(\ell_*, z)\}$, where $\ell \prec \ell_*$ $(\ell \succ \ell_*)$. Note that the result does not depend on the choice of ℓ. Let

$$v^{(1)}(\ell_*) = v^{(2)}(\ell_*) = \text{argmax}\{\ell_*' v : v \in \mathcal{Q}\}.$$

Since $\mathcal{Q}$ is a circle, $v^{(i)}(\ell_*)$, $i = 1, 2$, is a singleton.

If ℓ_* is a root of the first type, consider the vectograms $f(z, \xi^{(1)}(\ell_*, z), \mathcal{Q})$, and $f(z, \mathcal{P}, v^{(1)}(\ell_*))$. We have

$$\ell' f(z, \xi^{(1)}(\ell_*, z), v) \leq \max_{v \in \mathcal{Q}} \ell' f(z, \xi^{(1)}(\ell_*, z), v) = H(\ell, z) \leq \ell'\kappa < 0 \qquad (7)$$

for $v \in \mathcal{Q}$ and $\ell \prec \ell_*$. For $\ell \succ \ell_*$, it holds:

$$\max_{v \in \mathcal{Q}} \ell' v \leq \ell' v^{(1)}(\ell_*) + \ell'\kappa/2$$

because κ is orthogonal to ℓ_* (and hence to $v^{(1)}(\ell_*)$), and $\ell'\kappa > 0$. Thus,

$$H(\ell, z) \leq \min_{\xi \in \mathcal{P}} \ell' f(z, \xi, v^{(1)}(\ell_*)) + \ell'\kappa/2.$$

The last inequality yields

$$\ell' f(z, \xi, v^{(1)}(\ell_*)) \geq \min_{\xi \in \mathcal{P}} \ell' f(z, \xi, v^{(1)}(\ell_*)) \geq H(\ell, z) - \ell'\kappa/2 \geq \ell'\kappa/2 > 0 \quad (8)$$

for $\xi \in \mathcal{P}$ and $\ell \succ \ell_*$.

Relations (7) and (8) ensure that the vectograms $f(z, \xi^{(1)}(\ell_*, z), \mathcal{Q})$, and $f(z, \mathcal{P}, v^{(1)}(\ell_*))$ do not contain zero and are located with respect to the direction of the vector $f^{(1)} = f(z, \xi^{(1)}(\ell_*, z), v^{(1)}(\ell_*))$, as it is shown in Fig. 2a.

Therefore, the existence of a strict root ℓ_* of the first type at a point z ensures together with taking the control $\xi^{(1)}(\ell_*, z)$ $(v^{(1)}(\ell_*))$ by player P (E) that the velocity vector $f(z, \xi^{(1)}(\ell_*, z), v)$ $(f(z, \xi, v^{(1)}(\ell_*)))$ is directed to the right (to the left) with respect to the direction of the vector $f^{(1)}$ for any control v (ξ) of player E (P). Such a disposition of the vectograms means that player P (E) guarantees the trajectories do not go to the left (to the right) with respect to the direction of $f^{(1)}$. The direction of the vector $f^{(1)}$ is called the semipermeable direction of the first type. The vector $f^{(1)}$ is orthogonal to the vector ℓ_*, its direction can be obtained from ℓ_* by a clockwise rotation through the angle $\pi/2$.

Arguing similarly, we obtain that the existence of a strict root ℓ_* of the second type at a point z ensures together with taking the control $\xi^{(2)}(\ell_*, z)$ $(v^{(2)}(\ell_*))$ by player P (E) that the velocity vector $f(z, \xi^{(2)}(\ell_*, z), v)$ $(f(z, \xi, v^{(2)}(\ell_*)))$ is directed to the left (to the right) with respect to the direction of the vector $f^{(2)} = f(z, \xi^{(2)}(\ell_*, z), v^{(2)}(\ell_*))$ for any control of player E (P). This means that player P (E) guarantees the trajectories do not go to the right (to the left) with respect to the direction of $f^{(2)}$. The direction of the vector $f^{(2)}$ is called

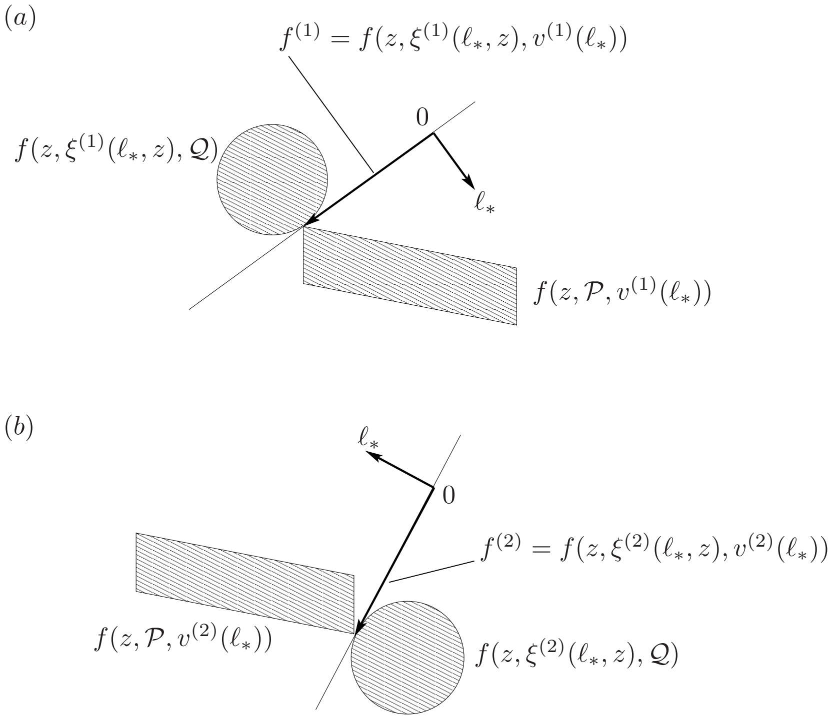

Figure 2: (a) Semipermeable direction of the first type. (b) Semipermeable direction of the second type.

the semipermeable direction of the second type. The vector $f^{(2)}$ is orthogonal to the vector ℓ_*, its direction can be obtained from ℓ_* by a counterclockwise rotation through the angle $\pi/2$ (see Fig. 2b).

Thus, there is a significant difference in the location of the vectograms for strict roots of the first and second type.

B. We distinguish semipermeable curves of the first and second types. A smooth curve is called semipermeable of the first (second) type if the direction of the tangent vector at any point along this curve is the semipermeable direction of the first (second) type.

The side of a semipermeable curve that player P (E) can keep is called positive (negative). The positive (negative) side of a semipermeable curve of the first type is on the right (on the left) when looking along the semipermeable direction. The opposite is valid for semipermeable curves of the second type.

Figure 3 illustrates the role of semipermeable curves of the first and second type in solving a game of kind with the classical homicidal chauffeur dynamics, the restriction $\mathcal{Q}$ is of a rather large radius. The objective of player P is to bring

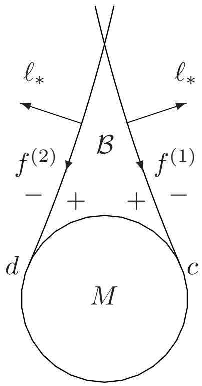

Figure 3: Semipermeable curves of the first and second type and capture set.

trajectories to the terminal set M which is a circle centered at the origin, the objective of player E is opposite. Denote by c and d the endpoints of the usable part [10] on the boundary of M. The semipermeable curves of the first and second type that are tangent to the set M and pass through the points c and d, respectively, define completely the capture set $\mathcal{B}$ of all points for which guaranteed time of attaining M is finite. The curves are faced towards each other with the positive sides.

C. In Fig. 4, the families of semipermeable curves for the classical homicidal chauffeur dynamics are presented. There are families $\Lambda^{(1),1}$ and $\Lambda^{(1),2}$ of the first type and families $\Lambda^{(2),1}$ and $\Lambda^{(2),2}$ of the second type. The second upper index in the notation $\Lambda^{(i),j}$ indicates those of two extremal values of control u that corresponds to this family: $j = 1$ is related to curves which are trajectories for $u = 1$; $j = 2$ is related to curves which are trajectories for $u = -1$. The arrows show the direction of motion in reverse time. Due to symmetry properties of the dynamics, all families can be obtained from only one of them (for example, $\Lambda^{(1),1}$) by means of reflections about the horizontal and vertical axes.

The construction of mentioned four families of smooth semipermeable curves can be explained as follows.

Let $a = 1$ in system (5). Assign the set

$$B_* = \{(x, y) : \; -y + v_x = 0, \; x - 1 + v_y = 0, \; v \in \mathcal{Q}\}$$

to the control $u = 1$, and the set

$$A_* = \{(x, y) : \; y + v_x = 0, \; -x - 1 + v_y = 0, \; v \in \mathcal{Q}\}$$

to the control $u = -1$. Hence, B_* is the set of all points in the plane x, y such that the velocity vector of system (5) for $w = 1$ vanishes at $u = 1$ and some $v \in \mathcal{Q}$. We have $A_* = -B_*$.

Consider two tangents to the sets A_*, B_* passing through the origin (see Fig. 5), and mark arcs $a_1a_2a_3$ and $b_1b_2b_3$ on ∂A_*, and ∂B_*, respectively.

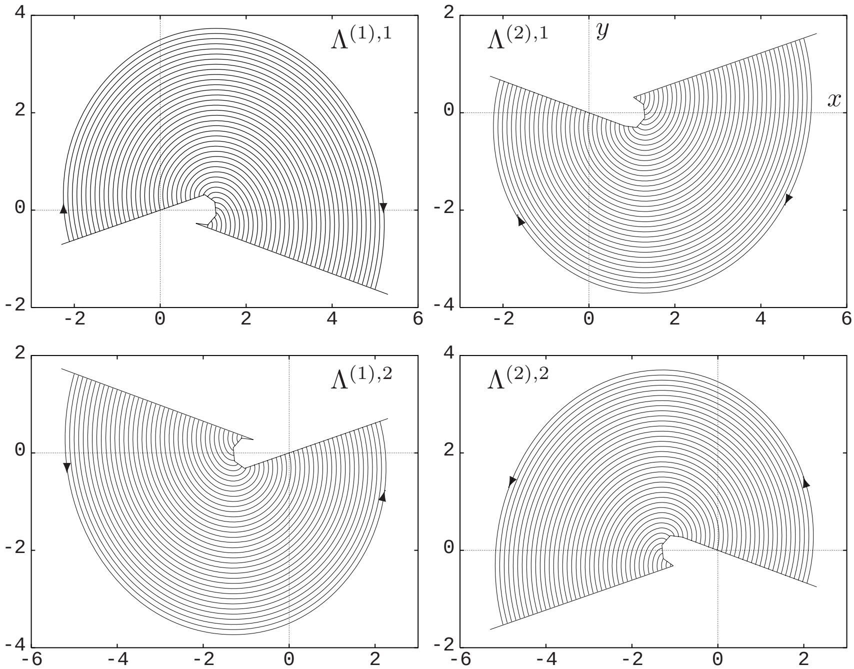

Figure 4: Families of smooth semipermeable curves for the classical homicidal chauffeur dynamics.

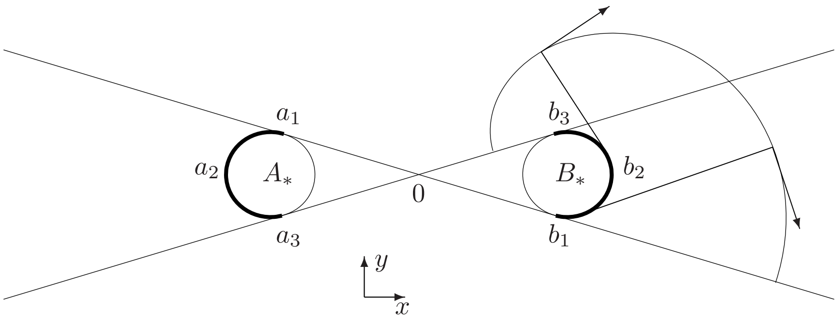

Figure 5: Auxiliary arcs generating the families of smooth semipermeable curves for the classical homicidal chauffeur dynamics.

Attach an inextensible string of a fixed length to the point b_1 and wind it up on the arc $b_1b_2b_3$. Then wind the string down keeping it taut in the clockwise direction. The end of the string traces an involute, which is a semipermeable curve of the family $\Lambda^{(1),1}$. A detailed proof of this fact is given in [17]. The complete family $\Lambda^{(1),1}$ is obtained by changing the length of the string.

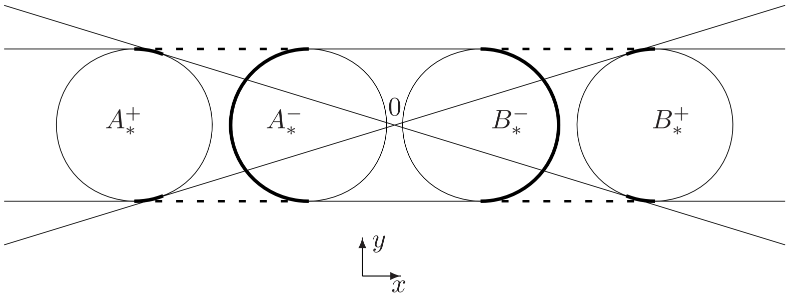

Figure 6: Auxiliary arcs generating the families of smooth semipermeable curves for dynamics (5); $a = 0.33$, $\nu = 0.3$.

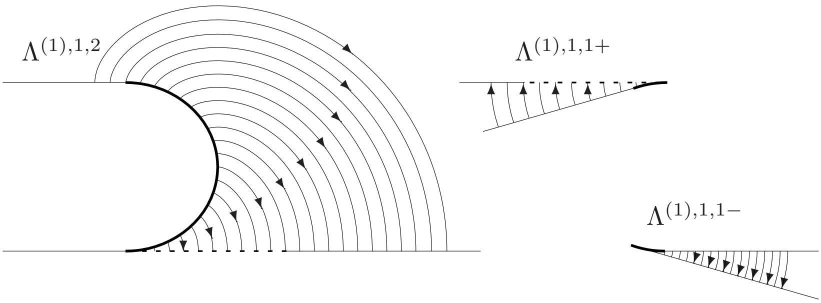

Figure 7: Three families of smooth semipermeable curves of the first type corresponding to $u = 1$ for dynamics (5); $a \geq -\nu$, $\nu = 0.3$.

The family $\Lambda^{(2),2}$ is obtained as the collection of the counterclockwise-involutes of the arc $a_1a_2a_3$ by attaching the string to the point a_3 .

The family $\Lambda^{(2),1}$ is generated by the clockwise-involutes of the arc $b_1b_2b_3$ provided the string is attached to the point b_3.

The family $\Lambda^{(1),2}$ is composed of the counterclockwise-involutes of the arc $a_1a_2a_3$ provided the string is attached to the point a_1.

D. Families of semipermeable curves corresponding to dynamics (5) are arranged in a more complicated way (see Figs. 6–10).

Figure 6 shows auxiliary lines used for the construction of the families. Here, A_*^+ and B_*^+ are the circles of radius ν with the centers at the points $(-1, 0)$ and $(1, 0)$, respectively. The circle A_*^+ (B_*^+) consists of all points (x, y) such that the right-hand side of system (5) becomes zero for $u = -1$ $(u = 1)$, $w = 1$, and some $v \in \mathcal{Q}$. The only difference in the definition of the circle A_*^- (B_*^-) is that $w = a$ instead of $w = 1$ is used. For Fig. 6, $a = 0.33$, $\nu = 0.3$. Six thick arcs on the boundaries of the circles are the only lines utilized for the construction of

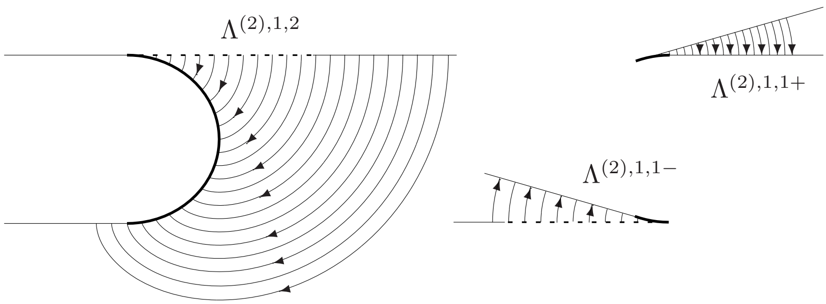

Figure 8: Three families of smooth semipermeable curves of the second type corresponding to $u = 1$ for dynamics (5); $a \geq -\nu$, $\nu = 0.3$.

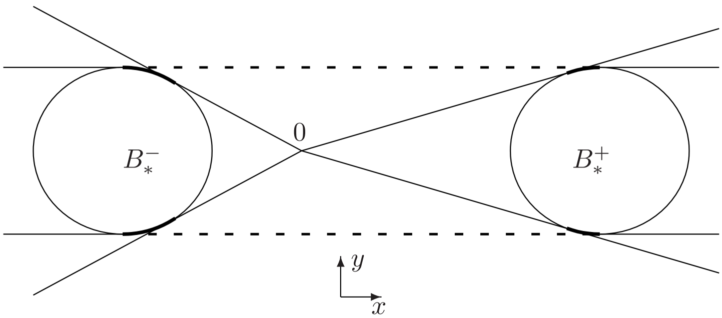

Figure 9: Auxiliary arcs generating the families of smooth semipermeable curves corresponding to $u = 1$ for dynamics (5); $a = -0.6$, $\nu = 0.3$.

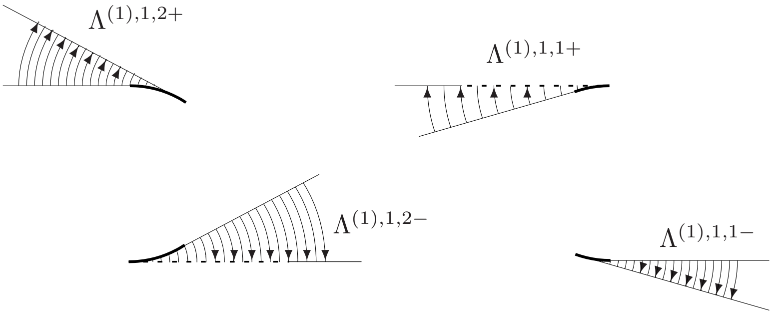

Figure 10: Four families of smooth semipermeable curves of the first type corresponding to $u = 1$ for dynamics (5); $a = -0.6$, $\nu = 0.3$.

families of smooth semipermeable curves.

In the problem considered, for $a \geq -\nu$, there are three families $\Lambda^{(1),1,1+}$, $\Lambda^{(1),1,1-}$, and $\Lambda^{(1),1,2}$ of semipermeable curves of the first type that correspond to $u = 1$ (instead of one family in the classical homicidal chauffeur problem). The families are drawn in Fig. 7. Every one of these three families consists of the involutes of one of the three arcs. The families $\Lambda^{(1),1,1+}$ and $\Lambda^{(1),1,2}$ overlap. There is also an overlap between the families $\Lambda^{(1),1,1-}$ and $\Lambda^{(1),1,2}$ over the part of the family $\Lambda^{(1),1,1-}$ in the region above the line $y = -\nu$. If $a \to 1$, the union of three families gives the family $\Lambda^{(1),1}$ of the classical homicidal chauffeur problem.

Figure 8 shows the families $\Lambda^{(2),1,1+}$, $\Lambda^{(2),1,1-}$, and $\Lambda^{(2),1,2}$ of the second type corresponding to $u = 1$. They are obtained by the reflection of corresponding families from Fig. 7 about the horizontal axis and by changing the direction of arrows. Families $\Lambda^{(2),2,1+}$, $\Lambda^{(2),2,1-}$, and $\Lambda^{(2),2,2}$ can be obtained by the reflection of corresponding families from Fig. 7 about the vertical axis (the direction of arrows does not change).

Families $\Lambda^{(1),2,1+}$, $\Lambda^{(1),2,1-}$, and $\Lambda^{(1),2,2}$ of the first type can be obtained by the reflection of families $\Lambda^{(2),1,1+}$, $\Lambda^{(2),1,1-}$, and $\Lambda^{(2),1,2}$ about the vertical axis.

For $-1 \leq a < -\nu$, the family $\Lambda^{(1),1,2}$ splits into two families: $\Lambda^{(1),1,2-}$ and $\Lambda^{(1),1,2+}$. Everything else is similar to the case $a \geq -\nu$. The lines that define the families of smooth semipermeable curves corresponding to $u = 1$ are depicted in Fig. 9. In Fig. 10, four families of the first type for $u = 1$ are presented.

4 Level sets of the value function

In this section, results of the computation of level sets W_τ of the value function $V(x, y)$ for time-optimal differential game with dynamics (5) will be presented. The collection of all points $(x, y) \in \partial W_\tau$ such that $V(x, y) = \tau$ is called the front corresponding to the reverse time τ. The computational procedure [16] for the construction of the level sets runs backward in time on the interval $[0, \tau_f]$. The value of τ_f is given below in the figure captions. For all figures, the horizontal axis is x, the vertical axis is y.

It is supposed in all examples that the set M is a circle of radius $r = 0.3$ with the center at the origin, and the control u is bounded as $|u| \leq 1$.

The following 6 variants of parameter values will be considered:

$$1.\ a{=}1, \quad 2.\ a = 0.25, \quad 3.\ a = -0.1, \quad 4.\ a{=}-0.4, \quad 5.\ a = -0.6, \quad 6.\ a = -1.$$

The variants are ordered by decreasing value of the parameter a. Remind that the constraint on the control w of player P is $a \leq w \leq 1$. Therefore, the capabilities of player P increase with the decrease of value a.

In every variant, the constraint on the control of player E is $|v| \leq \nu = 0.3$.

In the computations, the circle M is approximated by a regular 30-polygon, the circular constraint of player E is replaced by a regular octagon.

A. Let now describe level sets of the value function.

1) Figure 11 presents level sets W_τ for the classical homicidal chauffeur problem, variant 1. The value function is discontinuous on two barrier lines: the right barrier line is a semipermeable curve of the family $\Lambda^{(1),1}$, the left one is a semipermeable curve of the family $\Lambda^{(2),2}$ (see Fig. 4). The right barrier terminates at the lower tangent to the circle B_* passing through the origin, the left one terminates at the lower tangent to the circle A_*. After the termination, the barriers are continued by the lines formed of the corner points on the fronts of level sets. The value function is not differentiable on these lines.

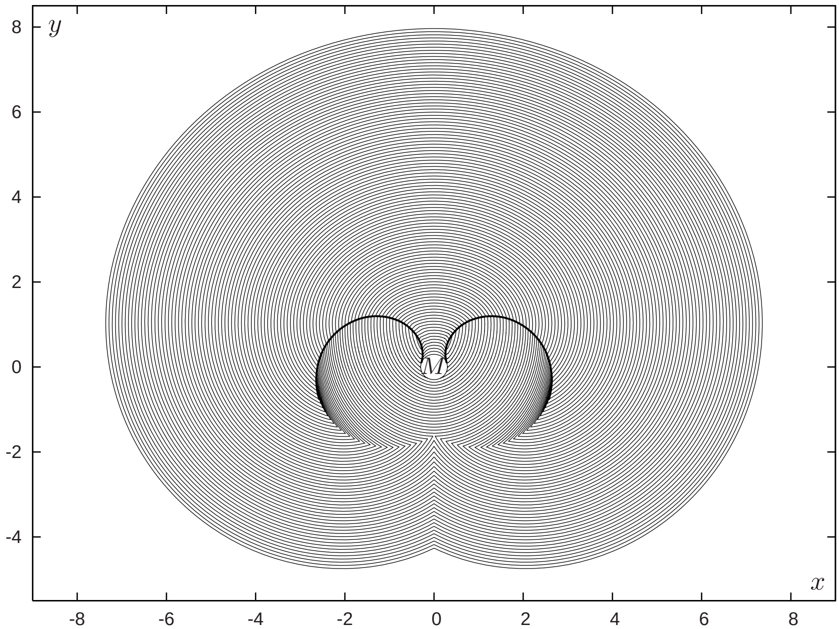

Figure 11: Level sets W_τ for the classical homicidal chauffeur problem; $\tau_f = 10.3$.

We can conclude about lines of discontinuity of the value function by analyzing the families of semipermeable curves before the construction of level sets of the value function. The knowledge of these families is utilized for the verification of numerical computation of level sets.

After bending round the right and left barriers, the right and left parts of the front meet on the vertical axis at time $\tau = 7.82$. A closed front occurs and a hole is generated which is completely filled out at time $\tau = 10.3$.

2) If $a \geq -\nu = -0.3$, then there is only one usable part on the boundary ∂M for the chosen radius $r = 0.3$ of the terminal circle M. It is located in the upper part of ∂M. Its right (left) endpoint is the intersection point of ∂M with the upper

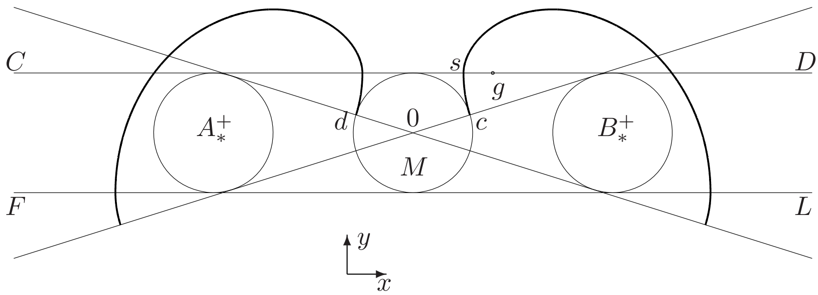

Figure 12: Long barrier emanating from the point c.

tangent to B_*^+ (A_*^+) passing through the origin. In Fig. 12, c and d are the endpoints of the usable part (from which the fronts are emitted in the backward procedure).

For chosen values $r = 0.3$ and $\nu = 0.3$, there exists a value $\bar{a} \in (0, 1)$ of the parameter a which separates the case of a "long" barrier from the case of a "short" one. The barrier lines emanate from the points c and d of the usable part. Due to the symmetry of the solution with respect to the vertical axis, one can only consider the right barrier emanating from the point c. The right long barrier is a semipermeable curve obtained by a smooth junction of a curve of the family $\Lambda^{(1),1,1+}$ going backward in time from the point c with a curve of the family $\Lambda^{(1),1,2}$ and, later on, with a curve of the family $\Lambda^{(1),1,1-}$ (see Fig. 7). The first junction can occur on the horizontal line $y = \nu = 0.3$ (denote it by CD), the second one is on the horizontal line $y = -\nu = -0.3$ (denote it by FL). The right and left long barriers are shown in Fig. 12. The right short barrier consists of a curve of the family $\Lambda^{(1),1,1+}$ going backward in time from the point c up to the line CD.

If the intersection point s of the last-mentioned semipermeable curve with the line CD is to the left from the point $g = (a, \nu)'$, being a tangent point of the line CD and the set B_*^-, then the barrier can be continued by a curve of the family $\Lambda^{(1),1,2}$. Therefore, we obtain a long barrier. If the point s is to the right from g, there is no any extension. In this case, a short barrier occurs. The coincidence $s = g$ defines a critical value $\bar{a}$. The transient case $a = \bar{a}$ is of great theoretical importance.

The value $a = 0.25$ in variant 2 is close to $\bar{a}$. The results of the computation of sets W_τ for this variant are presented in Fig. 13. The obtained right and left barriers are long. Since in the computation the constraint $|v| \leq \nu$ is a regular octagon, the sets B_*^+, B_*^-, A_*^+, and A_*^- are octagons as well. Figure 14 shows an enlarged fragment of Fig. 13 with additionally drawn sets A_*^+ and A_*^-.

3) Figure 15 presents computation results for variant 3. Enlarged fragments of Fig. 15 are shown in Figs. 16 and 17.

Here, the right barrier that emanates from the right endpoint c of the usable part on ∂M belongs to the family $\Lambda^{(1),1,1+}$. In the notations of Fig. 12, the point s is to the right of the point g that results in the short barrier. The left barrier

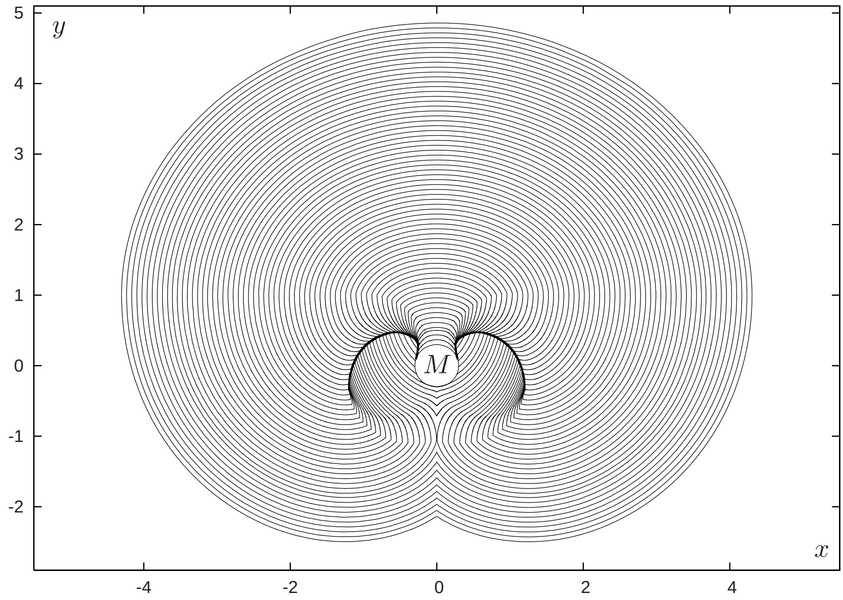

Figure 13: Level sets W_τ for differential game (5); $a = 0.25$, $\tau_f = 6.7$.

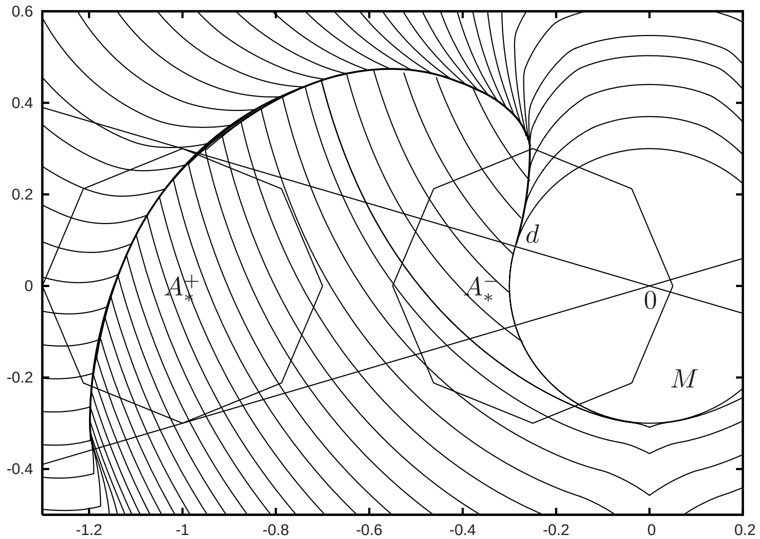

Figure 14: Enlarged fragment of Fig. 13.

which is symmetrical to the right one with respect to the vertical axis belongs to the family $\Lambda^{(2),2,1+}$. The termination of barriers on the horizontal line $y = 0.3$

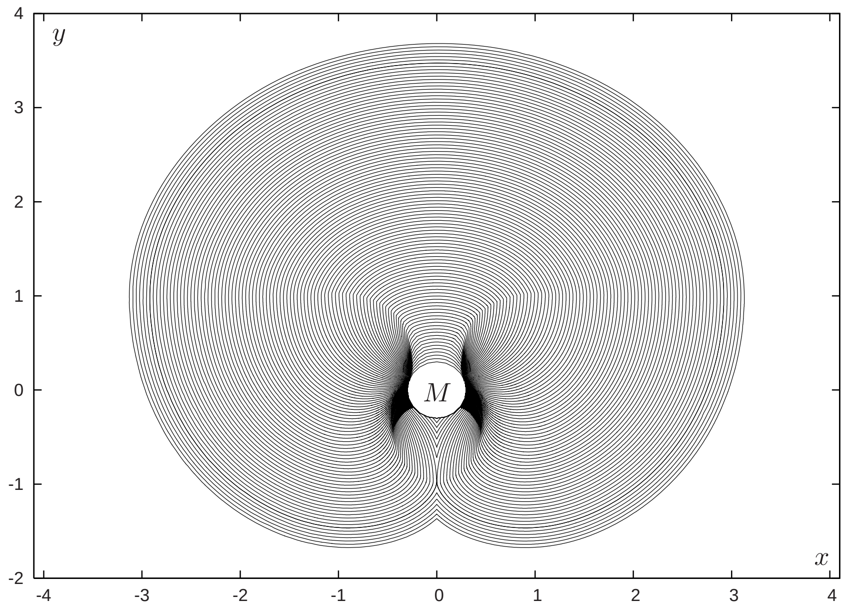

Figure 15: Level sets W_τ for differential game (5); $a = -0.1$, $\tau_f = 4.89$.

is completely in accordance with the results that follow from the knowledge of families of semipermeable curves.

After bending round the barrier lines, the ends of the front go down along the negative sides of the barriers and then move along the boundary of the terminal set. This is well seen in Figs. 16 and 17. Parts of the front near the right and left sides of the terminal set and far from it move with different velocities, which yields the generation of the right and left singular lines where the value function is non-differentiable.

Values τ_f given in the captions of Figs. 11, 13, and 15 correspond to the time instants at which the “inner hole” is filled out with the fronts.

4) Consider variant 4 (see Figs. 18 and 19). For this variant, two usable parts on ∂M arise. The upper usable part is larger than the lower one.

The right and left barriers emanating from the endpoints of the upper usable part are semipermeable curves of the families $\Lambda^{(1),1,1+}$ and $\Lambda^{(2),2,1+}$, respectively. They terminate, as expected, on the horizontal line $y = 0.3$. Outside set M, the value function is discontinuous on these two lines only.

The ends of the fronts which propagate from the lower usable part move very slowly along ∂M (see Fig. 20a). At some time, a local non-convexity with a kink is formed on a going down front. This kink rises upwards and after some time it

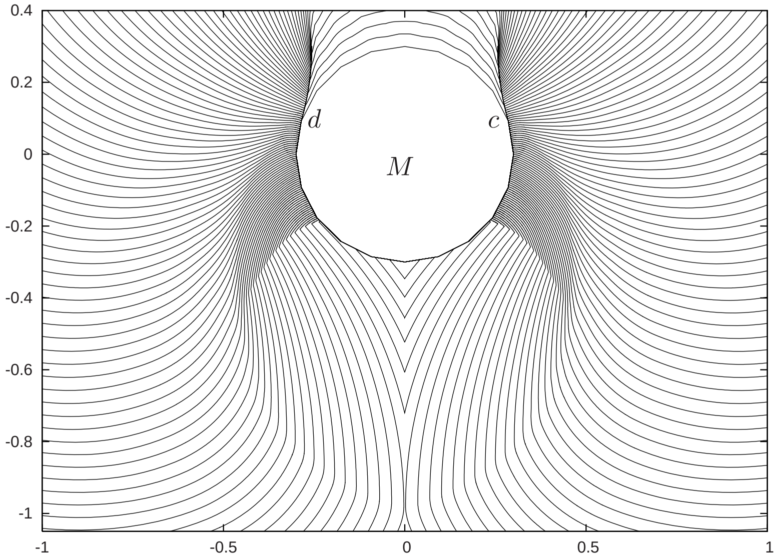

Figure 16: Enlarged fragment of Fig. 15.

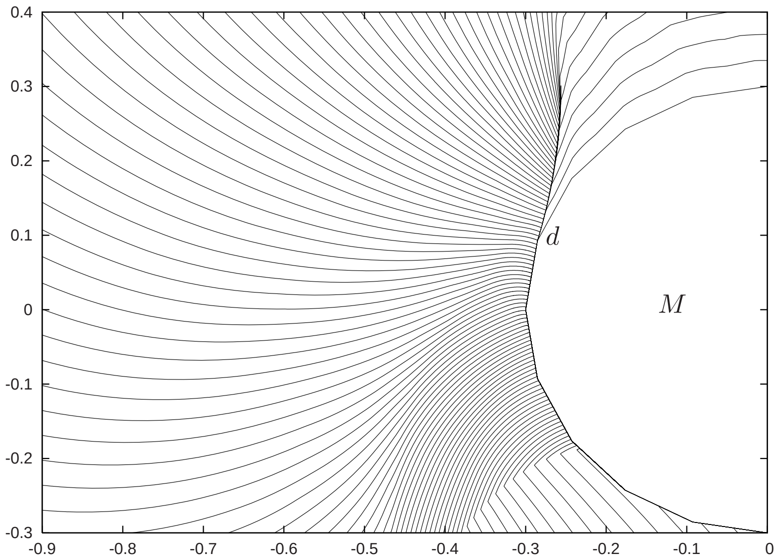

Figure 17: One more enlarged fragment of Fig. 15.

reaches the endpoint of the front on ∂M. In fact, a new endpoint of the front that moves along ∂M faster than before arises. Then, after some time, this endpoint

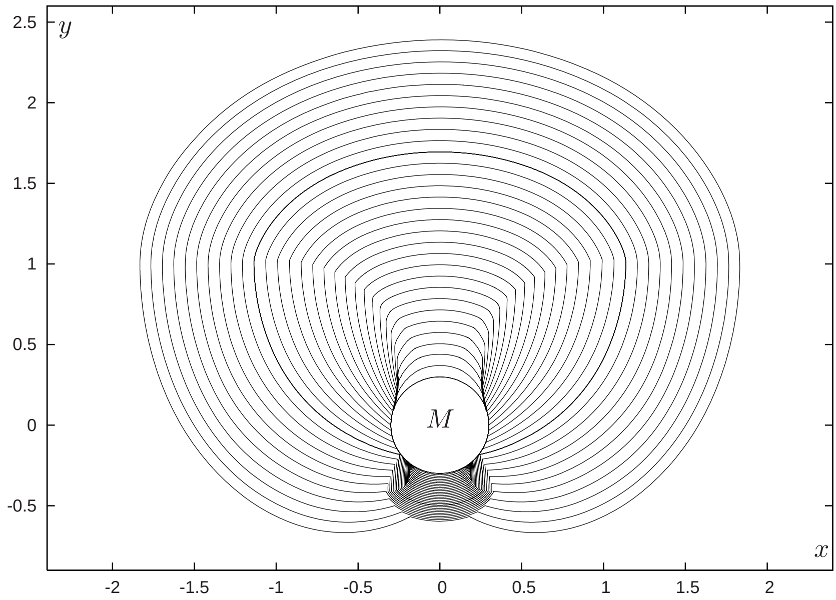

Figure 18: Level sets W_τ for differential game (5); $a = -0.4$, $\tau_f = 3$.

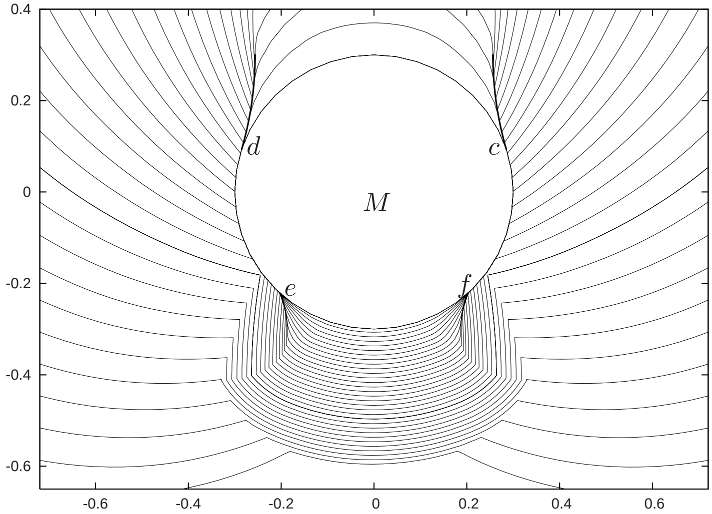

Figure 19: Enlarged fragment of Fig. 18.

meets the endpoint of the front going down along ∂M. A closed front is formed (Fig. 19).

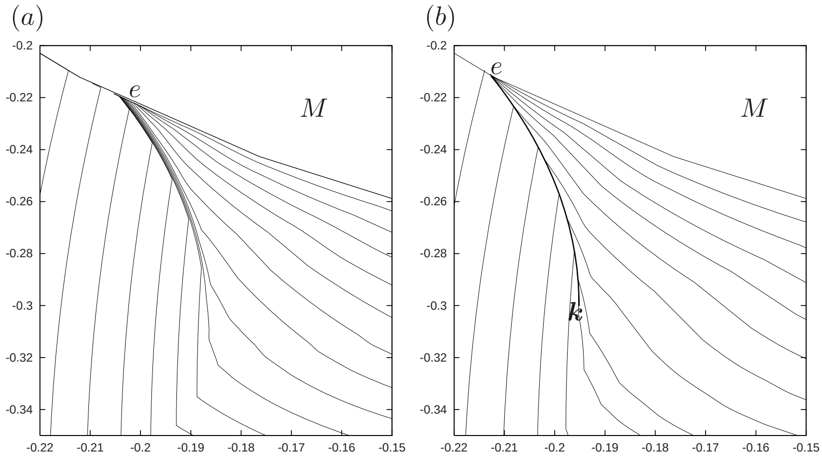

Figure 20: Continuous and discontinuous value function. (a) Enlarged fragment of Fig. 19, $a = -0.4$. The region of fronts accumulation is shown. The value function is continuous outside set M. (b) Level sets for $a = -0.425$. The value function is discontinuous outside set M on the barrier line ek.

5) For $a \leq \tilde{a} = -0.3\sqrt{2}$, there are barrier lines that emanate from the endpoints of the lower usable part and terminate at the horizontal straight line $y = -\nu = -0.3$. The left barrier line is a semipermeable curve of the family $\Lambda^{(1),1,2-}$, the right one is a semipermeable curve of the family $\Lambda^{(2),2,2-}$.

Figure 20b computed for $a = -0.425$ shows the lower fronts bending round the left barrier line. This can be compared with Fig. 20a computed for $a = -0.4$. At first sight, the pictures are very similar. However, the value function in Fig. 20b is discontinuous on the barrier line ek, while the value function in Fig. 20a is continuous in a similar region below the set M.

The behavior of lower fronts for values a close to $\tilde{a}$ initiates the following question which, to our opinion, can hardly be answered based on numerics only. Is it true that for $a = -0.3\sqrt{2}$ the endpoints of the lower fronts stay at the endpoints e, f of the usable part till some time $\hat{\tau} > 0$ and only for $\tau > \hat{\tau}$ begin to go down along the barrier lines?

Figures 21 and 22 show level sets of the value function for $a = -0.6$ (variant 5). Fronts generated by the upper and lower usable parts encounter at some time $\tau_* = 1.41$ (see Fig. 22). For $\tau > \tau_*$, the computation continues from the closed front.

6) For $a = -1$ (variant 6 being shown in Fig. 23), the computed sets are symmetrical with respect to both horizontal axis x and vertical axis y. Two short upper barrier lines like in variants 3, 4, and 5 are present. The lower barrier lines are symmetrical to the upper ones with respect to the axis x.

Note that, for $\nu = 0$, the set W_τ coincides with the reachable set $G(t, M)$ of

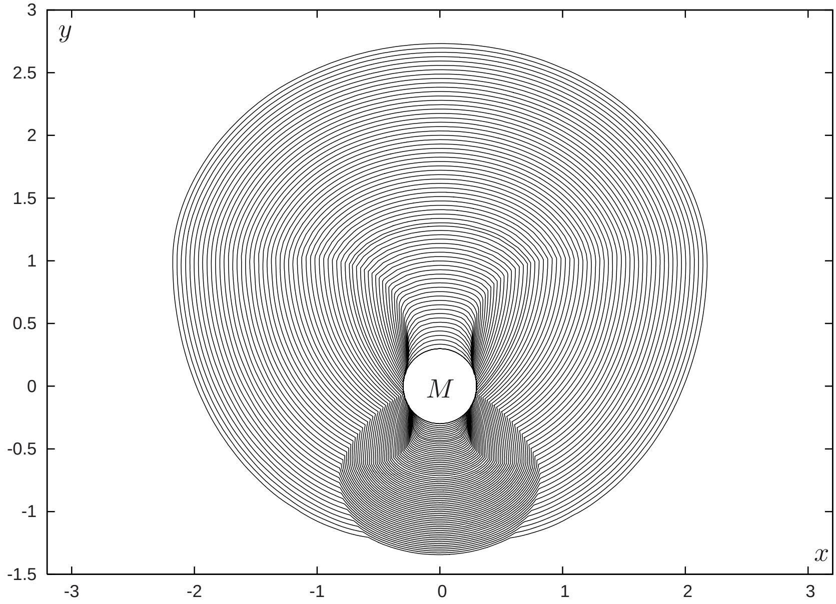

Figure 21: Level sets W_τ for differential game (5); $a = -0.6$, $\tau_f = 3.5$.

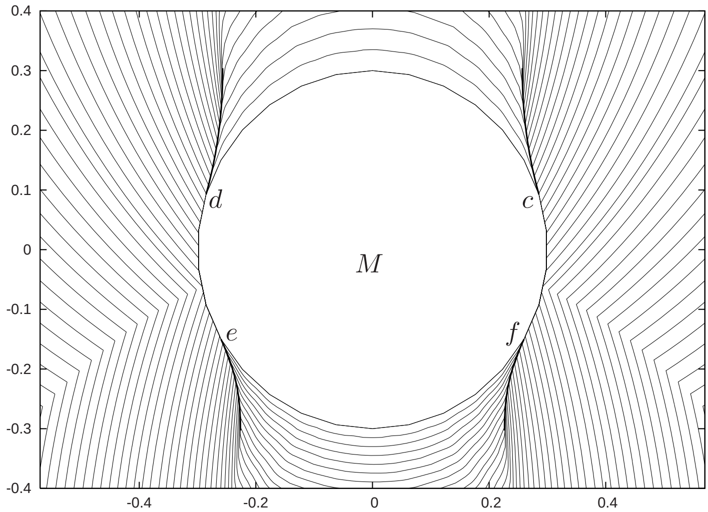

Figure 22: Enlarged fragment of Fig. 21.

system (3) in the plane of geometric coordinates at time $t = \tau$ with M being the initial set and with initial orientation of the forward motion direction h along

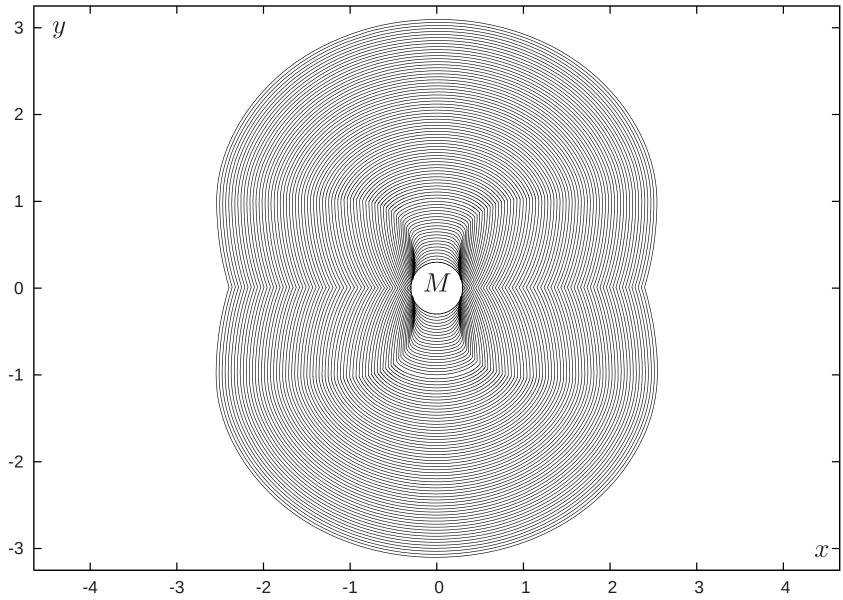

Figure 23: Level sets W_τ for differential game (5); $a = -1$, $\tau_f = 4$.

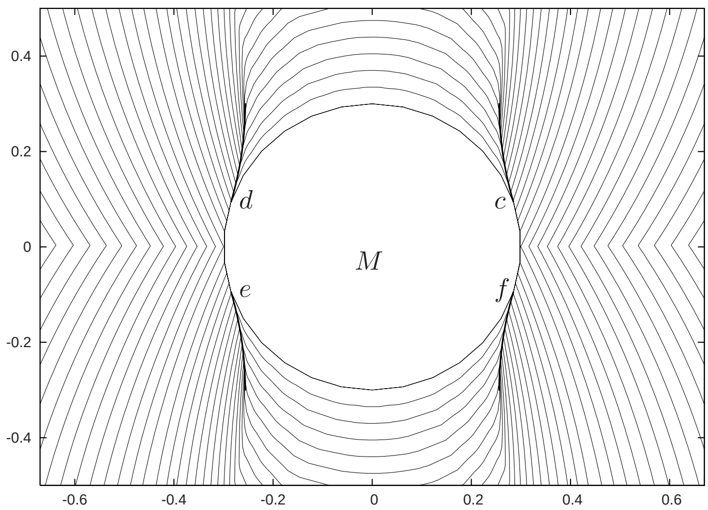

Figure 24: Enlarged fragment of Fig. 23.

the vertical axis. For Reeds and Shepp's model and the point set $M = \{0\}$, the boundary $\partial G(t, \{0\})$ is well investigated analytically in [21]. Besides, the structure of open-loop controls $u(\cdot)$, $w(\cdot)$ that bring the system to $\partial G(t, \{0\})$ is established. The sets W_τ shown in Fig. 23 are similar to the ones in Fig. 7 from [21]. But, to our opinion, the principal, and important fact for applications is that the optimal guaranteed time (the value function $V(x, y)$) is not continuous if disturbances are presented (i.e., $\nu > 0$). If $\nu = 0$, the value function is continuous.

The numerical procedure utilized for the computation of all level sets presented above uses an automatic adjustment of the step width Δ of the backward construction. The initial value of Δ, which is usually equal to 0.01, can decrease up to 10 times in the course of construction. The fronts are shown with the time step 0.1 in Figs. 11, 13, 14, 18–20, and with the time step 0.05 in Figs. 15–17, Figs. 21–24.

B. In Fig. 25, level sets W_τ computed for $\tau = 1$ are presented. Five sets are shown that correspond to the following values of parameter a: $1, 0.25, -0.1, -0.6, -1$. One can see how the set W_τ grows with the decrease of a, i.e., with the increasing length of the interval $[a, 1]$ from which the control w is being chosen. For $a = 1, 0.25, -0.1$, the set W_τ consists of an upper piece only; for $a = -0.6, -1$, a lower piece is added. The upper and lower pieces are symmetrical for $a = -1$.

Level sets $W(\tau)$ for $\tau = 3$ are given in Fig. 26.

Figures 25 and 26 demonstrate not only the dependence of the size of sets W_τ on the parameter a. The boundary of every set from Fig. 25 ($\tau = 1$) and every set from Fig. 26 ($\tau = 3$) contains the common upper smooth arc mn. This means that the optimal feedback control of player P for initial points on mn is the same as in the classical game, i.e., $w \equiv 1$. Any other common arcs on ∂W_τ do not exist. Therefore, for those initial points on ∂W_τ which do not belong to the arc mn, the optimal control of player P should utilize both the value $w = 1$ and $w = a$. The construction of optimal feedback control of player P on the base of level sets of the value function is a separate question which is not discussed in this paper. For Reeds and Shepp's problem (i.e., for $a = -1$), the optimal feedback control in the absence of disturbances (i.e., for $\nu = 0$) has been constructed in papers [22], [23].

C. Figure 27 demonstrates the discontinuity curves of the value function depending on the parameter a. As it was already mentioned, the upper usable part defined by the points c and d on the boundary of the circle M is the same for all values $a \in [-1, 1]$. The right and left barrier lines emanate from the points c and d. For $a \in [\overline{a}, 1]$, where $\overline{a} \approx 0.25$, we have long barriers (Fig. 27a). For $a \in [-1, \overline{a})$, the barriers are short (Fig. 27b). The short barriers terminate on the horizontal line $y = \nu = 0.3$. For $a \in [-0.3, 1]$, the lower part of the boundary of the circle M between the points c and d is a discontinuity curve of the value function.

If $a \in [-1, -0.3)$, there is a lower usable part ef (Fig. 27b) on the boundary of the circle M, which increases with decreasing a. The arcs cf and de on ∂M are discontinuity curves of the value function. If $a \in (\tilde{a}, -0.3)$, where $\tilde{a} = -0.3\sqrt{2}$, barrier lines emanating from the points e and f do not exist. For $a \in [-1, \tilde{a}]$, there is a barrier line being a first type semipermeable curve of the family $\Lambda^{(1),1,2-}$ that emanate from the left endpoint e of the usable part. The curve terminates on the horizontal line $y = -\nu = -0.3$. The right barrier emanated from the point f is

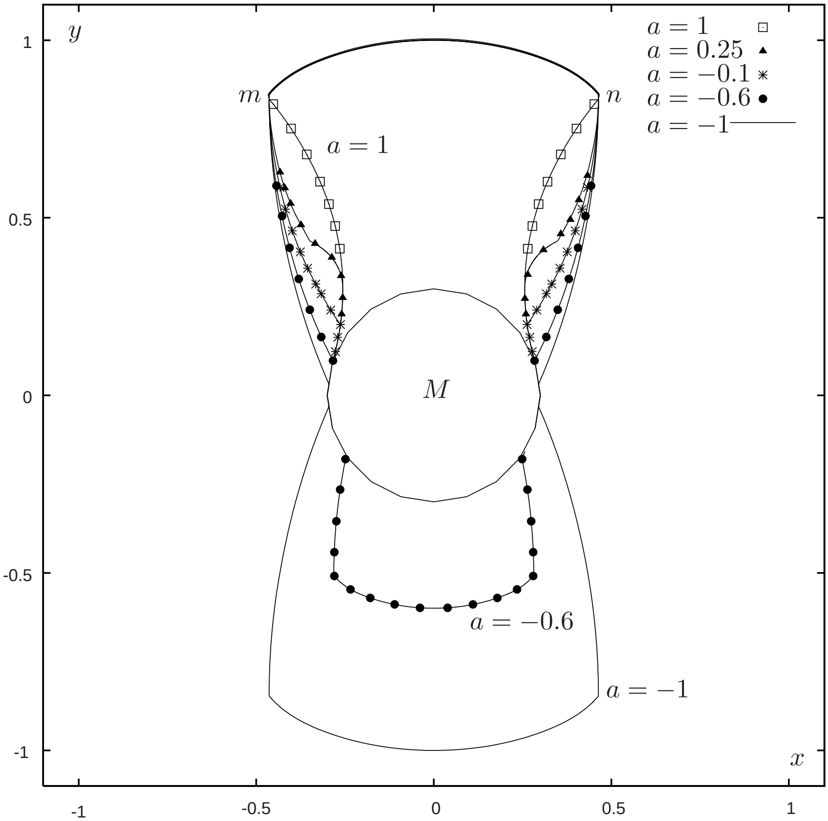

Figure 25: Level sets $W_\tau, \tau = 1$, of the value function for various magnitudes of the parameter a.

symmetrical to the left one with respect to the vertical axis and is a second type semipermeable curve.

5 Conclusion

It is known that the optimal control problems with dynamics describing an inertial car (see, e.g., [11]) are very complicated to study. This is why simplified models are widely used in mathematical literature. In particular, Reeds and Shepp's model is popular enough.

It is assumed in this model that the car can change instantaneously not only the angular velocity but also the direction of its motion to the opposite one without any loss in the velocity.

Therefore, we have two controls. The first control u is bounded as $|u| \leq 1$ and defines the angular velocity, the second control $w = \pm 1$ specifies the direc-

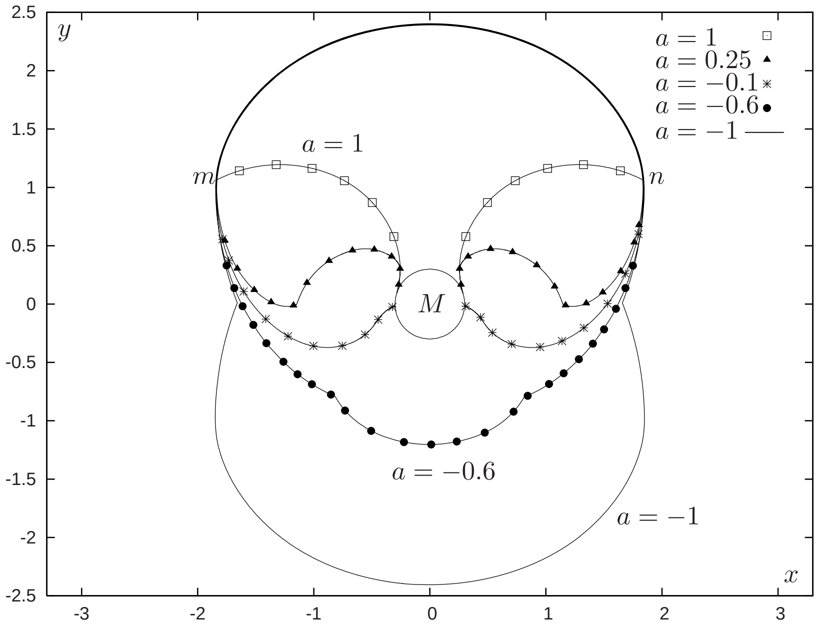

Figure 26: Level sets $W_\tau, \tau = 3$, of the value function for various magnitudes of the parameter a.

tion of motion. Using the convexification of vectograms, we can replace the constraint $w = \pm 1$ by the constraint $|w| \le 1$. Thus, we can speak about an instantaneous change of the linear velocity in the range from -1 to $+1$ for Reeds and Shepp's model.

It is supposed in the paper that control w which instantaneously changes the magnitude of the linear velocity is restricted as $a \le w \le 1$. Here, $a \in [-1, +1]$ is the parameter of the problem. For $a = 1$, we have the car with the constant magnitude of the linear velocity. For $a = -1$, we obtain Reeds and Shepp's car.

Using the above-mentioned dynamics of the car, we investigate a differential game which is similar to the famous homicidal chauffeur game by R. Isaacs. The objective of the player P that controls the car is to catch as soon as possible a non-inertial pedestrian E in a given circular neighborhood of P-state. For $a = 1$, this game becomes the classical one.

After we applied Isaacs' transformation, the construction of level sets of the value function is run in the two-dimensional plane of reduced coordinates. The level sets are computed using the algorithm developed by the authors for time-optimal differential games in the plane. This algorithm enables to construct the level sets of the value function (in other words, the fronts or isochrones) with good accuracy, which allows us to investigate in detail the regions of accumulation of fronts, the discontinuity lines of the value function, and the behavior of fronts near

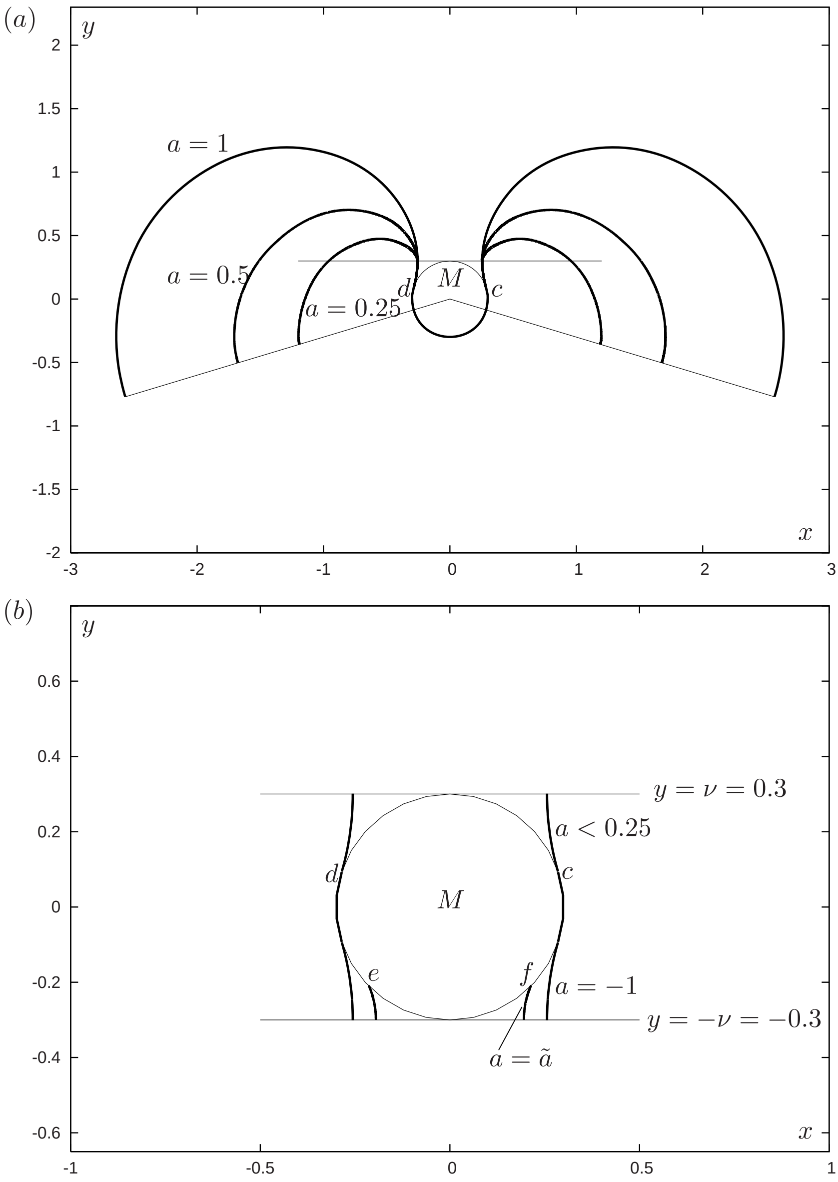

Figure 27: Discontinuity lines of the value function depending on the parameter a; $(a)\ a \geq \overline{a} \approx 0.25$; $(b)\ a < \overline{a}$.

such lines. In the paper, the structure of level sets is analyzed depending on the parameter a .

Additionally and independently on the construction of level sets of the value function, the families of smooth semipermeable curves defined by the dynamics of the problem are described in the plane of reduced coordinates. The knowledge of these families enables to validate the discontinuity lines computed in the course of run of the algorithm for the construction of level sets. Since the families of semipermeable curves depend on the dynamics only, they can be also used in other problems with mentioned dynamics, for example, in time-optimal problems with objectives of the players that are different from those considered in this paper.

Acknowledgements

This research was supported by RFBR Grants Nos. 06-01-00414 and 07-01-96085.

References

[1] Bardi M., Falcone M., Soravia P., *Numerical methods for pursuit-evasion games via viscosity solutions*, in M. Bardi, T.E.S. Raghavan, T. Parthasarathy (eds.), *Stochastic and Differential Games: Theory and Numerical Methods. Annals of the International Society of Dynamic Games* **4**, Birkhäuser, Boston, (1999), 105–175.

[2] Bernhard P., Larrouturou B., *Etude de la barriere pour un probleme de fuite optimale dans le plan*, Rapport de Recherche, INRIA, Sophia-Antipolis, (1989).

[3] Boissonnat J. D., Cerezo A., Leblond J., *Shortest paths of bounded curvature in the plane*, Proc. IEEE Conf. Robotics and Automation, Nice, France, (1992).

[4] Breakwell J. V., Merz A. W., *Toward a complete solution of the homicidal chauffeur game*, Proc. of the 1st Int. Conf. on the Theory and Applications of Differential Games, Amherst, Massachusetts, (1969).

[5] Breakwell J. V., *Zero-sum differential games with terminal payoff*, in P. Hagedorn, H. W. Knobloch, G. J. Olsder (eds.), *Differential Games and Applications. Lecture Notes in Control and Information Sciences*, Springer, Berlin, (1977), 70–95.

[6] Cardaliaguet P., Quincampoix M., Saint-Pierre P., *Numerical methods for optimal control and differential games*, Ceremade CNRS URA 749, University of Paris-Dauphine, (1995).

[7] Cockayne E. J., Hall G. W. C., *Plane motion of a particle subject to curvature constraints*, SIAM J. Control **13**(1), (1975), 197–220.

[8] Dubins L. E., *On curves of minimal length with a constraint on average curvature and with prescribed initial and terminal positions and tangents*, Amer. J. Math. **79**, (1957), 497–516.

[9] Isaacs R., Games of pursuit, Scientific report of the RAND Corporation, Santa Monica, (1951).

[10] Isaacs R., *Differential Games*, John Wiley, New York, (1965).

[11] Laumond J.-P. (ed.), *Robot Motion Planning and Control*, Lecture Notes in Control and Information Sciences **229**, Springer, New York, (1998).

[12] Lewin J., Breakwell J. V., *The surveillance-evasion game of degree*, JOTA **16**(3/4), (1975), 339–353.

[13] Lewin J., Olsder, G. J., *Conic surveillance-evasion*, JOTA **27**(1), (1979), 107–125.

[14] Markov A. A., *Some examples of the solution of a special kind of problem on greatest and least quantities*, Soobscenija Charkovskogo matematiceskogo obscestva, Ser.2,1, NN5,6, (1889), 250–276 (in Russian).

[15] Merz A. W., *The homicidal chauffeur – a differential game*, PhD Thesis, Stanford University, (1971).

[16] Patsko V. S., Turova V. L., *Level sets of the value function in differential games with the homicidal chauffeur dynamics*, Int. Game Theory Review **3**(1), (2001), 67–112.

[17] Patsko V. S., Turova V. L., *Families of semipermeable curves in differential games with the homicidal chauffeur dynamics*, Automatica **40**(12), (2004), 2059–2068.

[18] Pecsvaradi T., *Optimal horizontal guidance law for aircraft in the terminal area*, IEEE Trans. on Automatic Control **AC-17**(6), (1972), 763–772.

[19] Raivio T., Ehtamo H., *On the numerical solution of a class of pursuit-evasion games*, in J. Filar, K. Mizukami, V. Gaitsgory (eds.), *Advances in Dynamic Games and Applications. Annals of the International Society of Dynamic Games* **5**, Birkhäuser, Boston, (2000), 177–192.

[20] Reeds J. A., Shepp L. A., *Optimal paths for a car that goes both forwards and backwards*, Pacific J. Math. **145**(2), (1990), 367–393.

[21] Souères P., Fourquet J.-Y., Laumond J.-P., *Set of reachable positions for a car*, IEEE Trans. on Automatic Control **39**(8), (1994), 1626–1630.

[22] Souères P., Laumond J-P., *Shortest paths synthesis for a car-like robot*, IEEE Trans. on Automatic Control **41**(5), (1996), 672–688.

[23] Souères P., *Minimum-length trajectories for a car: an example of the use of Boltianskii's sufficient conditions for optimality*, IEEE Trans. on Automatic Control **52**(2), (2007), 323–327.

[24] Sussmann H. J., Tang W., *Shortest paths for the Reeds-Shepp car: a worked*

out example of the use of geometric techniques in nonlinear optimal control, Rutgers University, New Brunswick, NJ, Rep. SYCON-91-10, (1991).

[25] Vendittelli M., Laumond J.-P., Nissoux C., *Obstacle distance for car-like robots*, IEEE Trans. on Robotics and Automation **15**(4), (1999), 678–691.

Credible Linear-Incentive Equilibrium Strategies in Linear-Quadratic Differential Games

Guiomar Martín-Herrán
Departamento de Economía Aplicada (Matemáticas)
Universidad de Valladolid
Valladolid, Spain
guiomar@eco.uva.es

Georges Zaccour
Chair in Game Theory and Management and GERAD
HEC Montréal
Montréal (Québec) Canada H3T 2A7
georges.zaccour@gerad.ca

Abstract

In this paper we characterize the credibility of linear-incentive equilibrium strategies for the class of linear-quadratic differential games. This class is widely used in applications of differential games in economics and management science. We derive a general condition for credibility and illustrate its use on two examples: one is a general homogenous linear-quadratic game and the second is an environmental-economics game.

Key words. Linear-quadratic differential games, cooperation, incentive equilibria, credibility, environmental economics

AMS Subject Classifications. 91A23; 49N10.

1 Introduction

An important issue in cooperative differential games is the sustainability over time of the efficient solution agreed upon by the players at the initial date of the game. Intuitively, an agreement will last for its whole intended duration if, at any intermediate instant of time $\tau \in [0, T]$, where T is the end date of the agreement, each player prefers (payoffwise) to continue being part of the agreement rather than leaving it. Haurie (1976) was the first to point out that sustainability is not necessarily always at hand. He provided two possible reasons for an agreement to be broken at that intermediate instant of time $\tau \in [0, T]$:

(1) The original agreement may not be a solution of the cooperative game that starts out at τ, and, therefore, the players would not go for its continuation.
(2) A player may find it optimal to deviate (or cheat) by implementing a different strategy than the one prescribed by the agreement.

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_14,

Different avenues have been suggested in the literature to go around, at least to some extent, of such instabilities. One option is to assume that the agreement is binding (i.e., renegotiation-free), as done in some early work in cooperative differential games (see, e.g., Liu (1973)). Such an assumption does not, however, stand an empirical test. Indeed, simple observation shows that agreements, albeit duly signed, are broken every day (divorce is an illustrative example!). Another approach is to design a cooperative solution or an agreement which is time-consistent, i.e., along the cooperative state trajectory, no player finds it optimal to switch to her non-cooperative control at any intermediate instant of time (see, Yeung and Petrosyan (2005), Jørgensen and Zaccour (2001a, 2002), Petrosjan and Zaccour (2003), Petrosjan and Zenkevich (1996), and Petrosjan (1997)). Note that the test consists in comparing the cooperative and non-cooperative payoffs-to-go. An alternative possibility is to design an *agreeable* agreement. Agreeability requires the comparison condition of the payoffs-to-go to hold along any state trajectory and not only along the cooperative one, and thus it is a stronger sustainability concept than time-consistency (see, e.g., Kaitala and Pohjola (1990, 1995) and Jørgensen et al. (2003, 2005)). Finally, one can attempt to design an agreement which is self-enforceable, i.e., is an equilibrium[1]. One approach to embody the cooperative solution with the equilibrium property is to use so-called trigger strategies. These strategies are based on the past actions in the game and they include a threat to punish, credibly and effectively, any player who cheats on the agreement (see, e.g., Tolwinski et al. (1986) and Dockner et al. (2000)).

In a two-player differential-game setting, another approach, which is our topic, is to support the cooperative solution by incentive strategies (see, e.g., Ehtamo and Hämäläinen (1986, 1989, 1993), Jørgensen and Zaccour (2001b), and Martín-Herrán and Zaccour (2005)). Incentive strategies are functions of the possible deviation of the other player and recommend to each player to implement her part of the cooperative (desired or coordinated) solution whenever the other player is doing so. Although these strategies are relatively easy to construct, one concern is their credibility. By this we mean that each player will, indeed, implement her incentive strategy, and not the coordinated solution, if she observes that the other one has deviated from the coordinated solution. In Martín-Herrán and Zaccour (2005), we provided a condition to check for the credibility of incentive strategies for the class of linear-state differential games. We illustrated its use on two examples taken from the literature, showed that the proposed linear-incentive strategies are not always credible, and provided non-linear ones which are always credible. In this follow-up paper, we consider the credibility of incentive strategies for the class of linear-quadratic differential games (LQDG) and derive a credibility condition which can be (relatively easily) be checked.

[1] It may happen that the cooperative solution is in itself an equilibrium, as in some differential games of special structure (see, e.g., Chiarella et al. (1984), Rincón-Zapatero et al. (2000), and Martín-Herrán and Rincón-Zapatero (2005)). In such setting, the problem is solved.

The differential games literature in economics, management science, and other applied areas, has often adopted the linear-quandratic (LQ) structure. The main reason for its popularity is that this structure allows to capture some needed economic properties (e.g., marginal decreasing returns, saturation effects) in still a tractable manner. Many examples of applications of such class of differential games can be found in the recent books by Dockner et al. (2000), Jørgensen and Zaccour (2004), Engwerda (2005), Yeung and Petrosyan (2005), and the survey by Jørgensen and Zaccour (2007).

The rest of the paper is organized as follows. In Sec. 2, we recall the ingredients of linear-quadratic differential games and derive the coordinated solution. In Sec. 3, we define incentive strategies and equilibrium and provide a characterization formula for assessing their credibility. In Sec. 4, we illustrate our result on an environmental-economics LQ game. In Sec. 5, we conclude.

2 Linear-Quadratic Differential Games

We focus on two-player LQDG with one state variable. The game is played over an infinite time horizon. To save on notation, we assume that each player has a scalar control, although all results can be generalized to a multidimensional control case.

The optimization problem of player i is as follows:

$$\max_{u_i(t)} \left\{ W_i = \int_0^\infty e^{-\rho t} \left(\frac{1}{2}[m_i x^2(t) + u'(t) R^i u(t)] + q_i x(t) + \sum_{j=1}^{2} \left(r_{ij} + d_{ij} x(t)\right) u_j(t) \right) dt \right\},$$

$$\text{subject to: } \dot{x}(t) = \alpha x(t) + \sum_{i=1}^{2} \beta_i u_i(t), \quad x(0) = x_0, \tag{1}$$

where $x(t) \in \mathbb{R}$ is the state variable, x_0 is the initial state, $u_i(t)$ is the control variable of player $i \in \{1,2\}$, and $u'(t) = (u_1(t), u_2(t)) \in \mathbb{R}^2$. Here, $R^i \in \mathbb{R}^{2\times 2}$ are symmetric matrices for all $i \in \{1,2\}$. The element in row j and column k in matrix R^i is denoted by R^i_{jk}. Finally, $r_{ij}, d_{ij}, m_i, q_i, \alpha, \beta_i$ are constants and ρ is the discount rate ($\rho > 0$).

In what follows, we assume that the players implement stationary Markovian strategies, which is a standard assumption in autonomous, infinite-horizon differential games. Due to stationarity, equilibrium strategies and value functions do not depend explicitly on t.

2.1 Cooperative solution

Assume that the players agree to play a cooperative game in which they jointly maximize the aggregate payoff:

$$\sum_{i=1}^{2} W_i = \sum_{i=1}^{2} \int_0^{\infty} e^{-\rho t} \left(\frac{1}{2}[m_i x^2(t) + u'(t) R^i u(t)] + q_i x(t) + \sum_{j=1}^{2} \left(r_{ij} + d_{ij} x(t)\right) u_j(t) \right) dt \, .$$

The next proposition characterizes the solution of the cooperative problem that we wish to sustain by incentive equilibrium strategies.

Proposition 2.1. *Denote by* (u_1^c, u_2^c) *the cooperative solution*[2]*. The optimal controls* u_i^c, $i = 1, 2$, *which are linear in the state variable* x *are given by:*

$$u_i^c(x) = A_i^c x + B_i^c, \qquad i = 1, 2 \, . \tag{2}$$

The optimal state trajectory is:

$$x^c(t) = \left(x_0 + \frac{X_1^c}{X_2^c} \right) e^{X_2^c t} - \frac{X_1^c}{X_2^c} \, . \tag{3}$$

Constants A_i^c, B_i^c *and* X_i^c *are given by:*

$$\begin{aligned} A_i^c &= \frac{(d_{ij} + d_{jj} + a^c \beta_j)(R_{ij}^i + R_{ij}^j) - (d_{ii} + d_{ji} + a^c \beta_i)(R_{jj}^i + R_{jj}^j)}{(R_{ii}^i + R_{ii}^j)(R_{jj}^i + R_{jj}^j) - (R_{ij}^i + R_{ij}^j)^2}, \\ B_i^c &= \frac{(r_{ij} + r_{jj} + b^c \beta_j)(R_{ij}^i + R_{ij}^j) - (r_{ii} + r_{ji} + b^c \beta_i)(R_{jj}^i + R_{jj}^j)}{(R_{ii}^i + R_{ii}^j)(R_{jj}^i + R_{jj}^j) - (R_{ij}^i + R_{ij}^j)^2}, \\ X_1^c &= B_1^c \beta_1 + B_2^c \beta_2, \quad X_2^c = \alpha + A_1^c \beta_1 + A_2^c \beta_2 \, . \end{aligned}$$

Coefficients a^c, b^c*, and* c^c *are characterized in Appendix* A*.*

Proof. In the cooperative game the players jointly optimize the aggregate payoff. Let us denote by

$$V^c(x) = \frac{a^c}{2} x^2 + b^c x + c^c \, ,$$

the optimal value function of this game. Using the Hamilton-Jacobi-Bellman (HJB) equation, we obtain from the first-order optimality condition that the optimal controls satisfy the following system:

$$(R_{ii}^i + R_{ii}^j) u_i + (R_{ij}^i + R_{ij}^j) u_j + r_{ii} + r_{ji} + (d_{ii} + d_{ji}) x + \beta_i (V^c)'(x) = 0, \quad i = 1, 2 \, .$$

[2] From now on the time argument is omitted so no confusion can arise.

The solution to this system is u_i^c, $i = 1, 2$, given in (2). The expressions of these optimal controls depend on the coefficients of the value function a^c, b^c, which can be determined by identification by means of the HJB equation. Their expressions are collected in Appendix A.

Inserting the optimal controls in the state equation and solving leads to the optimal state trajectory in (3). □

Once the optimal cooperative solution is obtained, we can compute the part corresponding to each player of the total cooperative payoff.

Corollary 2.1. *Player i's optimal payoff under the cooperative strategy is given by:*

$$\begin{aligned} W_i(u_1^c, u_2^c) = {} & K_{1i}^c \left(x_0 + \frac{X_1^c}{X_2^c} \right)^2 \frac{1}{\rho - 2X_2^c} \\ & + \left(x_0 + \frac{X_1^c}{X_2^c} \right) \left(K_{2i}^c - 2K_{1i}^c \frac{X_1^c}{X_2^c} \right) \frac{1}{\rho - X_2^c} \\ & + \frac{X_1^c}{X_2^c} \left(K_{2i}^c - K_{1i}^c \frac{X_1^c}{X_2^c} \right) \frac{1}{\rho} \, , \end{aligned} \tag{4}$$

where $\rho - X_2^c$ and $\rho - 2X_2^c$ are assumed to be positive, and

$$\begin{aligned} K_{1i}^c &= A_i^c d_{ii} + A_j^c d_{ij} + \frac{m_i}{2} + \frac{(A_i^c)^2 R_{ii}^i}{2} + \frac{\left(A_j^c\right)^2 R_{jj}^i}{2} + A_i^c A_j^c R_{ij}^i \, , \\ K_{2i}^c &= q_i + A_i^c r_{ii} + A_j^c r_{ij} + B_i^c (d_{ii} + A_i^c R_{ii}^i + A_j^c R_{ij}^i) \\ & \quad + B_j^c (d_{ij} + A_i^c R_{ij}^i + A_j^c R_{jj}^i) \, , \end{aligned}$$

$i, j \in \{1, 2\}, i \neq j.$

Proof. Substitute in each player objective functional the control by $u_i^c, i = 1, 2$ given in (2), and the state variable by its optimal trajectory given in (3):

$$\begin{aligned} W_i(u_1^c, u_2^c) = \int_0^\infty & \left[\frac{1}{2} [m_i (x^c)^2(t) + (u^c)'(t) R^i u^c(t)] + q_i x^c(t) \right. \\ & \left. + \sum_{j=1}^2 (r_{ij} + d_{ij} x^c(t)) \, u_j^c(t) \right] e^{-\rho t} \, dt \, , \end{aligned}$$

where u^c denotes the vector (u_1^c, u_2^c). After some computations and arranging terms, the expression above can be rewritten as:

$$W_i(u_1^c, u_2^c) = \int_0^\infty \left[K_{1i}^c (x^c)^2(t) + K_{2i}^c x^c(t) + K_{3i} \right] e^{-\rho t} \, dt \, ,$$

where

$$K_{3i} = B_i^c r_{ii} + B_j^c r_{ij} + \frac{1}{2}\left[(B_i^c)^2 R_{ii}^i + \left(B_j^c\right)^2 R_{jj}^i + 2B_i^c B_j^c R_{ij}^i\right], \quad i,j \in \{1,2\},\ i \neq j\ .$$

Integrating over time leads to the expression in (4), where assumptions $\rho - X_2^c > 0$ and $\rho - 2X_2^c > 0$ are needed to guarantee a convergent improper integral. □

The following proposition provides the players' payoffs, and resulting state trajectory, for any pair of linear strategies. These results are needed in the next section.

Proposition 2.2. *If $(\hat{u}_1, \hat{u}_2)$ is any pair of linear controls, given by*

$$\hat{u}_i(x) = \hat{A}_i x + \hat{B}_i\ , \tag{5}$$

then the corresponding state trajectory is

$$\hat{x}(t) = \left(x_0 + \frac{\hat{X}_1}{\hat{X}_2}\right) e^{\hat{X}_2 t} - \frac{\hat{X}_1}{\hat{X}_2}\ ,$$

and player i's payoff is given by:

$$\begin{aligned} W_i(\hat{u}_1, \hat{u}_2) = {} & \hat{K}_{1i}\left(x_0 + \frac{\hat{X}_1}{\hat{X}_2}\right)^2 \frac{1}{\rho - 2\hat{X}_2} \\ & + \left(x_0 + \frac{\hat{X}_1}{\hat{X}_2}\right)\left(\hat{K}_{2i} - 2\hat{K}_{1i}\frac{\hat{X}_1}{\hat{X}_2}\right)\frac{1}{\rho - \hat{X}_2} \\ & + \frac{\hat{X}_1}{\hat{X}_2}\left(\hat{K}_{2i} - \hat{K}_{1i}\frac{\hat{X}_1}{\hat{X}_2}\right)\frac{1}{\rho}\ , \end{aligned}$$

where $\rho - \hat{X}_2$ and $\rho - 2\hat{X}_2$ are assumed to be positive and constants $\hat{X}_k, \hat{K}_{1l}, k = 1,2;\ l = 1,\dots 3$ are obtained changing A_i^c, B_i^c by $\hat{A}_i$, $\hat{B}_i$ in X_k^c, $K_{1l}^c, k = 1,2;\ l = 1,\dots 3$.

Proof. Replacing $(\hat{u}_1, \hat{u}_2)$ in the state equation and solving leads to the state trajectory.

Player i's payoff when the pair of linear controls $(\hat{u}_1, \hat{u}_2)$, given by (5), is used can be obtained along the same lines as in Corollary 2.1. □

3 Incentive Equilibria

The cooperative solution characterized above is not, in general, an equilibrium. As mentioned in the Introduction, our aim is to sustain the coordinated solution in time by means of incentive equilibrium strategies and characterize their credibility. For the sake of completeness, let us recall the definition of an incentive equilibrium strategy.

Let $(u_1^c, u_2^c) \in \mathbb{R} \times \mathbb{R}$ be the desired cooperative solution. Denote by

$$\Psi_1 = \{\psi_1 \,|\, \psi_1 : \mathbb{R} \longrightarrow \mathbb{R}\},\ \Psi_2 = \{\psi_2 \,|\, \psi_2 : \mathbb{R} \longrightarrow \mathbb{R}\}\ ,$$

the sets of admissible incentive strategies.

Definition 3.1. A strategy pair $\psi_1 \in \Psi_1, \psi_2 \in \Psi_2$ is an incentive equilibrium at (u_1^c, u_2^c) if

$$\begin{aligned} W_1(u_1^c, u_2^c) &\geq W_1(u_1^c, \psi_2(u_1)), && \forall u_1 \in \mathbb{R}\ , \\ W_2(u_1^c, u_2^c) &\geq W_2(\psi_1(u_2), u_2^c), && \forall u_2 \in \mathbb{R}\ , \\ \psi_1(u_2^c) &= u_1^c,\ \psi_2(u_1^c) = u_2^c\ . \end{aligned}$$

In order to keep things simple, we shall confine in what follows ourselves to linear-incentive strategies, that is:

$$\psi_j(u_i) = u_j^c + D_j(u_i - u_i^c), \quad i,j = 1,2,\ i \neq j\ , \tag{6}$$

with $D_j,\ j = 1,2$, denoting a non-zero constant.

To characterize an incentive equilibrium we need to solve the following pair of optimal control problems where each player assumes that the other is using the incentive strategy given in (6):

$$\max_{u_i} W_i = \int_0^\infty \left(\frac{1}{2}[m_i x^2 + u' R^i u] + q_i x + \sum_{j=1}^{2} (r_{ij} + d_{ij} x)\, u_j \right) e^{-\rho t}$$

$$\text{s.t.: } \dot{x} = \alpha x + \sum_{i=1}^{2} \beta_i u_i, \quad x(0) = x_0, \quad \rho > 0\ , \tag{7}$$

$$u_j = \psi_j(u_i) = u_j^c + D_j(u_i - u_i^c), \quad i,j = 1,2,\ i \neq j\ .$$

The next proposition characterizes the solutions of these optimal control problems.

Proposition 3.1. *An interior solution u_i^* $(i = 1,2)$ of the optimal control problem in (7) is given by:*

$$u_i^*(x) = A_i^* x + B_i^*, \quad i = 1,2\ , \tag{8}$$

where:

$$A_i^* = -\frac{d_{ii} + d_{ij} D_j + (A_j^c - A_i^c D_j)(R_{ij}^i + R_{jj}^i D_j + a_i^*(\beta_i + \beta_j D_j))}{R_{ii}^i + D_j(2R_{ij}^i + R_{jj}^i D_j)}\ ,$$

$$B_i^* = -\frac{r_{ii} + r_{ij} D_j + (B_j^c - B_i^c D_j)(R_{ij}^i + R_{jj}^i D_j + b_i^*(\beta_i + \beta_j D_j))}{R_{ii}^i + D_j(2R_{ij}^i + R_{jj}^i D_j)}\ .$$

Coefficients a_i^*, b_i^*, and c_i^* are characterized in Appendix A.

Proof. Let us denote by

$$V_i^*(x) = \frac{a_i^*}{2}x^2 + b_i^* x + c_i^*$$

the optimal value function of the optimal control described below. Using the Hamilton-Jacobi-Bellman (HJB) equation from the first-order optimality condition and after some easy computations, we obtain the optimal control u_i^* in (8). Note that the expression of u_i^* is written in terms of the coefficients of the value function a_i^*, b_i^*, which can be determined by identification using the HJB equation. Their expressions are printed in Appendix A. □

To determine an incentive equilibrium we need to impose

$$u_i^* = u_i^c,\ i = 1, 2\ .$$

Using this fact, the following proposition establishes necessary conditions which must be satisfied by the incentive equilibrium strategies.

Proposition 3.2. *To be an incentive equilibrium at* (u_1^c, u_2^c)*, a strategy pair* $\psi_1 \in \Psi_1, \psi_2 \in \Psi_2$ *given by* (6) *must satisfy the following condition:*

$$D_j = \frac{T_{23}^i T_{11}^i - T_{13}^i T_{21}^i}{T_{12}^i T_{21}^i - T_{11}^i T_{22}^i}, \quad j, i = 1, 2,\ i \neq j\ , \tag{9}$$

where coefficients $T_{kl}^i,\ k = 1, 2, l = 1, \ldots, 3$ *are given in Appendix* A *and* T_{21}^i *and* $T_{12}^i T_{21}^i - T_{11}^i T_{22}^i$ *for* $i = 1, 2$ *are assumed to be non-null.*

Proof. Since the cooperative control of each player, u_i^c, must equal u_i^*, from expressions (2) and (8) we have:

$$A_i^c = A_i^*, \quad B_i^c = B_i^*, \quad i = 1, 2\ .$$

For each player these two equations can be rewritten as a second-order polynomials in D_j, as follows:

$$T_{11}^i D_j^2 + T_{12}^i D_j + T_{13} = 0\ , \tag{10}$$

$$T_{21}^i D_j^2 + T_{22}^i D_j + T_{23} = 0\ . \tag{11}$$

The coefficients $T_{kl}^i,\ k = 1, 2, l = 1, \ldots, 3$ are collected in Appendix A. From Eq. (11), under assumption $T_{21}^i \neq 0$, we derive:

$$D_j^2 = -\frac{T_{22} D_j + T_{23}}{T_{21}}\ .$$

Replacing this expression in (10) we obtain a linear equation in D_j. Under assumption $T_{12}^i T_{21}^i - T_{11}^i T_{22}^i \neq 0$, the solution of this equation is given by (9). □

3.1 Credibility of Incentive Strategies

For an incentive equilibrium to be credible, it must be in the best interest of each player to implement her incentive strategy if the other player deviates from the coordinated solution rather than to play her part of the cooperative solution. A formal definition follows.

Definition 3.2. The incentive equilibrium strategy pair $\psi_1 \in \Psi_1, \psi_2 \in \Psi_2$ at (u_1^c, u_2^c) is credible in $U_1 \times U_2$ if the following inequalities are satisfied:

$$W_1(\psi_1(u_2), u_2) \geq W_1(u_1^c, u_2) , \qquad \forall u_2 \in U_2 , \tag{12}$$
$$W_2(u_1, \psi_2(u_1)) \geq W_2(u_1, u_2^c) , \qquad \forall u_1 \in U_1 . \tag{13}$$

Note that the above definition characterizes the credibility of equilibrium strategies for any possible deviation in the set $U_1 \times U_2$. However, in order to preserve the linear-quadratic structure of the game, we restrict our analysis to the linear-incentive strategies defined by (6). Moreover, we assume that if one player deviates from the cooperative solution, then she will use a control $\tilde{u}_i$ which is supposed to be linear and given by:

$$\tilde{u}_i(x) = \tilde{A}_i x + \tilde{B}_i , \tag{14}$$

where $\tilde{A}_i, \tilde{B}_i$ are two constants. Let us denote by

$$F_i = \{u_i \text{ admissible} \,|\, u_i(x) = \tilde{A}_i x + \tilde{B}_i\}, \quad i = 1, 2 . \tag{15}$$

The next proposition states the credibility conditions for a linear-quadratic differential game under the previous assumptions.

Proposition 3.3. *Consider the differential game defined by* (1) *and denote by* (u_1^c, u_2^c) *its cooperative solution. The incentive equilibrium strategy pair* $\psi_1 \in \Psi_1, \psi_2 \in \Psi_2$ *at* (u_1^c, u_2^c) *given by* (6) *is credible for any deviation* $u_1 \in F_1$ *and* $u_2 \in F_2$ *if the following conditions hold:*

$$\begin{aligned}
&\Big[\tilde{K}_{1i}\big(\bar{X}_{2i}\big)^2(\rho - 2\tilde{X}_{2i})\Big(\tilde{X}_{1i} + x_0\tilde{X}_{2i}\Big)^2 \\
&\quad -\bar{K}_{1i}\Big(\tilde{X}_{2i}\Big)^2(\rho - 2\bar{X}_{2i})\big(\bar{X}_{1i} + x_0\bar{X}_{2i}\big)^2\Big]\,(\rho - \tilde{X}_{2i})(\rho - \bar{X}_{2i}) \\
&\quad + \Big[(\rho - \bar{X}_{2i})\big(\bar{X}_{2i}\big)^2(\tilde{K}_{2i}\tilde{X}_{2i} - \tilde{K}_{1i}\tilde{X}_{1i})((\rho - \tilde{X}_{2i})\tilde{X}_{1i} + x_0\tilde{X}_{2i}) \\
&\quad - (\rho - \tilde{X}_{2i})\Big(\tilde{X}_{2i}\Big)^2(\bar{K}_{2i}\bar{X}_{2i} - \bar{K}_{1i}\bar{X}_{1i})((\rho - \bar{X}_{2i})\bar{X}_{1i} + x_0\bar{X}_{2i})\Big] \\
&\qquad\qquad (\rho - 2\tilde{X}_{2i})(\rho - 2\hat{X}_{2i}) > 0 ,
\end{aligned}$$

$i=1,2$, *where*

$$\tilde{X}_{1i} = \beta_i(B_i^c - D_i(B_j^c - \tilde{B}_j)) + \beta_j\tilde{B}_j\ ,$$
$$\tilde{X}_{2i} = \alpha + \beta_i(A_i^c - D_i(A_j^c - \tilde{A}_j)) + \beta_j\tilde{A}_j\ ,$$
$$\hat{X}_{1i} = \beta_i B_i^c + \beta_j\tilde{B}_j\ ,$$
$$\hat{X}_{2i} = \alpha + \beta_i A_i^c + \beta_j\tilde{A}_j\ ,$$

and constants $\tilde{K}_{ki}, \bar{K}_{ki}$, *for* $k=1,2$ *can be obtained replacing in* $\hat{K}_{ki}$ *constants* $\hat{A}_i$, $\hat{A}_j$, $\hat{B}_i$, $\hat{B}_j$ *by* $A_i^c - D_i(A_j^c - \tilde{A}_j), \tilde{A}_j, B_i^c - D_i(B_j^c - \tilde{B}_j), \tilde{B}_j$ *in case of* $\tilde{K}_{ki}$, *and by* $A_i^c, \tilde{A}_j, B_i^c, \tilde{B}_j$ *in case of* $\bar{K}_{ki}$, *with* $i,j=1,2, i\neq j$.

Proof. We have to check that inequalities (12) and (13) are satisfied for $u_2 \in F_2$ and $u_1 \in F_1$, respectively. Note that the linear-incentive equilibrium strategy pair $\psi_1 \in \Psi_1, \psi_2 \in \Psi_2$ at (u_1^c, u_2^c) given by (6) for $u_1 \in F_1, u_2 \in F_2$ reads:

$$\begin{aligned}\psi_i(u_j) &= u_i^c + D_i(u_j - u_j^c) = A_i^c x + B_i^c + D_i(u_j - (A_j^c x + B_j^c))\\ &= A_i^c x + B_i^c + D_i(\tilde{A}_j x + \tilde{B}_j - (A_j^c x + B_j^c))\\ &= [A_i^c - D_i(A_j^c - \tilde{A}_j)]x + B_i^c - D_i(B_j^c - \tilde{B}_j), \quad i,j=1,2, i\neq j\ .\end{aligned}$$

It suffices to compute the expressions of the different payoffs appearing in these inequalities taking into account the expression of player i's payoff along a given pair of linear controls established in Proposition 2.2. The payoffs can be expressed as:

$$\begin{aligned}W_i(\psi_i(u_j), u_j) = \tilde{K}_{1i}&\left(x_0 + \frac{\tilde{X}_{1i}}{\tilde{X}_{2i}}\right)^2 \frac{1}{\rho - 2\tilde{X}_{2i}}\\ &+ \left(x_0 + \frac{\tilde{X}_{1i}}{\tilde{X}_{2i}}\right)\left(\tilde{K}_{2i} - 2\tilde{K}_{1i}\frac{\tilde{X}_{1i}}{\tilde{X}_{2i}}\right)\frac{1}{\rho - \tilde{X}_{2i}}\\ &+ \frac{\tilde{X}_{1i}}{\tilde{X}_{2i}}\left(\tilde{K}_{2i} - \tilde{K}_{1i}\frac{\tilde{X}_{1i}}{\tilde{X}_{2i}}\right)\frac{1}{\rho},\ i,j=1,2,\ i\neq j\ .\end{aligned}$$

$$\begin{aligned}W_i(u_i^c, u_j) = \bar{K}_{1i}&\left(x_0 + \frac{\bar{X}_{1i}}{\bar{X}_{2i}}\right)^2 \frac{1}{\rho - 2\bar{X}_{2i}}\\ &+ \left(x_0 + \frac{\bar{X}_{1i}}{\bar{X}_{2i}}\right)\left(\bar{K}_{2i} - 2\bar{K}_{1i}\frac{\bar{X}_{1i}}{\bar{X}_{2i}}\right)\frac{1}{\rho - \bar{X}_{2i}}\\ &+ \frac{\bar{X}_{1i}}{\bar{X}_{2i}}\left(\bar{K}_{2i} - \bar{K}_{1i}\frac{\bar{X}_{1i}}{\bar{X}_{2i}}\right)\frac{1}{\rho},\ i,j=1,2,\ i\neq j\ .\end{aligned}$$

Expressions $\rho - \tilde{X}_2$, $\rho - 2\tilde{X}_2$, $\rho - \bar{X}_2$ and $\rho - 2\bar{X}_2$ are assumed to be positive in order to have convergent improper integrals. Easy, but tedious, computations lead to the inequalities in the proposition. □

The two inequalities provide conditions which could be checked once the specific functions involved are at hand.

To illustrate the analysis of the credibility of incentive equilibrium strategies, we consider two examples. One is a simple homogeneous LQDG where the objective functionals do not involve a state-control interaction term, nor the other player's control. This example does not refer to a particular economics application and is fully treated in Appendix B. It is meant to show that, even in a relatively simple setting, the credibility conditions are tedious. The following example corresponds to a global environmental problem.

4 An Environmental-Economics Example

Let player i's optimization problem be given by:

$$\max_{u_i} \left\{ W_i = \int_0^\infty e^{-\rho t} \left[u_i \left(\gamma_i - \frac{1}{2} u_i \right) - \frac{1}{2} \varphi_i x^2 \right] dt \right\} ,$$
$$\text{s.t.: } \dot{x} = \beta(u_i + u_j) - \alpha x, \quad x(0) = x_0 ,$$

where β, γ_i and $\varphi_i, i, j \in \{1, 2\}, i \neq j$ are positive parameters and $0 < \alpha < 1$. This problem has the following features:

- The objective functional of player i is quadratic in control and state and depends on the player's own control only. Note that, contrary to the homogenous example treated in Appendix 2, here a linear term in the control is included.
- There is no interaction neither between the controls nor between the controls and the state.

Admittedly, this structure is restrictive but fits actually many models belonging to an active area of applications of differential games, namely global environmental problems. In such a context the control u_i is the emissions level and the state x is the accumulated stock of pollution. Assuming that emissions are a proportional by-product of industrial activities, revenues generated from these activities can be expressed equivalently in terms of emissions seen here as an input in production. Thus, the function $f_i(u_i) = u_i \left(\gamma_i - \frac{1}{2} u_i \right)$ represents the concave revenue function of player i. Revenues depend only on own industrial activities, and not on the other players' ones. Pollution induces damage costs, given by $\frac{1}{2} \varphi_i x^2$, assumed to depend on accumulated pollution. The state dynamics describes pollution accumulation. It says that the variation in its level is the sum of emissions, scaled by a parameter, minus what is absorbed by nature (α is termed the absorption rate).

The environmental models in Van der Ploeg and De Zeeuw (1992) and Dockner and Long (1993) are particular cases of our specification. Indeed, both models are completely symmetric ($\gamma_i = \gamma, \varphi_i = \varphi$ for all players). The first paper studies an n-country model, while the second one concentrates on a two-region differential game. Both papers compare cooperative and noncooperative emissions strategies, but are not interested in the issue of sustainability. List and Mason (2001) reconsider the same problem with two asymmetric players. To derive the Markov-perfect noncooperative strategies, they assume $\gamma_j = \theta\gamma_i$ and $\varphi_j = \omega\varphi_i$ (which are more restrictive than our structure). Fernández (2002) studies transboundary water pollution and takes into account asymmetry of the two countries' costs and benefits. In Fernández (2002), each player has three control variables and the characterization of optimal equilibrium is carried out by means of numerical simulations. Another paper, which uses a similar specification to analyze a pollution control problem, is Kaitala and Pohjola (1995). Their model is asymmetric since one group of countries is assumed to be non-vulnerable to global warming. This assumption leads to an analytically solvable model, since one only needs to solve a decoupled system of algebraic Riccati equations.

In the cooperative (joint maximization) solution the linear optimal strategies and quadratic value function are given by:

$$u_i^c(x) = \gamma_i + \beta(a^c x + b^c), \quad V^c(x) = \frac{a^c}{2}x^2 + b^c x + c^c\,, \tag{16}$$

where

$$\begin{aligned} a^c &= \frac{2\alpha + \rho \pm \sqrt{(2\alpha+\rho)^2 + 8\beta^2(\varphi_1+\varphi_2)}}{4\beta^2}\,, \\ b^c &= \frac{\beta(\gamma_1+\gamma_2)a^c}{\alpha+\rho-2\beta^2 a^c}\,, \\ c^c &= \frac{\gamma_1^2+\gamma_2^2+2\beta b^c(\gamma_1+\gamma_2+\beta b^c)}{2\rho}\,. \end{aligned}$$

The optimal state trajectory is:

$$x^c(t) = \left(x_0 - \frac{\beta(\gamma_1+\gamma_2+2\beta b^c)}{\alpha-2\beta^2 a^c}\right) e^{-(\alpha-2\beta^2 a^c)t} + \frac{\beta(\gamma_1+\gamma_2+2\beta b^c)}{\alpha-2\beta^2 a^c}\,.$$

The state dynamics of the game has a globally asymptotically stable steady state if $\alpha - 2\beta^2 a^c > 0$. It can be shown that to guarantee this inequality, the expression of a^c where the square root is affected by a negative sign has to be chosen. From now on, by a^c we refer to this root. Let us note that with this choice coefficients a^c and b^c are negative and c^c is positive. Therefore, when the cooperative game is played the stock of pollution converges in the long run to the positive steady-state value, given by:

$$\frac{\beta(\gamma_1+\gamma_2+2\beta b^c)}{\alpha-2\beta^2 a^c} = \frac{X_1^c}{X_2^c}\,.$$

Player i's optimal payoff under the cooperative strategy is given by:

$$W_i(u_1^c, u_2^c) = \left[C_{1i}^c - \frac{X_1^c}{X_2^c}\left(C_{2i}^c + C_{3i}^c \frac{X_1^c}{X_2^c}\right)\right]\frac{1}{\rho}$$
$$- \left(x_0 - \frac{X_1^c}{X_2^c}\right)\left(C_{2i}^c + 2C_{3i}^c \frac{X_1^c}{X_2^c}\right)\frac{1}{\rho + X_2^c}$$
$$- C_{3i}^c \left(x_0 - \frac{X_1^c}{X_2^c}\right)^2 \frac{1}{\rho + 2X_2^c} ,$$

where

$$C_{1i}^c = \frac{1}{2}(\gamma_i + \beta b^c)(\gamma_i - \beta b^c), \quad C_{2i}^c = \beta^2 a^c b^c, \quad C_{3i}^c = \frac{1}{2}[(\beta a^c)^2 + \varphi_i] .$$

If $(\hat{u}_1, \hat{u}_2)$ is any pair of linear controls, given by

$$\hat{u}_i(x) = \hat{A}_i x + \hat{B}_i ,$$

the state trajectory associated is

$$\hat{x}(t) = \left(x_0 - \frac{\beta(\hat{B}_1 + \hat{B}_2)}{\alpha - \beta(\hat{A}_1 + \hat{A}_2)}\right)e^{-(\alpha - \beta(\hat{A}_1 + \hat{A}_2))t} + \frac{\beta(\hat{B}_1 + \hat{B}_2)}{\alpha - \beta(\hat{A}_1 + \hat{A}_2)} .$$

The state dynamics of the game has a globally asymptotically stable steady state if $\alpha - \beta(\hat{A}_1 + \hat{A}_2) > 0$.

Furthermore, player i's payoff is given by:

$$W_i(\hat{u}_1, \hat{u}_2) = \left[(\gamma_i - \hat{B}_i)\hat{A}_i \frac{\hat{X}_1}{\hat{X}_2} + \left(\gamma_i - \frac{1}{2}\hat{B}_i\right)\hat{B}_i - \frac{1}{2}(\hat{A}_i^2 + \varphi_i)\left(\frac{\hat{X}_1}{\hat{X}_2}\right)^2\right]\frac{1}{\rho}$$
$$+ \left(x_0 - \frac{\hat{X}_1}{\hat{X}_2}\right)\left[(\gamma_i - \hat{B}_i)\hat{A}_i - (\hat{A}_i^2 + \varphi_i)\frac{\hat{X}_1}{\hat{X}_2}\right]\frac{1}{\rho + \hat{X}_2}$$
$$- \frac{1}{2}(\hat{A}_i^2 + \varphi_i)\left(x_0 - \frac{\hat{X}_1}{\hat{X}_2}\right)^2 \frac{1}{\rho + 2\hat{X}_2} ,$$

where $\hat{X}_2$ is assumed to be positive in order to have convergent improper integrals and

$$\hat{X}_1 = \beta(\hat{B}_1 + \hat{B}_2), \quad \hat{X}_2 = \alpha - \beta(\hat{A}_1 + \hat{A}_2) .$$

4.1 Linear-Incentive Strategies

We focus now on the characterization of linear-incentive strategies of the type established in (6). An interior solution u_i^* $(i = 1, 2)$ of the optimal control problem (7) characterizing the incentive strategies satisfies:

$$u_i^*(x) = \gamma_i + \beta(1 + D_j)(a_i^* x + b_i^*), \quad i, j = 1, 2, i \neq j , \tag{17}$$

where constants a_i^*, b_i^* are determined using the Hamilton-Jacobi-Bellman equation associated with the optimal control problem. Assuming that the value function of problem (7) is given by $V_i^* = \frac{a_i^*}{2}x^2 + b_i^* x + c_i^*$, and using the standard technique, coefficients a_i^*, b_i^*, and c_i^* are determined:

$$a_i^* = \frac{2a^c\beta^2(D_j-1)+2\alpha+\rho \pm \sqrt{(2a^c\beta^2(D_j-1)+2\alpha+\rho)^2+4(1+D_j)^2\beta^2\varphi_i}}{2(1+D_j)^2\beta^2},$$

$$b_i^* = \frac{a_i^*\beta(\gamma_i+\gamma_j-b^c(D_j-1)\beta)}{a^c(D_j-1)\beta^2-a_i^*(1+D_j)^2\beta^2+\alpha+\rho},$$

$$c_i^* = \frac{\gamma_i(\gamma_i+2b_i^*\beta)+b_i^*\beta[2\gamma_j+(-2b^c(D_j-1)+b_i^*(1+D_j)^2)\beta]}{2\rho}.$$

To determine a linear-incentive equilibrium we need to impose

$$u_i^* = u_i^c, \quad i = 1,2,$$

where u_i^* and u_i^c are given by (17) and (16), respectively. From the above equality we obtain the following conditions:

$$(1+D_j)a_i^* = a^c, \quad (1+D_j)b_i^* = b^c, \quad i,j = 1,2, i \neq j. \tag{18}$$

Therefore, the following ratios must be satisfied:

$$\frac{a^c}{b^c} = \frac{a_i^*}{b_i^*}, \quad i = 1,2.$$

The previous equalities once the expressions of b^c and b_i^* in terms of a_c and a_i^*, respectively, have been replaced, after some computations simplify as follows:

$$-2a^cD_j + a_i^*(1+D_j)^2 = 0, \quad i = 1,2.$$

Taking into account the first condition in (18), one has that the slope of the linear-incentive equilibrium strategy is unitary, that is,

$$D_j = 1, \quad j = 1,2,$$

and the incentive strategies are given by:

$$u_j = \psi_j(u_i) = u_j^c - u_i^c + u_i = \gamma_j - \gamma_i + u_i, \ \ i,j = 1,2, \ i \neq j.$$

From the last equality we have that firm j's emissions are greater (lower) than those of firm i if and only if A_j is greater (lower) than A_i.

Coefficients a_i^*, b_i^*, and c_i^* under assumption $D_j = 1$ read:

$$a_i^* = \frac{2\alpha + \rho \pm \sqrt{(2\alpha+\rho)^2 + 16\beta^2\varphi}}{8\beta^2} ,$$
$$b_i^* = \frac{\beta(\gamma_i + \gamma_j)a_i^*}{\alpha + \rho - 4\beta^2 a_i^*} ,$$
$$c_i^* = \frac{\gamma_i(\gamma_i + 2\beta b_i^*) + 2\beta b_i^*(\gamma_j + 2\beta b_i^*)}{2\rho} .$$

Note that, in this case, these expressions are independent of the coefficients of the value function associated with the cooperative game.

From the conditions in (18), one has $a^c = 2a_i^*, b^c = 2b_i^*$. An easy computation establishes that these conditions are satisfied if and only if $\varphi_i = \varphi_j$. That is, both firms must have the same damage cost parameter. From now on, let denote by φ this common damage cost.

We focus on credible linear-incentive strategies for any deviation in set F_i as defined in (15). The fulfillment of inequalities in (23) guarantees that the linear-incentives strategies are credible when linear deviations in set F_i are considered.

The two sides of the inequality for player i after some computations read:

$$\begin{aligned} W_i(\psi_i(u_j), u_j) = &\left[\tilde{C}_{1i} - \frac{\tilde{X}_1}{\tilde{X}_2}\left(\tilde{C}_{2i} + \tilde{C}_{3i}\frac{\tilde{X}_1}{\tilde{X}_2}\right)\right]\frac{1}{\rho} \\ &- \left(x_0 - \frac{\tilde{X}_1}{\tilde{X}_2}\right)\left(\tilde{C}_{2i} + 2\tilde{C}_{3i}\frac{\tilde{X}_1}{\tilde{X}_2}\right)\frac{1}{\rho + \tilde{X}_2} \\ &- \tilde{C}_{3i}\left(x_0 - \frac{\tilde{X}_1}{\tilde{X}_2}\right)^2 \frac{1}{\rho + 2\tilde{X}_2} , \\ W_i(u_i^c, u_j) = &\left[C'_{1i} - \frac{X'_1}{X'_2}\left(C'_{2i} + C'_{3i}\frac{X'_1}{X'_2}\right)\right]\frac{1}{\rho} \\ &- \left(x_0 - \frac{X'_1}{X'_2}\right)\left(C'_{2i} + 2C'_{3i}\frac{X'_1}{X'_2}\right)\frac{1}{\rho + X'_2} \\ &- C'_{3i}\left(x_0 - \frac{X'_1}{X'_2}\right)^2 \frac{1}{\rho + 2X'_2} , \end{aligned}$$

where

$$\tilde{C}_{1i} = \frac{1}{2}(\gamma_i^2 - (\gamma_j - \tilde{B}_j)^2), \quad \tilde{C}_{2i} = -(\gamma_j - \tilde{B}_j)\tilde{A}_j, \quad \tilde{C}_{3i} = \frac{1}{2}(\tilde{A}_j^2 + \varphi) ,$$
$$\tilde{X}_1 = \beta(\gamma_i - \gamma_j + 2\tilde{B}_j), \quad \tilde{X}_2 = \delta - 2\beta\tilde{A}_j ,$$
$$C'_{1i} = \frac{1}{2}(\gamma_i^2 - (\beta b^c)^2), \quad C'_{2i} = \beta^2 a^c b^c, \quad C'_{3i} = \frac{1}{2}((\beta a^c)^2 + \varphi) ,$$
$$X'_1 = \beta(\gamma_i + \beta b^c + \tilde{B}_j), \quad X'_2 = \delta - \beta^2 a^c - \beta\tilde{A}_j ,$$

and $\tilde{X}_2$ and X'_2 are assumed to be positive in order to have globally stable steady states, which also guarantee convergent improper integrals.

In the noncooperative case, it is easy to establish that the stationary Markovian Nash equilibrium strategies are given by $u_i^N(x) = \gamma_i + \beta(a^N x + b^N)$, where the value functions are by $V_i^{nc}(x) = \frac{a^N}{2}x^2 + b^N x + c_i^N$ and the superscript N stands for Nash equilibrium. Note that the quadratic and linear coefficients are the same for all the players, due to symmetry in objectives and dynamics.

Following the standard approach for solving HJB equations, we obtain the following coefficients for the noncooperative value functions:

$$a^N = \frac{\rho + 2\alpha - \sqrt{(\rho + 2\alpha)^2 + 12\beta^2\varphi}}{6\beta^2} < 0 ,$$
$$b^N = -\frac{a^N\beta(\gamma_1 + \gamma_2)}{3\beta^2 a^N - (\rho + \alpha)} < 0 ,$$
$$c_i^N = \frac{\gamma_i^2 + 2\beta b^N(\gamma_1 + \gamma_2) + 3(\beta b^N)^2}{2\rho} .$$

The equilibrium noncooperative state trajectory is given by:

$$x^N(t) = \left(x_0 - \frac{\beta\left(\gamma_1 + \gamma_2 + 2\beta b^N\right)}{\alpha - 2\beta^2 a^N}\right) e^{(\alpha - 2\beta^2 a^N)t} + \frac{\beta\left(\gamma_1 + \gamma_2 + 2\beta b^N\right)}{\alpha - 2\beta^2 a^N} .$$

The state dynamics of the game has a globally asymptotically stable steady state if $\alpha - 2\beta^2 a^N > 0$. It can be proved that to guarantee this inequality and therefore, the globally asymptotically stable steady state, the only possibility is to choose $a^N < 0$.

As usual, in this kind of model, the non-cooperative solution leads to emission levels greater than those prescribed by the cooperative solution. Then, one can suppose that if one player deviates from the cooperative solution, he is going to choose an emission level greater than that corresponding to the cooperative solution.

Therefore, we try to characterize emission levels that correspond with linear deviations in set F_i that are greater than the cooperative ones and for which the credibility conditions in (23) are fulfilled.

These last inequalities are quite complex and involve all the parameters of the model as well as $\tilde{A}_j$, $\tilde{B}_j$ coefficients of the emission levels along the deviation from the cooperative solution. For that reason, it is not possible to analytically characterize the feasible set for $\tilde{A}_j$ and $\tilde{B}_j$ which ensure credibility of the incentive strategies. We resort to numerical simulations to carry out this analysis.

Let consider the following values of the model parameters:

$$\varphi_1 = \varphi_2 = \varphi = 1,\ \beta = 1,\ \rho = 0.1,\ \alpha = 0.2,\ \gamma_1 = 1.01,\ \gamma_2 = 1,\ S_0 = 0.05\ .$$

Under this assumption, the coefficients of the cooperative and non-cooperative game are given by

$$a^c = -0.8828, \quad b^c = -0.8590, \quad a^N = -0.5, \quad b^N = -0.5583\ ,$$

and the cooperative and non-cooperative emission strategies are:

$$\begin{aligned} u_1^c(x) &= 0.1510 - 0.8828x, \qquad & u_2^c(x) &= 0.1410 - 0.8828x\ , \\ u_1^N(x) &= 0.4517 - 0.5x, \qquad & u_2^N(x) &= 0.4417 - 0.5x\ . \end{aligned}$$

The incentive equilibrium strategies associated with linear deviations of the form $\tilde{u}_i(x) = \tilde{A}_i x + \tilde{B}_i, i = 1, 2$, are:

$$\begin{aligned} u_1(x) &= \psi_1(\tilde{u}_2(x)) = \gamma_1 - \gamma_2 + \tilde{u}_2(x) = 0.01 + \tilde{A}_2 x + \tilde{B}_2\ , \\ u_2(x) &= \psi_2(\tilde{u}_1(x)) = \gamma_2 - \gamma_1 + \tilde{u}_1(x) = -0.01 + \tilde{A}_1 x + \tilde{B}_1\ . \end{aligned}$$

The credibility conditions for player 1 and 2 read:

$$\begin{aligned} 0.0010\tilde{A}_2^5 &+ (-4.4360 + 0.04\tilde{B}_2)\tilde{A}_2^4 + 0.4(-72.7103 + \tilde{B}_2)(-0.2635 + \tilde{B}_2)\tilde{A}_2^3 \\ &+ 11.2426(-0.0501 + \tilde{B}_2)(4.9011 + \tilde{B}_2)\tilde{A}_2^2 \\ &+ 30.4507(-0.5656 + \tilde{B}_2)(-0.0194 + \tilde{B}_2)\tilde{A}_2 \\ &- 21.3055(-0.0423 + \tilde{B}_2)(-0.0148 + \tilde{B}_2) \geq 0\ , \end{aligned}$$

$$\begin{aligned} 0.0010\tilde{A}_1^5 &+ (-4.4364 + 0.01\tilde{B}_1)\tilde{A}_1^4 + 0.4(-72.7203 + \tilde{B}_1)(-0.2735 + \tilde{B}_1)\tilde{A}_1^3 \\ &+ 11.2426(-0.0601 + \tilde{B}_1)(4.8911 + \tilde{B}_1)\tilde{A}_1^2 \\ &+ 30.4507(-0.5756 + \tilde{B}_1)(-0.0248 + \tilde{B}_1)\tilde{A}_1 \\ &- 21.3055(-0.0526 + \tilde{B}_1)(-0.0248 + \tilde{B}_1) \geq 0\ . \end{aligned}$$

Plotting the left-hand side of the previous inequalities allows us to determine the intervals for parameters $\tilde{A}_i, \tilde{B}_i, i = 1, 2$ that lead to credible incentive strategies. The credible ranges are as follows:

$$\tilde{A}_1 \in [-0.85, -0.7],\ \tilde{B}_1 \in [0.15, 0.2],\ \tilde{A}_2 \in [-0.9, -0.7],\ \tilde{B}_2 \in [0.15, 0.2]\ .$$

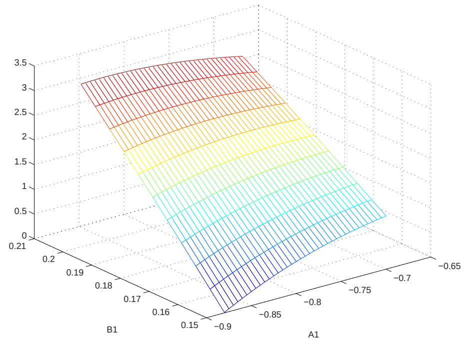

Figure 1: Credible set for deviations of player 1.

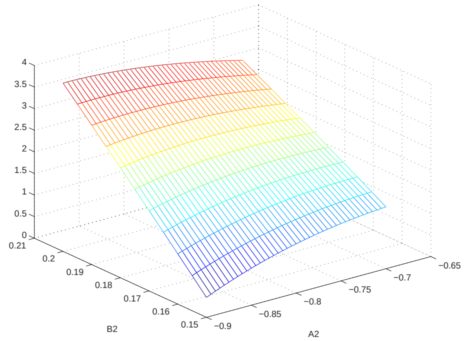

Figure 2: Credible set for deviations of player 2.

Figures 1 and 2 show the credibility conditions for player 2 and 1 for deviations of the opponent in the credible ranges. Figures 3 and 4 show these conditions for wider ranges of the deviations parameters, and therefore, leading to non-credible strategies out of the credible ranges.

For linear deviations associated with parameter values in these intervals the following results apply. First, the emission levels as well as the firms' benefits are positive for any value of the pollution stock. Second, along the deviation the emission levels of both players are greater than those agreed in the cooperative game but lower than those prescribed by the completely non-cooperative Nash game.

Note that the admissible set of deviations for player 1 leading to credible incentive strategies is wider than that of player 2. This result can be explained by the asymmetric assumption for coefficients γ_1 and γ_2 associated to the revenue functions of both players.

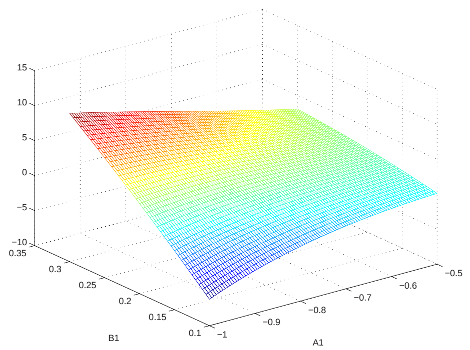

Figure 3: Credibility condition of player 2 for deviations of player 1.

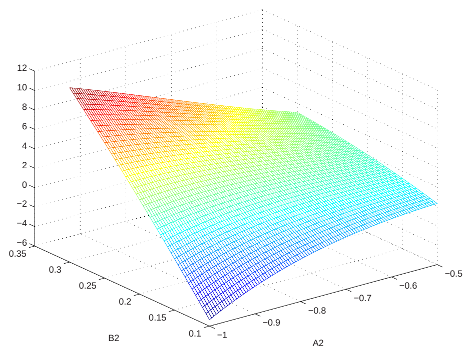

Figure 4: Credibility condition of player 1 for deviations of player 2.

Tables 1 and 2 show the emission levels of both players under cooperation (superscript c), non-cooperation (Nash game, superscript N) and when the incentive strategies are implemented ($\psi_i(\tilde{u}_j), i, j = 1, 2, i \neq j$) for a linear deviation of the rival ($\tilde{u}_j(x) = \tilde{A}_j x + \tilde{B}_j$). Both tables collect the results for different deviation levels for player 1, Table 1, and player 2, Table 2.

From these results we can conclude that only small deviations from the cooperative emission levels lead to credible strategies.

5 Concluding Remarks

We characterized in this paper the conditions to have credible linear-incentive equilibrium strategies in the class of linear-quadratic differential games. The lat-

Table 1: Emission levels for an initial pollution stock $S_0 = 0.05$.
Different deviations of player 1.

	$\tilde{A}_1 = -0.85$ $\tilde{B}_1 = 0.15,\ \tilde{B}_1 = 0.2$		$\tilde{A}_1 = -0.8$ $\tilde{B}_1 = 0.15,\ \tilde{B}_1 = 0.2$		$\tilde{A}_1 = -0.7$ $\tilde{B}_1 = 0.15,\ \tilde{B}_1 = 0.2$	
$\tilde{u}_1$	0.1075	0.1575	0.11	0.16	0.115	0.165
$u_2 = \psi_2(\tilde{u}_1)$	0.0975	0.1475	0.10	0.15	0.105	0.105
u_1^c	0.1068		0.1068		0.1068	
u_2^c	0.0968		0.0968		0.0968	
u_1^N	0.4267		0.4267		0.4267	
u_2^N	0.4167		0.4167		0.4167	

Table 2: Emission levels for an initial pollution stock $S_0 = 0.05$.
Different deviations of player 2.

	$\tilde{A}_2 = -0.9$ $\tilde{B}_2 = 0.15,\ \tilde{B}_2 = 0.2$		$\tilde{A}_2 = -0.8$ $\tilde{B}_2 = 0.15,\ \tilde{B}_2 = 0.2$		$\tilde{A}_2 = -0.7$ $\tilde{B}_2 = 0.15,\ \tilde{B}_2 = 0.2$	
$\tilde{u}_2$	0.105	0.155	0.11	0.16	0.115	0.165
$u_1 = \psi_1(\tilde{u}_2)$	0.115	0.165	0.12	0.17	0.125	0.175
u_1^c	0.1068		0.1068		0.1068	
u_2^c	0.0968		0.0968		0.0968	
u_1^N	0.4267		0.4267		0.4267	
u_2^N	0.4167		0.4167		0.4167	

ter is popular in economics and management science applications of differential games, because it allows the inclusion in the models of important features such as economics of scale, decreasing marginal returns, etc., while still being tractable (i.e., admitting analytical solutions). One message of this paper is that obtaining credibility for even linear-incentive strategies is tedious. A second one is that credibility may be at hand for only small deviations with respect of the cooperative (or agreed-upon) strategies.

Appendix A

Coefficients a^c, b^c, and c^c of the cooperative value function (Proposition 2.1)

Coefficient a^c satisfies the following quadratic equation:

$$\Xi_1 (a^c)^2 + \Xi_2 a^c + \Xi_3 = 0 \ ,$$

where

$$\Xi_1 = -\left(R^i_{jj} + R^j_{jj}\right) {\beta_i}^2 + 2\left(R^i_{ij} + R^j_{ij}\right) \beta_i \beta_j - \left(R^i_{ii} + R^j_{ii}\right) {\beta_j}^2 \ ,$$

$$\Xi_2 = -2\left(d_{ij}+d_{jj}\right)\left(-\left(R^i_{ij}+R^j_{ij}\right)\beta_i+\left(R^i_{ii}+R^j_{ii}\right)\ \beta_j\right)$$
$$-\,2\left(d_{ii}+d_{ji}\right)\left(\left(R^i_{jj}+R^j_{jj}\right)\beta_i-\left(R^i_{ij}+R^j_{ij}\right)\beta_j\right)$$
$$+\left(\left(R^i_{ij}+R^j_{ij}\right)^2-\left(R^i_{ii}+R^j_{ii}\right)\left(R^i_{jj}+R^j_{jj}\right)\right)(-2\alpha+\rho)\ ,$$

$$\Xi_3 = (m_i+m_j)\left(-\left(R^i_{ij}+R^j_{ij}\right)^2+\left(R^i_{ii}+R^j_{ii}\right)\ \left(R^i_{jj}+R^j_{jj}\right)\right)$$
$$-\left(d_{ij}+d_{jj}\right)^2\left(R^i_{ii}+R^j_{ii}\right)+2\left(d_{ii}+d_{ji}\right)\left(d_{ij}+d_{jj}\right)\left(R^i_{ij}+R^j_{ij}\right)$$
$$-\left(d_{ii}+d_{ji}\right)^2\left(R^i_{jj}+R^j_{jj}\right)\ .$$

Coefficient b^c solves the following equation:

$$\Sigma_1 b^c+\Sigma_2=0\ ,$$

where

$$\Sigma_1=\left(d_{ij}+d_{jj}\right)\left(\left(R^i_{ij}+R^j_{ij}\right)\beta_i-\left(R^i_{ii}+R^j_{ii}\right)\beta_j\right)$$
$$-\left(d_{ii}+d_{ji}\right)\left((R^i_{jj}+R^j_{jj})\beta_i-\left(R^i_{ij}+R^j_{ij}\right)\beta_j\right)$$
$$+\,a^c\left(-\left(R^i_{jj}+R^j_{jj}\right){\beta_i}^2+2\left(R^i_{ij}+R^j_{ij}\right)\beta_i\beta_j-\left(R^i_{ii}+R^j_{ii}\right)\ {\beta_j}^2\right)$$
$$+\left(\left(R^i_{ij}+R^j_{ij}\right)^2-\left(R^i_{ii}+R^j_{ii}\right)\left(R^i_{jj}\beta_i+R^j_{jj}\right)\right)(-\alpha+\rho)\ ,$$
$$\Sigma_2=\left(d_{ij}+d_{jj}\right)\left(\left(R^i_{ij}+R^j_{ij}\right)(r_{ii}+r_{ji})-\left(R^i_{ii}+R^j_{ii}\right)(r_{ij}+r_{jj})\right)$$
$$+\left(q_i+q_j\right)\left(-\left(R^i_{ij}+R^j_{ij}\right)^2+\left(R^i_{ii}+R^j_{ii}\right)\left(R^i_{jj}\beta_i+R^j_{jj}\right)\right)$$
$$+\left(d_{ii}+d_{ji}\right)\left(\left(R^i_{ij}+R^j_{ij}\right)(r_{ij}+r_{jj})-(r_{ii}+r_{ji})\left(R^i_{jj}+R^j_{jj}\right)\right)$$
$$+\,a^c\Big[\left(\left(R^i_{ij}+R^j_{ij}\right)(r_{ij}+r_{jj})-(r_{ii}+r_{ji})\left(R^i_{jj}+R^j_{jj}\right)\right)\beta_i$$
$$+\left(\left(R^i_{ij}+R^j_{ij}\right)(r_{ii}+r_{ji})-\left(R^i_{ii}+R^j_{ii}\right)(r_{ij}+r_{jj})\right)\beta_j\Big]\ .$$

Coefficient c^c satisfies the following equation:

$$\rho\left(1+c^c\right)\left(\left(R^i_{ij}+R^j_{ij}\right)^2-\left(R^i_{ii}+R^j_{ii}\right)\left(R^i_{jj}+R^j_{jj}\right)\right)$$
$$+\,b^c\Big[\left(\left(R^i_{ij}+R^j_{ij}\right)(r_{ij}+r_{jj})-(r_{ii}+r_{ji})\left(R^i_{jj}+R^j_{jj}\right)\right)\beta_i$$
$$+\left(\left(R^i_{ij}+R^j_{ij}\right)(r_{ii}+r_{ji})-\left(R^i_{ii}+R^j_{ii}\right)(r_{ij}+r_{jj})\right)\beta_j\Big]=0\ .$$

Coefficients a_i^*, b_i^*, and c_i^* of the optimal value function associated to the incentive optimal control problem (Proposition 3.1)

Coefficient a_i^* satisfies the following quadratic equation:

$$\Delta_1 (a_i^*)^2 + \Delta_2 a_i^* + \Delta_3 = 0 \,,$$

where

$$\begin{aligned}
\Delta_1 &= (\beta_i + D_j\beta_j)^2, \\
\Delta_2 &= 2\left((d_{ii} + d_{ij}D_j)(\beta_i + D_j\beta_j)\right. \\
&\quad + \left(A_j^c - A_i^c D_j\right)\left(\left(R_{ij}^i + D_j R_{jj}^i\right)\beta_i - \left(R_{ii}^i + D_j R_{ij}^i\right)\beta_j\right) \\
&\quad \left. + \left(R_{ii}^i + D_j\left(2R_{ij}^i + D_j R_{jj}^i\right)\right)(-2\alpha + \rho)\right) \,, \\
\Delta_3 &= (d_{ii} + d_{ij}D_j)^2 \\
&\quad + \left(A_j^c - A_i^c D_j\right)^2 \left({R_{ij}^i}^2 - R_{ii}^i R_{jj}^i\right) - m_i\left(R_{ii}^i + D_j\left(2R_{ij}^i + D_j R_{jj}^i\right)\right) \\
&\quad + 2\left(A_j^c - A_i^c D_j\right)\left(-d_{ij}\left(R_{ii}^i D_j R_{ij}^i\right) + d_{ii}\left(R_{ij}^i + D_j R_{jj}^i\right)\right) \,.
\end{aligned}$$

Coefficient b_i^* solves the following equation:

$$\Lambda_1 b_i^* + \Lambda_2 a_i^* + \Lambda_3 = 0 \,,$$

where:

$$\begin{aligned}
\Lambda_1 &= (d_{ii} + d_{ij}D_j)(\beta_i + D_j\beta_j) + a_i^*(\beta_i + D_j\beta_j)^2 \\
&\quad + \left(R_{ii}^i + D_j\left(2R_{ij}^i + D_j R_{jj}^i\right)\right)(-\alpha + \rho) \\
&\quad + \left(A_j^c - A_i^c D_j\right)\left(\left(R_{ij}^i + D_j R_{jj}^i\right)\beta_i - \left(R_{ii}^i + D_j R_{ij}^i\right)\beta_j\right) \,, \\
\Lambda_2 &= (r_{ii} + D_j r_{ij})(\beta_i + D_j\beta_j) \\
&\quad + \left(B_j^c - B_i^c D_j\right)\left(R_{ij}^i\beta_i - R_{ii}^i\beta_j + D_j\left(R_{jj}^i\beta_i - R_{ij}^i\beta_j\right)\right) \,, \\
\Lambda_3 &= r_{ij}\left(D_j(d_{ii} + d_{ij}D_j) - \left(A_j^c - A_i^c D_j\right)\left(R_{ii}^i + D_j R_{ij}^i\right)\right) \\
&\quad + q_i\left(-R_{ii}^i - D_j\left(2R_{ij}^i + D_j R_{jj}^i\right)\right) \\
&\quad + \left(B_j^c - B_i^c D_j\right)\left(-d_{ij}\left(R_{ii}^i + D_j R_{ij}^i\right) + d_{ii}\left(R_{ij}^i + D_j R_{jj}^i\right)\right. \\
&\quad + \left(A_j^c - A_i^c D_j\right)\left(\left(R_{ij}^i\right)^2 - R_{ii}^i R_{jj}^i\right)\Big) \\
&\quad + r_{ii}\left(d_{ii} + d_{ij}D_j + \left(A_j^c - A_i^c D_j\right)\left(R_{ij}^i + D_j R_{jj}^i\right)\right) \,.
\end{aligned}$$

Coefficient c_i^* is characterized by the following equation:

$$\Theta_1 c_i^* + \Theta_2 b_i^* + \Theta_3 = 0 \,,$$

where

$$\begin{aligned}
\Theta_1 &= 2\,\left(R^i_{ii} + D_j\left(2R^i_{ij} + D_j R^i_{jj}\right)\right)\ ,\\
\Theta_2 &= 2\left(r_{ii} + D_j r_{ij}\right)\left(\beta_i + D_j\beta_j\right)\\
&\quad + 2\left(B^c_j - B^c_i D_j\right)\left(D_j R^i_{jj}\beta_i - R^i_{ii}\beta_j + R^i_{ij}\left(\beta_i - D_j\beta_j\right)\right)\ ,\\
\Theta_3 &= \ {r_{ii}}^2 + 2D_j r_{ij} + D_j r_{ij}\left(-2 + 2r_{ii} + D_j r_{ij}\right)\\
&\quad + \left(B^c_j - B^c_i D_j\right)^2\left(\ \left(R^i_{ij}\right)^2 - R^i_{ii}R^i_{jj}\right)\\
&\quad + 2\left(B^c_j - B^c_i D_j\right)\ \left(-\left(r_{ij}\left(R^i_{ii} + D_j R^i_{ij}\right)\right) + r_{ii}\left(R^i_{ij} + D_j R^i_{jj}\right)\right)\ .
\end{aligned}$$

Coefficients $T^i_{kl},\ k = 1,2,\ l = 1,\ldots,3, i = 1,2$ of the equations characterizing the incentive equilibrium strategies (Proposition 3.2)

$$\begin{aligned}
T^i_{11} &= -R^i_{jj}\left(\left(R^i_{jj} + R^j_{jj}\right)\left(d_{ii} + d_{ji} + A^c_i\left(R^i_{ii} + R^j_{ii}\right) + a^c\beta_i\right)\right.\\
&\quad \left. -\ \left(R^i_{ij} + R^j_{ij}\right)\left(d_{ij} + d_{jj} + A^c_i\left(R^i_{ij} + R^j_{ij}\right) + a^c\beta_j\right)\right)\ ,\\
T^i_{12} &= 2d_{jj}R^i_{ij}\left(R^i_{ij} + R^j_{ij}\right) - 2\left(d_{ii} + d_{ji}\right)R^i_{ij}\left(R^i_{jj} + R^j_{jj}\right)\\
&\quad - 2a^c R^i_{ij}\left(R^i_{jj} + R^j_{jj}\right)\beta_i\\
&\quad + \left(A^c_i R^i_{ij} - A^c_j R^i_{jj}\right)\left(\left(R^i_{ij} + R^j_{ij}\right)^2 - \left(R^i_{ii} + R^j_{ii}\right)\left(R^i_{jj} + R^j_{jj}\right)\right)\\
&\quad + d_{ji}\left(\left(R^i_{ij} - R^j_{ij}\right)\left(R^i_{ij} + R^j_{ij}\right) + \left(R^i_{ii} + R^j_{ii}\right)\ \left(R^i_{jj} + R^j_{jj}\right)\right)\\
&\quad + \left(\left(2a^c - a^*_i\right)R^i_{ij}\left(R^i_{ij} + R^j_{ij}\right)\right.\\
&\quad \left. + a^*_i\left(-{R^j_{ij}}^2 + \left(R^i_{ii} + R^j_{ii}\right)\left(R^i_{jj} + R^j_{jj}\right)\right)\right)\beta_j\ ,\\
T^i_{13} &= -\left(R^i_{jj} + R^j_{jj}\right)\left(-d_{ii}R^j_{ii} + R^i_{ii}\left(d_{ji} + a^c\beta_i\right)\right.\\
&\quad \left. -\left(R^i_{ii} + R^j_{ii}\right)\left(-A^c_j R^i_{ij} + a^*_i\beta_i\right)\right) - \left(R^i_{jj} + R^j_{jj}\right)\\
&\quad \left((d_{ji} + a^c\beta_i)R^i_{ii} - d_{ii}R^j_{ii} - (a^*_i\beta_i - A^c_j R^i_{ij})(R^i_{ii} + R^j_{ii})\right)\ ,\\
T^i_{21} &= R^i_{jj}\left(\left(R^i_{jj} + R^j_{jj}\right)\left(r_{ii} + B^c_i\left(R^i_{ii} + R^j_{ii}\right) + r_{ji} + b^c\beta i\right)\right.\\
&\quad \left. -\ \left(R^i_{ij} + R^j_{ij}\right)\left(-r_{ij} + B^c_i\left(R^i_{ij} + R^j_{ij}\right) + r_{jj} + b^c\beta_j\right)\right)\ ,
\end{aligned}$$

$$
\begin{aligned}
T_{22}^i = &-\left(R_{ij}^i + R_{ij}^j\right)\left(r_{ij}\left(R_{ij}^i + R_{ij}^j\right) + 2R_{ij}^i r_{jj}\right) \\
&- \left(\left(R_{ii}^i + R_{ii}^j\right) r_{ij} - 2R_{ij}^i\left(r_{ii} + r_{ji}\right)\right)\left(R_{jj}^i + R_{jj}^j\right) \\
&+ \left(\left(R_{ij}^i + R_{ij}^j\right)^2 - \left(R_{ii}^i + R_{ii}^j\right)\left(R_{jj}^i + R_{jj}^j\right)\right) \times \\
&\left(-B_i^c R_{ij}^i + B_j^c R_{jj}^i + b_i^* \beta_j\right) \\
&+ 2b^c R_{ij}^i \left(\left(R_{jj}^i + R_{jj}^j\right)\beta_i - \left(R_{ij}^i + R_{ij}^j\right)\beta_j\right)\ , \\
T_{23}^i = &\left(\left(R_{ij}^i + R_{ij}^j\right)^2 - R_{ii}^j\left(R_{jj}^i + R_{jj}^j\right)\right)\left(r_{ii} + B_j^c R_{ij}^i + b_i^* \beta_i\right) \\
&- R_{ii}^i\left(-\left(R_{jj}^i + R_{jj}^j\right)\left(-B_j^c R_{ij}^i + r_{ji} + \left(b^c - b_i^*\right)\beta_i\right)\right. \\
&+ \left.\left(R_{ij}^i + R_{ij}^j\right)\left(r_{ij} + r_{jj} + b^c\beta_j\right)\right)\ .
\end{aligned}
$$

Appendix B: A Homogeneous LQ Example

The class of homogeneous LQDG is characterized by the absence of linear terms in the objective functional of player i, i.e., $q_i = r_{ij} = 0$ in (1) for all $i, j \in \{1, 2\}$. Player i's optimization problem then becomes:

$$
\max_{u_i}\left\{W_i = \int_{t_0}^{\infty} e^{-\rho t}\left[\frac{1}{2}[m_i x^2 + u' R_i u] + \sum_{j=1}^{2} d_{ij} u_j x\right] dt\right\} \tag{19}
$$

$$
\dot{x} = \alpha x + \sum_{j=1}^{2} \beta_j u_j, \quad x(t_0) = x_0\ .
$$

Economic applications which considered homogeneous LQDG include Neck and Dockner (1987), Cohen and Michel (1988), and Aarle et al. (1995). The model in Neck and Dockner (1987) is fully symmetric and assumes that the control and state variables do not interact (i.e., $d_{ij} = 0\ \forall i, j \in \{1, 2\}$). This last assumption is also made in the two-player differential game in Aarle et al. (1995). Cohen and Michel (1988) study a game that fits exactly the formulation in (19).

It is well known that if the players cooperate by maximizing the sum of their objective functionals in a homogeneous LQDG, then:

- the optimal strategies are linear in the state: $u_i^c(x) = \gamma_i^c x$;
- the value function is quadratic, $V^c(x) = r^c x^2$, where r^c is a solution to a second-degree polynomial.

Example 5.1. Let player i's optimization problem be given by

$$\max_{u_i}\left\{W_i=\int_{t_0}^{\infty} e^{-\rho t}[g_i u_i^2+m_i x^2]dt\right\} ,$$
$$\text{s.t.: } \dot{x}=\alpha x+\sum_{j=1}^{2}\beta_j u_j,\ \ x(t_0)=x_0 ,$$

where $g_i \neq 0$.

In the cooperative (joint maximization) case, it is easy to check that the optimal linear strategies and quadratic value function are

$$u_i^c(x)=-\frac{r^c\beta_i}{g_i}x, \qquad V^c(x)=r^c x^2 , \tag{20}$$

where r^c is real and given by

$$r^c=\frac{2\alpha-\rho\pm\sqrt{(2\alpha-\rho)^2+4\sum_{j=1}^{2}m_j\sum_{j=1}^{n}\frac{\beta_j^2}{g_j}}}{2\sum_{j=1}^{2}\frac{\beta_j^2}{g_j}} ,$$

if $(2\alpha-\rho)^2+4\sum_{j=1}^{2}m_j\sum_{j=1}^{2}\frac{\beta_j^2}{g_j}\geq 0$.

The optimal state trajectory is

$$x^c(t)=x_0 e^{\left(-r^c\sum_{j=1}^{2}\frac{\beta_j^2}{g_j}+\alpha\right)t} .$$

The state dynamics of the game has a globally asymptotically stable steady state if $-r^c\sum_{j=1}^{2}\frac{\beta_j^2}{g_j}+\alpha<0$.

Player i' s optimal payoff in the cooperative game is given by:

$$W_i(u_1^c,u_2^c)=\left((r^c)^2\frac{\beta_i^2}{g_i}+m_i\right)\frac{1}{\rho-2\alpha+2r^c\sum_{j=1}^{2}\frac{\beta_j^2}{g_j}}x_0^2 .$$

Let $(\hat{u}_1,\hat{u}_2)$ be any pair of linear controls, given by

$$\hat{u}_i(x)=\hat{A}_i x+\hat{B}_i .$$

The corresponding state trajectory is then

$$\hat{x}(t)=x_0 e^{(\beta_1 A_1+\beta_2 A_2+\alpha)t} .$$

The state dynamics of the game has a globally asymptotically stable steady state if $\beta_1 A_1+\beta_2 A_2+\alpha<0$.

Furthermore, player i's payoff is given by:

$$W_i(\hat{u}_1, \hat{u}_2) = (g_i A_i^2 + m_i) \frac{1}{\rho - 2(\beta_1 A_1 + \beta_2 A_2 + \alpha)} x_0^2 ,$$

where $\rho - 2(\beta_1 A_1 + \beta_2 A_2 + \alpha)$ is assumed to be positive in order to have a convergent improper integral.

We focus now on the characterization of linear-incentive strategies of the type established in (6).

An interior solution u_i^* $(i = 1, 2)$ of the optimal control problem (7) characterizing the incentive strategies satisfies:

$$u_i^*(x) = -\frac{r_i^*(\beta_i + \beta_j D_j)x}{g_i}, \quad i, j = 1, 2, , i \neq j , \tag{21}$$

where constant r_i^* is determined using the Hamilton-Jacobi-Bellman equation associated with the optimal control problem. Assuming that the value function of problem (7) is given by $V_i^* = r_i^* x^2$, and using the standard technique, the coefficient r_i^* is given by:

$$r_i^* = \frac{2\alpha - \rho - 2\beta_j r^c \left(\frac{\beta_j}{g_j} - \beta_i\right) \pm \sqrt{\left[2\alpha - \rho - 2\beta_j r^c \left(\frac{\beta_j}{g_j} - \beta_i\right)\right]^2 + 4m_i \frac{(\beta_i + \beta_j D_j)^2}{g_i}}}{2\frac{(\beta_i + \beta_j D_j)^2}{g_i}} .$$

To determine a linear-incentive equilibrium we need to impose

$$u_i^* = u_i^c, \quad i = 1, 2 ,$$

where u_i^* and u_i^c are given by (21) and (20), respectively. From the above equality, we obtain the slope of the linear-incentive equilibrium strategy:

$$D_j = \left(\frac{r^c}{r_i^*} - 1\right) \frac{\beta_i}{\beta_j} . \tag{22}$$

Replacing this expression in (21), u_i^* can be rewritten as:

$$r_i^* = \frac{m_i g_i - (r^c)^2 \beta_i^2}{g_i \left[\rho + 2\beta_j r^c \left(\frac{\beta_j}{g_j} - \beta_i\right) - 2\alpha\right]} .$$

Substituting this last expression in (22), D_j can be expressed as a function only of r^c as follows:

$$D_j = \left(\frac{r^c g_i \left[\rho - 2\alpha + 2\beta_j \left(\frac{\beta_j}{g_j} - \beta_i\right) r^c\right]}{m_i g_i - \beta_i^2 (r^c)^2} - 1\right) \frac{\beta_i}{\beta_j} .$$

As in Sec. 3.1, in order to study the credibility of the incentive equilibrium strategies we assume that if one player deviates from the cooperative solution she will use a linear control as given in (14). Therefore, we focus on credible linear-incentive strategies for any deviation in set F_i as defined in (15).

To assess the credibility of the linear-incentive strategies for linear deviations in set F_i, the following inequalities must to be satisfied:

$$W_i(\psi(u_j), u_j) \geq W_i(u_i^c, u_j) \quad \forall u_j \in F_j,\ i,j = 1,2, i \neq j\ . \tag{23}$$

The left- and right-hand sides of last inequality for player i after easy computations read:

$$W_i(\psi(u_j), u_j) = \left\{ g_i \left[-\frac{\beta_i}{g_i} r^c + D_i \left(A_j + \frac{\beta_j}{g_j} r^c \right) \right]^2 + m_i \right\} \frac{1}{\rho - 2\Gamma_i} x_0^2\ ,$$

$$W_i(u_i^c, u_j) = \left(\beta_i^2 \frac{(r^c)^2}{g_i} + m_i \right) \frac{1}{\rho - 2\Lambda_i} x_0^2\ ,$$

where $\rho - 2\Gamma_i$ and $\rho - 2\Lambda_i$ are assumed to be positive in order to have a convergent improper integral, and

$$\Gamma_i = \beta_i \left[D_i \frac{\beta_j}{g_j} - \frac{\beta_i}{g_i} \right] r^c + A_j(\beta_i D_i + \beta_j) + \alpha\ ,$$

$$\Lambda_i = -\beta_i^2 \frac{r^c}{g_i} + \beta_j A_j + \alpha\ .$$

Easy but tedious computations allow us to rewrite the credibility conditions in (23) as follows:

$$D_i \left(A_j + \frac{\beta_j}{g_j} r^c \right) \left\{ (\rho - 2\Lambda) \left[-2\beta_i r^c + g_i D_i \left(A_j + \frac{\beta_j}{g_j} r^c \right) \right] + 2\beta_i \left(\beta_i^2 \frac{(r^c)^2}{g_i} + m_i \right) \right\} \geq 0\ ,$$

$i, j = 1, 2, i \neq j$.

The linear-incentive strategies are credible for any deviation in (15) if both inequalities are satisfied simultaneously. An exhaustive study of the inequality corresponding to player i leads to the following result. The credibility condition for player i is fulfilled if and only if one of the following conditions is satisfied:

(1) $D_i > 0$

(a) $A_j \geq -r^c \frac{\beta_j}{g_j}$ and

$$C_{2i}^2 - 4C_{1i}C_{3i} < 0 \quad \text{and} \quad C_{3i} \geq 0$$

or

$$C_{2i}^2 - 4C_{1i}C_{3i} \geq 0 \quad \text{and} \quad \left\{ \begin{array}{c} A_j \geq \max\{A_j^+, A_j^-\} \\ \text{or} \\ A_j \leq \min\{A_j^+, A_j^-\} \end{array} \right. .$$

(b) $A_j \leq -r^c \frac{\beta_j}{g_j}$ and

$$C_{2i}^2 - 4C_{1i}C_{3i} < 0 \quad \text{and} \quad C_{3i} < 0$$

or

$$C_{2i}^2 - 4C_{1i}C_{3i} \geq 0 \quad \text{and} \quad A_j \in [\min\{A_j^+, A_j^-\}, \max\{A_j^+, A_j^-\}] .$$

(2) $D_i < 0$

(a) $A_j \geq -r^c \frac{\beta_j}{g_j}$ and

$$C_{2i}^2 - 4C_{1i}C_{3i} < 0 \quad \text{and} \quad C_{3i} < 0$$

or

$$C_{2i}^2 - 4C_{1i}C_{3i} \geq 0 \quad \text{and} \quad A_j \in [\min\{A_j^+, A_j^-\}, \max\{A_j^+, A_j^-\}] .$$

(b) $A_j \leq -r^c \frac{\beta_j}{g_j}$ and

$$C_{2i}^2 - 4C_{1i}C_{3i} < 0 \quad \text{and} \quad C_{3i} \geq 0$$

or

$$C_{2i}^2 - 4C_{1i}C_{3i} \geq 0 \quad \text{and} \quad \left\{ \begin{array}{c} A_j \geq \max\{A_j^+, A_j^-\} \\ \text{or} \\ A_j \leq \min\{A_j^+, A_j^-\} \end{array} \right. .$$

Constants $C_k, k = 1, \ldots, 3$ are given by:

$$C_{1i} = -2\beta_j g_i D_i ,$$
$$C_{2i} = g_i D_i \left(\rho + 2\beta_i^2 \frac{r^c}{g_i} - 2\alpha \right) - 2\beta_j r^c \left(-2\beta_i + g_i D_i \frac{\beta_j}{g_j} \right) ,$$
$$C_{3i} = r^c \left(\rho + 2\beta_i^2 \frac{r^c}{g_i} - 2\alpha \right) \left(-2\beta_i + g_i D_i \frac{\beta_j}{g_j} \right) + 2\beta_i \left(\beta_i^2 \frac{(r^c)^2}{g_i} + m_i \right) .$$

The lower and upper bounds for A_j are:

$$A_j^{(+,-)} = \frac{-C_{2i} \pm \sqrt{C_{2i}^2 - 4C_{1i}C_{3i}}}{2C_{1i}} ,$$

where A_j^+ is the expression where the square root is affected by a positive sign.

Similar conditions ensuring the credibility for player j can be derived analogously, changing subscript i by j. Merging the conditions for both players we obtain those which must be satisfied to guarantee that the linear-incentive strategies are credible.

Acknowledgements

Research started when the second author was visiting professor at Universidad de Valladolid, under grant SAB2004-0162. The first author's research was partially supported by MEC and JCYL under projects SEJ2005-03858/ECON and VA045A06, co-financed by FEDER funds. The second author's research is supported by NSERC, Canada.

References

[1] Aarle B., Bovenberg L., Raith M., *Monetary and Fiscal Policy Interaction and Government Debt Stabilization*, Journal of Economics **62** (1995), 111–140.

[2] Chiarella C., Kemp M.C., Van Long N., *On the Economics of International Fisheries*, International Economic Review **25** (1984), 85–92.

[3] Cohen D., Michel P., *How should Control Theory Be Used to Calculate a Time-Consistent Government Policy*, Review of Economic Studies **LV** (1988), 263–274.

[4] Dockner E.J., Long N.V., *International Pollution Control: Cooperative versus Noncooperative Strategies*, Journal of Environmental Economics and Management **25** (1993), 13–29.

[5] Dockner E., Jørgensen S., Long N.V., Sorger G., *Differential Games in Economics and Management Science*, Cambridge University Press, Cambridge, U.K., (2000).

[6] Ehtamo H., Hämäläinen R.P., *On Affine Incentives for Dynamic Decision Problems*, in T. Basar (ed.), *Dynamic Games and Applications in Economics*, Springer-Verlag, Berlin, (1986), 47–63.

[7] Ehtamo H., Hämäläinen R.P., *Incentive Strategies and Equilibria for Dynamic Games with Delayed Information*, Journal of Optimization Theory and Applications **63** (1989), 355–369.

[8] Ehtamo H., Hämäläinen R.P., *Cooperative Incentive Equilibrium for a Resource Management Problem*, Journal of Economic Dynamics and Control **17** (1993), 659–678.

[9] Engwerda J., *Linear-Quadratic Dynamic Optimization and Differential Games*, John Wiley, New York, (2005).

[10] Fernández L., *Trade's Dynamic Solutions to Transboundary Pollution*, Journal of Environmental Economics and Management **43** (2002), 386–411.

[11] Haurie A., *A Note on Nonzero-sum Differential Games with Bargaining Solution*, Journal of Optimization Theory and Applications **18** (1976), 31–39.

[12] Jørgensen S., Martín-Herrán G., Zaccour G., *Agreeability and Time-Consistency in Linear-State Differential Games*, Journal of Optimization Theory and Applications **119(1)** (2003), 49–63.

[13] Jørgensen S., Martín-Herrán G., Zaccour G., *Sustainability of Cooperation Overtime in Linear-Quadratic Differential Games*, International Game Theory Review **7(4)** (2005), 395–406.

[14] Jørgensen S., Zaccour G., *Time Consistent Side Payments in a Dynamic Game of Downstream Pollution*, Journal of Economic Dynamics and Control **25(2)** (2001), 1973–1987.

[15] Jørgensen S., Zaccour G., *Incentive Equilibrium Strategies and Welfare Allocation in a Dynamic Game of Pollution Control*, Automatica **37** (2001), 29–36.

[16] Jørgensen S., Zaccour G., *Time Consistency in Cooperative Differential Game*, in G. Zaccour (ed.), *Decision and Control in Management Science: In honor of Professor Alain Haurie*, Kluwer Academic Publishers, Boston, (2002), 349–366.

[17] Jørgensen S., Zaccour G., *Differential Games in Marketing*, Kluwer Academic Publishers, Boston, (2004).

[18] Jørgensen S., Zaccour G., *Developments in Differential Game Theory and Numerical Methods: Economic and Management Applications*, Computational Management Science, **4** (2007), 159–181.

[19] Kaitala V., Pohjola M., *Economic Development and Agreeable Redistribution in Capitalism: Efficient Game Equilibria in a Two-Class Neo-Classical Growth Model*, International Economic Review **32** (1990), 421–438.

[20] Kaitala V., Pohjola M., *Sustainable International Agreements on Green House Warming: A Game Theory Study*, Annals of the International Society of Dynamic Games **2** (1995), 67–88.

[21] List J.A., Mason C.F., *Optimal Institutional Arrangements for Transboundary Pollutants in a Second-Best World: Evidence from a Differential Game with Asymmetric Players*, Journal of Environmental Economics and Management **42** (2001), 277–296.

[22] Liu P.T., *Non-Zero Sum Differential Games with Bargaining Solution*, Journal of Optimization Theory and Applications **11(3)** (1973), 284–292.

[23] Martín-Herrán G., Rincón-Zapatero J.P., *Efficient Markov Perfect Nash Equilibria: Theory and Application to Dynamic Fishery Games*, Journal of Economic Dynamics and Control **29** (2005), 1073–1096.

[24] Martín-Herrán G., Zaccour G., *Credibility of Incentive Equilibrium Strategies in Linear-State Differential Games*, Journal of Optimization Theory and Applications **126(2)** (2005), 1–23.

[25] Neck R., Dockner E.J., *Can the Gains from International Cooperation be Secured without Policy Coordination*, Ricerche Economiche **3–4** (1987), 373–391.

[26] Petrosjan L.A., *Agreeable Solutions in Differential Games*, International Journal of Mathematics, Game Theory and Algebra **7(2/3)** (1997), 165–177.

[27] Petrosjan L., Zaccour G., *Time-Consistent Shapley Value of Pollution Cost Reduction*, Journal of Economic Dynamics and Control **27(3)** (2003), 381–398.

[28] Petrosjan L.A., Zenkevich N.A., *Game Theory*, World Scientific, Singapore, (1996).

[29] Rincón-Zapatero J.P., Martín-Herrán G., Martínez J., *Identification of Efficient Subgame-perfect Nash Equilibria in a Class of Differential Games*, Journal of Optimization Theory and Applications **104** (2000), 235–242.

[30] Tolwinski B., Haurie A., Leitmann G., *Cooperative Equilibria in Differential Games*, Journal of Mathematical Analysis and Applications **119** (1986), 182–202.

[31] Van der Ploeg F., De Zeeuw A.J., *International Aspects of Pollution Control*, Environmental and Resource Economics **2** (1992), 117–139.

[32] Yeung D.W.K., Petrosjan L., *Cooperative Stochastic Differential Games*, Springer, New York, (2005).

Robust Control Approach to Digital Option Pricing: Synthesis Approach

Stéphane Thiery
Laboratory I3S
University of Nice Sophia Antipolis and CNRS
Sophia Antipolis, France
thiery@i3s.unice.fr

Pierre Bernhard
Laboratory I3S
University of Nice Sophia Antipolis and CNRS
Sophia Antipolis, France
Pierre.Bernhard@polytech.unice.fr

Geert Jan Olsder
Delft Institute of Applied Mathematics
Delft University
P.O. Box 5031, 2600 GA Delft, The Netherlands
g.j.olsder@tudelft.nl

Abstract

In the framework of an interval market model already used in [6,7,5], we tackle the case of a digital option with its discontinuous terminal payment. In this paper we develop the synthesis approach via the investigation of the trajectory field à la Isaacs-Breakwell.

Key words. Differential games, Isaacs' equation, option pricing.

AMS Subject Classifications. Primary 91A23; Secondary 49L20, 91B28.

1 Introduction

In [3,4,6], we introduced a robust control approach to option pricing using a non stochastic model of the market uncertainties and a minimax approach. Several authors used simultaneously and independently similar approaches. The same market model has previously been used in [12][1], who coined the phrase "interval model", and also in [9], where the author stresses that due to the incompleteness of the market, the hedge is a super-replication, and the premium computed is a "seller's price". While these references use tools quite different from ours, a very similar theory (actually strictly more general, and consequently less detailed) has

[1] Available as a preprint as early as 2000.

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_15,

been developed in [11,1] and related papers. See also a further development of similar ideas in [13]. A rather exhaustive survey of our own theory can be found in [5]. A discussion of the strengths and weaknesses of this approach, as compared to the classical Black and Scholes theory, can be found in [6].

We tackle here the case of a digital call. This is a contract whereby the writer (also called here the *trader*) pledges to pay the buyer a fixed amount D at a prescribed final time (or *exercise time*) T if the market price of a specified underlying stock at this time is not less than a given price K (also called the *strike*). This defines a *payoff* of the option which is a discontinuous function of the final price of the underlying stock. This is to be contrasted with classical—or "vanilla"— options where the payoff is continuous and even convex.

In this paper, we develop the analysis in terms of the field of extremal trajectories, reconstructing the Value function via a classical synthesis approach à la Isaacs-Breakwell. This allows us to completely solve the problem and to exhibit a representation theorem similar to that of [8,5], with the same linear vector PDE, but with a different coefficient q^+, and different sets of initial conditions and gradient discontinuities.[2]

2 The model

We quickly recall the overall framework. The reader is referred to our previous works, i.e., [5], for a more detailed description.

Parameters The parameters of our problem are the exercise time T, strike K, amount[3] D, transaction cost rates C^+ and C^-, depending on whether it is a buy or a sale, and the similar c^+ and c^- that apply to the closure costs, with $C^- \leq c^- \leq 0 \leq c^+ \leq C^+$, and $C^-C^+ \neq 0$. We let $\xi > 0$ denote the amount of a buy of underlying stock, and $\xi < 0$ denote a sale of an amount $-\xi$. The transaction costs are $C^+\xi$ or $C^-\xi$, respectively, both positive. Moreover, we shall (realistically) assume that $1 + C^- > 0$.

We shall use the notation $C^\varepsilon\xi$ to mean $C^+\xi$ if $\xi > 0$ and $C^-\xi$ if $\xi < 0$. A similar convention will apply to such notations as $q^\varepsilon(\check{v} - v)$ where q^ε will mean q^+ or q^- (to be defined) depending on the sign ε of $(\check{v} - v)$.

Furthermore, two constant bounds $\tau^- < 0$ and $\tau^+ > 0$ on the relative stock price rate of change are known. See the market model below.

We assume that via the classical change of variables to end-time values, the riskless interest rate has been factored out.

State variables In these end-time values, the state variables are:

- the **u**underlying stock price u,

---

[2]A much more detailed and complete analysis is to appear in [14].

[3]Without loss of generality, this could have been taken as 1. Keeping it as D helps one keep track of the physical dimensions in the calculations. The meaningful dimensionless quantity is the ratio D/K.

- the **v**alue of the portfolio's underlying stock content v,
- the total **w**orth of the portfolio w.

The underlying stock price u is by essence positive. Moreover, since we consider a call (an increasing payment function), there is no point in considering negative v's either, so that we shall always assume that u and v are both non negative.

Market model We use the *interval model*. In that model, it is assumed that the stock price is an absolutely continuous function, and furthermore that two numbers, τ^+ and τ^-, with $\tau^- < 0 < \tau^+$, are known, such that:

$$\forall t_1, t_2\,, \quad \mathrm{e}^{\tau^-(t_2-t_1)} \leq \frac{u(t_2)}{u(t_1)} \leq \mathrm{e}^{\tau^+(t_2-t_1)}\,. \tag{1}$$

We call Ω the set of such admissible stock price trajectories. Alternatively, we shall let $\tau = \dot{u}/u$ and let $\tau(t)$ be a measurable function with for all t, $\tau^- \leq \tau(t) \leq \tau^+$. We shall call Ψ the set of all such admissible rate functions.

Trader controls and strategy The control of the trader is through buying or selling underlying stocks. He may either do so in a *continuous trading* mode, a classical fictitious mode whereby one buys stocks at a continuous rate $\xi(t)$, (a sale if $\xi < 0$,) or in an impulsive mode or *discrete trading*. In that later mode, the trader buys (or sells) lump sums at finitely many freely chosen time instants. We shall call these times t_k, $k = 1, 2, \ldots$ and the amount traded at these instants ξ_k, of signs ε_k. Hence, the function $\xi(\cdot)$ will be considered as a sum of a measurable *continuous component* and finitely many weighted translated Dirac impulses. We shall call Ξ the set of such admissible trader's controls.

The trader acts knowing the market situation (fictitiously), with no time delay. The mathematical metaphor of that hypothesis is as follows: a strategy is a non-anticipative function $\phi : \Omega \to \Xi$. (The initial portfolio content $v(0) = v_0$ will usually be considered as zero. Yet, to be more general, we must let ϕ also depend on it.) In practice, we shall implement it as a state feedback $\xi(t) = \varphi(t, u(t), v(t))$. We do not attempt to describe all admissible state feedbacks, being content to check that the one we exhibit actually yields an admissible non-anticipative strategy. We let Φ be the set of admissible non-anticipative strategies, and use the notation $\varphi \in \Phi$ to mean that a feedback strategy φ generates an admissible strategy ϕ in Φ.

Dynamics The dynamics are:

$$\dot{u} = \tau u\,, \tag{2}$$

$$\dot{v} = \tau v + \xi\,, \tag{3}$$

$$\dot{w} = \tau v - C^{\varepsilon}\xi\,, \tag{4}$$

and in case the trader decides to make a block buy or sale of stocks of magnitude ξ_k at time t_k,

$$v(t_k^+) = v(t_k) + \xi_k\,, \tag{5}$$

$$w(t_k^+) = w(t_k) - C^{\varepsilon}\xi_k\,. \tag{6}$$

Payoff At the terminal time T, the trader sells any remaining underlying stock $v(T)$ in its portfolio, at a closure cost $-c^- v(T)$, and pays its due D to the buyer if $u(T) \geq K$. Thus, its incurred cost is $N(u(T), v(T))$ with

$$N(u,v) = \begin{cases} -c^- v & \text{if } u < K\,, \\ D - c^- v & \text{if } u \geq K\,. \end{cases} \tag{7}$$

The total expense born by the trader due to possible losses (or gains) due to stock price variations, transaction costs, and terminal costs, if the contract has been written at time t_0 with $u(t_0) = u_0$ and $v(t_0) = v_0$ is therefore

$$J(t_0, u_0, v_0; \xi(\cdot), \tau(\cdot)) = N(u(T), v(T)) + \int_{t_0}^{T} (C^\varepsilon \xi(t) - \tau(t) v(t)) + \sum_k C^{\varepsilon_k} \xi_k\,. \tag{8}$$

Let

$$W(t_0, u_0, v_0) := \inf_{\phi \in \Phi} \sup_{\tau \in \Psi} J(t_0, u_0, v_0; \phi(u(\cdot)), \tau(\cdot)) \tag{9}$$

be the Value function of that differential game problem. The premium that the writer should charge for this contract, written at time 0, is $P(u(0)) = W(0, u(0), 0)$.

The limit case $u(T) = K$ We chose to let the terminal payment be u.s.c., deciding that the amount D is owed if $u(T) = K$. This is no serious restriction in terms of modelization. But it simplifies the analysis in that it lets the supremum in (9) be a maximum.

For a given pair (t,u), we shall call $\Psi_K(t,u)$ the subset of Ψ of controls $\tau(\cdot) : [t,T] \to [\tau^-, \tau^+]$ that drive the price from $u(t) = u$ to $u(T) = K$, (i.e, $\exp \int_t^T \tau(s)\,\mathrm{d}s = K/u$). Finally, we call $\Lambda \subset [0,T] \times \mathbb{R}_+$ the set of (t,u)'s for which $\Psi_K(t,u) \neq \emptyset$. This is equivalent to $u \in [u_\ell(t), u_r(t)]$, with

$$u_\ell(t) = K\mathrm{e}^{-\tau^+(T-t)}\,, \qquad u_r(t) = K\mathrm{e}^{-\tau^-(T-t)}\,. \tag{10}$$

Further notations We give here for ease of reference some (strange) notations that we shall use all along hereafter. We let

$$q^-(t) = \max\{(1+c^-)\mathrm{e}^{\tau^-(T-t)} - 1\,, C^-\}\,, \tag{11}$$

$$q^+(u) = \min\Big\{\max\{(1+c^-)\frac{K}{u} - 1\,, C^-\}\,, C^+\Big\}\,. \tag{12}$$

These two numbers will appear as the opposite of the partial derivative $\partial W/\partial v$, or loss coefficients associated to the fact of having too much, respectively, not enough, of the underlying stock in the portfolio as compared to an "ideal" content $\check{v}(t,u)$.

We let also

$$t_- = T - \tfrac{1}{\tau^-} \ln\left(\tfrac{1+C^-}{1+c^-}\right)\,, \quad u_- = K\tfrac{1+c^-}{1+C^-}\,, \tag{13}$$

$$t_+ = T - \tfrac{1}{\tau^+}\ln\left(\tfrac{1+C^+}{1+c^-}\right), \quad u_+ = K\,\tfrac{1+c^-}{1+C^+}\,. \tag{14}$$

t_-, u_-, and u_+ are the switch values in the definitions (11),12 of q^- and q^+:

$$q^-(t) = \begin{cases} C^- & \text{if } t \le t_-\,, \\ (1+c^-)\mathrm{e}^{\tau^-(T-t)} - 1\,, & \text{if } t \ge t_-\,, \end{cases}$$

and

$$q^+(u) = \begin{cases} C^+ & \text{if } u \le u_+\,, \\ (1+c^-)\frac{K}{u} - 1 & \text{if } u_+ \le u \le u_-\,, \\ C^- & \text{if } u \ge u_-\,, \end{cases}$$

and t_+ is related to u_+ as shown in Figs. 1 and 2.

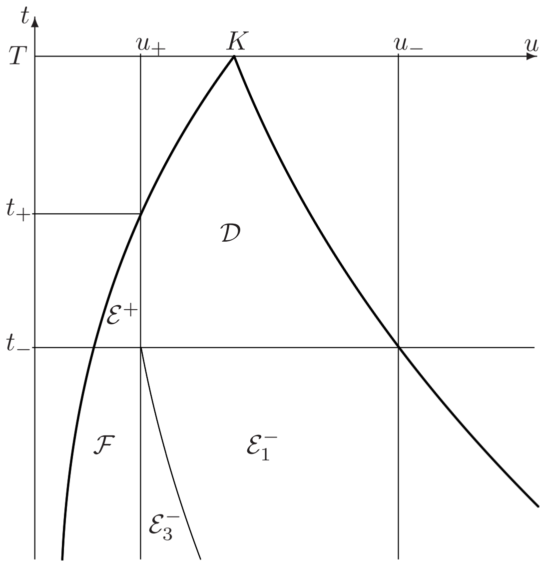

Figure 1: The various regions, $t_- < t_+$.

We shall also let

$$Q^\varepsilon = (q^\varepsilon \quad 1)\,,\ \varepsilon = \pm\,, \quad Q = \begin{pmatrix} Q^+ \\ Q^- \end{pmatrix}, \tag{15}$$

and, wherever $q^- \ne q^+$ (and hence Q invertible),

$$\mathcal{T} = \frac{1}{q^+ - q^-}\begin{pmatrix} \tau^+q^+ - \tau^-q^- & \tau^+ - \tau^- \\ -(\tau^+ - \tau^-)q^+q^- & \tau^-q^+ - \tau^+q^- \end{pmatrix} = Q^{-1}\begin{pmatrix} \tau^+ & 0 \\ 0 & \tau^- \end{pmatrix} Q\,. \tag{16}$$

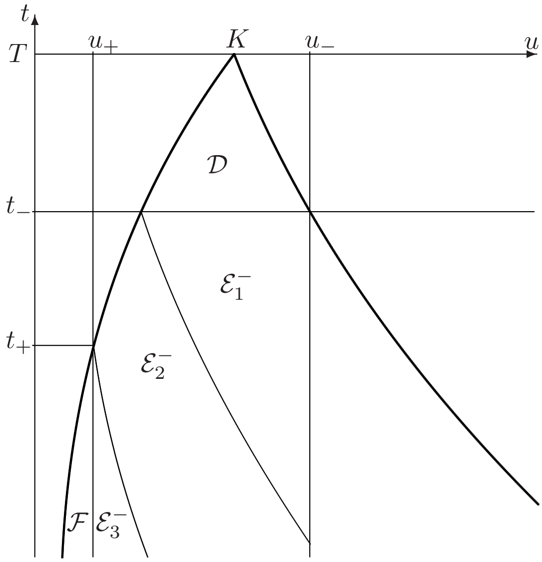

Figure 2: The various regions, $t_+ < t_-$.

This matrix appears in Theorem 3.1 below, and, as a consequence, in the representation Theorem 4.1. The second form in the equation above makes it, if not intuitive, at least logically connected to the context.

Finally, we set

$$\mathbb{1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathcal{S} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}. \tag{17}$$

$$\mathcal{V} = \begin{pmatrix} v \\ w \end{pmatrix}, \quad \check{\mathcal{V}} = \begin{pmatrix} \check{v} \\ \check{w} \end{pmatrix}. \tag{18}$$

The notations $\check{v}$ and $\check{w}$ will appear later.

3 Analysis

3.1 Isaacs equation

One of the issues in investigating that game is to decide how to cope with the impulse controls. One avenue is to consider the three-dimensional game of degree with state space (t, u, v) and payoff (8), and to use the differential quasi-variational inequality (DQVI) associated with that game, as done in [7]. The other avenue, that we follow here, and also used in [7], makes use of the 4-dimensional representation in (t, u, v, w), and finds the graph of W as the boundary of the *capturable states*. By this, we mean states that can be driven by the trader to the target set $\{w \geq N(u,v)\}$ at the terminal time T. In that framework, we use the tools of semipermeability to construct that boundary, and jump trajectories are just trajectories orthogonal to the (t,u) plane.

We shall use normals to the boundary of the capturable states of the general form

$$\nu = \begin{pmatrix} n \\ p \\ q \\ r \end{pmatrix}$$

with usually $r = 1$, i.e., an *inward* normal. The extremalizing controls will be given by Isaacs main equation. It can be written in terms of the Hamiltonian

$$H(t,u,v,w,n,p,q,\tau,\xi) = n + \tau[pu + (1+q)v] + (q - C^{\varepsilon})\xi$$

as

$$\sup_{\xi} \min_{\tau \in [\tau^-,\tau^+]} H(t,u,v,w,n,p,q,\tau,\xi) = 0\,. \tag{19}$$

The minimum in τ is always reached at $\tau = \tau^{\varepsilon}$, with $\varepsilon = \mathrm{sign}(\sigma)$ decided by the switch function $\sigma = -pu - (1+q)v$. Singular trajectories will involve $\sigma = 0$.

The supremum in ξ is reached at $\xi = 0$ if $C^- \leq q \leq C^+$. Otherwise, the supremum is $\pm\infty$, according to the sign of q. (This corresponds to a jump, of the same sign as q.)

It follows directly from the dynamics that jump trajectories lie in a (v,w) plane, and have a slope C^{ε}, ε the sign of the jump. Consequently, any hypersurface made up of such trajectories has a normal of the form $\nu^t = (n \;\; p \;\; C^{\varepsilon} \;\; 1)$. This yields a singular Hamiltonian in ξ.

The adjoint equations read

$$\begin{aligned} \dot{n} &= 0\,, \\ \dot{p} &= -\tau p\,, \\ \dot{q} &= -\tau(q+1)\,, \\ \dot{r} &= 0\,. \end{aligned}$$

Hence, we may set $r = 1$ if we manage to choose it such at terminal time.

3.2 Primary field

3.2.1 The sheets (τ_{ℓ}^-) and (τ_r^-)

Trajectories ending in the region $u < K$, $w + c^- v = 0$ must be with $p(T) = 0$, $q(T) = c^-$, yielding $\sigma(T) = -(1+c^-)v(T) < 0$. Therefore, the final τ should be τ^- and $\xi = 0$. The switch function is constant along such a trajectory. But these trajectories cannot be integrated beyond t_- (see (13)). As a matter of fact, the adjoint equations yield $q(t) = q^-(t)$ (see (11)), so that before t_-, we would have $q < C^-$, and the control $\xi = 0$ would no longer maximize the Hamiltonian.

This 2-D set of trajectories (parametrized by $u(T), v(T)$) creates a 3-D manifold, a semi-permeable hypersurface, that we call the *sheet* (τ_ℓ^-). This manifold obeys the equation

$$(\tau_\ell^-) : \quad q^- v + w = 0 \,,$$

and its normal ν_{τ^-} is

$$\nu^t_{\tau^-} = (-\tau^-(q^- + 1)v \quad 0 \quad q^- \quad 1) \tag{20}$$

or, using the notations (15) (18), as:

$$Q^-\mathcal{V} = 0 \,, \qquad \text{and} \qquad \nu^t_{\tau^-} = (-\tau^- Q^- \mathbb{1} v \quad 0 \quad q^- \quad 1) \,.$$

A similar construction holds for trajectories ending in the region $u > K$, resulting in a semi-permeable surface (τ_r^-): $Q^-\mathcal{V} = D$, with the same normal (20).

3.2.2 The singular sheet (K)

At $u = K$, the final payment N is non differentiable. The semi-permeable normal can be any element of the super-differential, hence of the form $(n \; p \; c^- \; 1)$, $p \le 0$. The final switch function is thus of the form $\sigma(T) = -[pK + (1 + c^-)v]$. Again, it is constant along a trajectory. If $\sigma < 0$, we get the trajectory $u_r(t)$ (see (10)) this is the "right" boundary of the sheet (τ_ℓ^-). If $\sigma > 0$, we get $u_\ell(t)$. This will be seen to be the "left" boundary of the sheet (K) in the region $t \ge t_+$.

If $p(T)$ is chosen to make $\sigma(T) = 0$, then sigma will remain zero along any trajectory (as long as $\xi = 0$), and hence any $\tau(t) \in [\tau^-, \tau^+]$ is permissible. Let $\theta := \int_t^T \tau(s)\mathrm{d}s$. We generate that way a new 3-D manifold parametrized by $(v(T), \theta, t)$.

It is a valid semi-permeable hypersurface as long as q remains between C^- and C^+. One easily sees that $q = (1 + c^-)K/u - 1$, so that the condition $q \le C^+$ translates into $u \ge u_+$ (see (14))—this is the left boundary of that sheet for $t \le t_+$—, and $q \ge C^-$ translates into $u \le u_-$—this is the right boundary of that sheet for $t \le t_-$. For $t \ge t_+$, the left boundary is the trajectory $u_\ell(t)$. For $t \ge t_-$, the right boundary is on $u = u_r(t)$. On this whole sheet, $q = q^+(u)$ (see (12)).

We call this semi-permeable hypersurface the *sheet* (K). It is also characterized by $Q^+\mathcal{V} = D$, and its normal is $\nu^t_K = (0 \quad -Q^+\mathbb{1}v/u \quad q^+ \quad 1)$.

3.2.3 Projection in the (t, u) space

It is useful to look at the projection of the different sheets in the (t, u) space. The domain of validity of each sheets is:

$$\begin{aligned}
(\tau_\ell^-) \quad &: \quad \{t \ge t_-\} \cap \{u < u_r(t)\} \,, \\
(\tau_r^-) \quad &: \quad \{t \ge t_-\} \cap \{u \ge u_r(t)\} \,, \\
(K) \quad &: \quad \max\{u_+, u_l(t)\} \le u \le \min\{u_-, u_r(t)\} \,.
\end{aligned}$$

This shows that (K) and (τ_l^-) coexist in the domain

$$\{t \geq t_-\} \cap \{\max\{u_+, u_l(t)\} \leq u \leq u_r(t)\} \ ,$$

and that (K) and (τ_r^-) exist in disjoint (t,u) domains, except along their common trajectory $u = u_r(t)$ for $t \geq t_-$ where they join smoothly since they share $q = (1+c^-)K/u_r - 1 = q^-$.

3.2.4 The dispersal manifold $\mathcal{D} = (\tau_\ell^-) \cap (K)$

In the region $t \geq t_-$, $\max\{u_+, u_\ell(t)\} \leq u < u_r(t)$, capturable states are characterized by the two conditions $w \geq -q^- v$ and $w \geq D - q^+ v$. The boundary of capturable states (the graph of W) is made of the two sheets (τ_ℓ^-) and (K), between the end time and their intersection $\mathcal{D}$ characterized by $v = \check{v}(t,u)$, $w = \check{w}(t,u)$ with

$$\check{v}(t,u) = \frac{D}{q^+(u) - q^-(t)} \ , \qquad \check{w}(t,u) = -q^-(t)\check{v}(t,u) \, . \tag{21}$$

This is a dispersal manifold $\mathcal{D}$, somewhat degenerate in that, on the one hand, one of the outgoing fields, the sheet (K), is traversed by trajectories generated by any control $\tau(\cdot)$, and, on the other hand, the trajectories of the other "outgoing" field, the sheet (τ_ℓ^-), actually traverse the dispersal manifold $\mathcal{D}$ itself.

This manifold, thus, is born by the sheet (K), and traversed by trajectories generated by $\tau = \tau^-$. We shall show hereafter the following fact.

Theorem 3.1. *If a manifold $\mathcal{V} = \check{\mathcal{V}}(t,u)$ is either*

(1) traversed by trajectories τ^- and lying on the sheet (K),
(2) traversed by trajectories τ^+ and lying on a sheet (τ^-),
(3) traversed by trajectories τ^+ and trajectories τ^-,

it satisfies the partial differential equation

$$\check{\mathcal{V}}_t + \mathcal{T}(\check{\mathcal{V}}_u u - \mathcal{S}\check{\mathcal{V}}) = 0 \, . \tag{22}$$

In the present case, as for all other closed formulas obtained thereafter, the PDE (22) can be checked by direct differentiation.

This PDE was first introduced in [7] for the Focal manifold. In [8], we showed that it is satisfied by all the singular surfaces of that game, this coming as a surprise. In [6], we showed that it is a necessary consequence of (23) if this formula is to be that of a viscosity solution of the Isaacs DIQV. The present paper will prove the theorem above via the investigation of the field of optimal trajectories, thus explaining it, perhaps, in a more natural way.

The Value function in that region is therefore

$$W(t,u,v) = \check{w}(t,u) + q^\varepsilon(t,u)(\check{v}(t,u) - v) \, , \tag{23}$$

where we recall that in such an expression, $\varepsilon = \mathrm{sign}(\check{v}(t,u) - v)$.

3.2.5 Trivial regions

Outside of Λ If $u(t) < u_\ell(t)$ or $u(t) \geq u_r(t)$, the terminal $u(T)$ is less than K or, respectively, larger or equal to K irrespectively of what the players (market and trader) do. Hence, the final payment to the buyer is certain, the option is actually without any merit. The only hedge is $v = 0$, and $w = 0$ if $u(t) < u_\ell(t)$, or $w = D$ if $u(t) \geq u_r(t)$. If at a given time t, $v(t)$ happens to be positive, then the trader should sell it, either immediately at a cost $-C^- v$ if $t \leq t_-$, or at terminal time at a cost $-q^- v$ at worst if $t \geq t_-$. Hence, it always incurs a cost $-q^- v$ (see (11)), and thus the Value function is

$$W(t,u,v) = -q^-(t)v\,, \quad \text{or} \quad W(t,u,v) = D - q^-(t)v\,,$$

depending on whether u lies to the left of Λ or to its right. For $t \geq t_-$, these regions are each covered by a single sheet (τ^-), which is the graph of the function W. Formally, this can be written as (23) with $\check{v} = 0$ and $\check{w} = 0$ or D. One may notice that this $\check{\mathcal{V}}$ still trivially satisfies (22), or at least $\check{\mathcal{V}}_t = 0 = \check{\mathcal{V}}_u u - \mathcal{S}\check{\mathcal{V}}$.

Region $t \leq t_-$, $u \geq u_-$ Again, that region needs no big theory. There the loss resulting from keeping a positive v if $u(T) = K$ is larger than $C^- v$. Hence, the only sensible strategy is to sell any v at once. The value is $W(t,u,v) = D - C^- v$, as already seen for larger u's. The representation (23) can be preserved as in the previous paragraph.

3.3 Equivocal manifolds

To further analyze that game, we need to distinguish whether $t_- < t_+$ or $t_- > t_+$. We choose here to show the detailed analysis for the case $t_- < t_+$ because it displays a richer set of singularities, although it is the less likely in a real life application. We shall only sketch the (more realistic) case $t_- > t_+$, stressing the main difference. (See Figs. 1 and 2 at the end.)

3.3.1 The equivocal manifold $\mathcal{E}^+ = (\tau_\ell^-) \cap (\uparrow)$

In the region $t \in [t_-, t_+]$, $u \leq u_+$, the sheet (τ_ℓ^-) still exists, but not the sheet (K). We have reached (backward) $q = C^+$, therefore we may expect a positive jump manifold $(\uparrow)$. Such a 3-D manifold must join onto the sheet (τ_ℓ^-) along a 2-D junction manifold that we call $\mathcal{E}^+$.

We again call $\check{v}(t,u)$, $\check{w}(t,u)$ the values of v and w on $\mathcal{E}^+$. This way, the boundary of the capturable states, and hence the graph of W, will still be described by (23), but now $q^+ = C^+$.

Staying on (τ^-) for a $\tau \neq \tau^-$ requires that $\mathrm{d}(Q^- \mathcal{V})/\mathrm{d}t = 0$, hence that (assuming $\xi \geq 0$)

$$-\tau^-(1+q^-)v + q^-(\tau v + \xi) + \tau v - C^+ \xi = 0\,,$$

that yields

$$\xi = (\tau - \tau^-)\frac{1+q^-}{C^+ - q^-}\check{v} \tag{24}$$

which is actually non-negative. We conjecture (and will check later) that the junction will actually be with $\tau = \tau^+$, and let $\xi = \xi^+$ be the corresponding control.

The requirement that $\mathcal{E}^+$ be traversed by trajectories τ^+ implies that the dynamics are satisfied with τ^+, hence that

$$\check{v}_t + \check{v}_u \tau^+ u = \tau^+ \check{v} + \xi^+ ,$$
$$\check{w}_t + \check{w}_u \tau^+ u = \tau^+ \check{v} - C^+ \xi^+ .$$

Multiplying the first equation by C^+ and summing, we get:

$$C^+\check{v}_t + \check{w}_t + \tau^+[(C^+\check{v}_u + \check{w}_u)u - (C^+ + 1)\check{v}] = 0 .$$

We now notice that it follows from (16), that for $\varepsilon = \pm$, it holds that $Q^\varepsilon \mathcal{T} = \tau^\varepsilon Q^\varepsilon$. Thus this equation can also be written

$$Q^+[\check{\mathcal{V}}_t + \mathcal{T}(\check{\mathcal{V}}_u u - \mathcal{S}\check{\mathcal{V}})] = 0 . \tag{25}$$

Now, $\mathcal{E}^+$ admits the two tangent vectors

$$\begin{pmatrix} 1 \\ 0 \\ \check{v}_t \\ \check{w}_t \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \\ \check{v}_u \\ \check{w}_u \end{pmatrix} .$$

They must be orthogonal to ν_{τ^-} as given by (20). This yields two equations:

$$-\tau^-(q^- + 1)\check{v} + q^-\check{v}_t + \check{w}_t = Q^-\check{\mathcal{V}}_t - \tau^- Q^- \mathcal{S}\check{\mathcal{V}} = 0 , \tag{26}$$
$$q^-\check{v}_u + \check{w}_u = Q^-\check{\mathcal{V}}_u = 0 .$$

Multiplying the second one by $\tau^- u$ and summing and using again $\tau^- Q^- = Q^- \mathcal{T}$, we get:

$$Q^-[\check{\mathcal{V}}_t + \mathcal{T}(\check{\mathcal{V}}_u u - \mathcal{S}\check{\mathcal{V}})] = 0 . \tag{27}$$

The two Eqs. (25) and (27) together can be written

$$Q[\check{\mathcal{V}}_t + \mathcal{T}(\check{\mathcal{V}}_u u - \mathcal{S}\check{\mathcal{V}})] = 0 ,$$

and as Q is invertible, we get (22) for $\mathcal{E}^+$. *This proves assertion 2 in Theorem 3.1.*

We expect this manifold to join continuously with $\mathcal{D}$ on the boundary $u = u_+$, and as this is not a characteristic curve of that equation, it specifies uniquely its solution. But we shall also find a more explicit construction via the trajectories.

We still have to check that the junction "does not leak", or that for a $\tau < \tau^+$, the trajectory remaining on (τ^-) drifts towards the capturable states, or equivalently that $\tau = \tau^+$ minimizes the relevant Hamiltonian.

For an intermediate value of ξ to be optimal, we need that at the junction, $q = C^+$ on the incoming trajectory. (We have seen that anyhow, the normal to a positive jump manifold has to have $q = C^+$.) Hence, this will be an equivocal junction, in the parlance of Isaacs. Let $\nu^+ = (n \quad p \quad C^+ \quad 1)^t$ be the normal to the (positive) jump manifold on $\mathcal{E}^+$. Isaacs main equation therefore reads

$$\max_\xi \min_\tau \{n + \tau[pu + (1 + C^+)\check{v}]\} = 0\,,$$

and $\tau = \tau^+$ is indeed the minimizing τ if $[pu + (1 + C^+)\check{v}]$ as computed along our trajectories τ^+ is non-positive, or equivalently, $n \geq 0$, since the Hamiltonian remains zero by construction.

We shall derive ν^+ from the theory of the generalized adjoint equations developed in [2]. Let us first investigate the conditions at the boundary $u = u_+$. There $\check{v}$ and $\check{w}$ are given by (21). Let s be the time at which a trajectory reaches that boundary. The boundary of $\mathcal{D}$ and its tangent are obtained as follows:

$$\mathcal{D} \cap \{u = u_+\} : \begin{pmatrix} s \\ u_+ \\ \frac{D}{C^+ - q^-(s)} \\ -q^- \frac{D}{C^+ - q^-(s)} \end{pmatrix}, \quad \text{tangent:} \begin{pmatrix} 1 \\ 0 \\ -\frac{D\tau^-(1+q^-)}{(C^+ - q^-)^2} \\ C^+ \frac{D\tau^-(1+q^-)}{(C^+ - q^-)^2} \end{pmatrix}. \tag{28}$$

Since that manifold is to be contained in $\mathcal{E}$, its tangent must be orthogonal to ν^+. This simply yields $n = 0$, which is therefore a terminal condition for the generalized adjoint equation. This equation reads, writing x for (t, u, v, w)

$$\dot{\nu}^+ = -\frac{\partial H}{\partial x} + \alpha(t)(\nu^+ - \nu_{\tau^-}),$$

where $\alpha(t)$ must be chosen so as to maintain the singularity while integrating backward, i.e., here, $q = C^+$, hence $\dot{q} = 0$. This gives

$$\begin{aligned} \dot{n} &= \alpha(n + \tau^-(1 + q^-)\check{v})\,, \\ \dot{p} &= (-\tau^+ + \alpha)p\,, \\ \dot{q} &= -\tau^+(1 + C^+) + \alpha(C^+ - q^-)\,, \\ \dot{r} &= 0\,. \end{aligned}$$

The requirement that $\dot{q} = 0$ gives $\alpha = \tau^+(1 + C^+)/(C^+ - q^-) > 0$. And the differential equation for n gives $n \geq 0$, since for $n = 0$, $\dot{n} < 0$ and we integrate *backward* from $n = 0$. This provides the sign information we needed.

A final remark, useful in checking that this is a viscosity solution (to appear in a forthcoming paper) is that Eq. (26) shows that $Q^-\check{\mathcal{V}}_t < 0$, and that, writing that the two tangents to $\mathcal{E}^+$ are orthogonal to ν^+, we get that $Q^+\check{\mathcal{V}}_t = -n \leq 0$.

Placing (24) in the dynamics, and using the initial conditions (28), the equations for the trajectories can be integrated in closed form with $\tau = \tau^+$, leading to the closed form formulas:

$$\begin{array}{rcl} \tilde{u}(t,s) &=& u_+ e^{\tau^+(t-s)} \quad \text{hence} \quad s = t - \frac{1}{\tau^+}\ln\left(\frac{u}{u_+}\right) \\ \check{v}(t,\tilde{u}(t,s)) = \tilde{v}(t,s) &=& \frac{D}{u_+(C^+ - q^-(s))}\left(\frac{C^+ - q^-(s)}{C^+ - q^-(t)}\right)^{\frac{\tau^+ - \tau^-}{-\tau^-}} \tilde{u}(t,s) \\ \check{w}(t,\tilde{u}(t,s)) = \tilde{w}(t,s) &=& -q^-(t)\tilde{v}(t,s) \end{array} .$$

3.3.2 The equivocal manifold $\mathcal{E}^- = (\downarrow) \cap (K)$

We now investigate the region $t \le t_-$, $u \in [u_+, u_-]$. The situation is somewhat symmetrical to that of the preceding paragraph, as in that region the sheet (K) exists while the sheet (τ_ℓ^-) does not. We therefore expect that a negative jump manifold $(\downarrow)$ joins onto the sheet (K). We conduct a similar analysis of the junction $\mathcal{E}^-$, still calling $\check{v}(t,u)$, $\check{w}(t,u)$ the equations of that 2-D manifold, and formula (23) will still hold, but now with $q^- = C^-$, and $q^+ = (1+c^-)K/u - 1$.

Trajectories staying on (K) must satisfy $\mathrm{d}(Q^+\mathcal{V})/\mathrm{d}t = 0$. Hence (assuming $\xi \le 0$), by differentiation $(q^+ - C^-)\xi = 0$, thus $\xi = 0$.

We conjecture (and shall check later on) that the junction is with $\tau = \tau^-$. The fact that $\mathcal{E}^-$ be traversed by trajectories τ^- now yields

$$\check{v}_t + \check{v}_u \tau^- u = \tau^- v \,,$$
$$\check{w}_t + \check{w}_u \tau^- u = \tau^- v \,.$$

Proceeding as in the previous case, we infer that $Q^-[\check{\mathcal{V}}_t + \mathcal{T}(\check{\mathcal{V}}_u u - \mathcal{S}\check{\mathcal{V}})] = 0$.

Writing that the two natural tangent vectors to $\mathcal{E}^-$ are orthogonal to the normal $\nu_K^t = (0 \quad -(1+q^+)v/u \quad q^+ \quad 1)$ yields the two equations $Q^+\check{\mathcal{V}}_t = 0$ and $Q^+(\check{\mathcal{V}}_u - \mathbb{1}v/u) = 0$. Multiplying the second one by $\tau^+ u$ and adding, we get $Q^+[\check{\mathcal{V}}_t + \mathcal{T}(\check{\mathcal{V}}_u u - \mathcal{S}\check{\mathcal{V}})] = 0$. This together with the previous similar equation again yields (22). *This proves assertion 1 of Theorem 3.1.*

We expect this manifold to join continuously with $\mathcal{D}$ at $t = t_-$. This initial condition can be seen to uniquely specify $\mathcal{E}^-$ in the region $u \ge u_+ \mathrm{e}^{-\tau^-(t_- - t)}$. We call $\mathcal{E}_1^-$ the manifold thus generated. Indeed, integrating with $\tau = \tau^-$ and $\xi = 0$ yields the same formulas for $\check{v}$, $\check{w}$ as in $\mathcal{D}$, but with q^- replaced by $\tilde{q}^- := [(1+c^-)\exp(\tau^-(T-t)) - 1]$. This formula *is* indeed that of q^- in the region $t \ge t_-$, while here, $q^- = C^-$.

A direct calculation then shows that $Q^-\check{\mathcal{V}}_t < 0$. Writing that the tangents to $\mathcal{E}^-$ are orthogonal to the normal ν^- to the jump manifold again yields $n + Q^-\check{\mathcal{V}}_t = 0$, showing that our construction does give $n > 0$ which shows that $\tau = \tau^-$ indeed minimizes the Hamiltonian. (The corner does not leak.)

In the region $u \le u_+ \exp(-\tau^-(t_- - t))$, the trajectories of $\mathcal{E}^-$, if they are still generated by $\tau = \tau^-$ as we shall show, end up on the line $u = u_+$. Therefore, the terminal conditions to integrate them retrogressively will be provided by the analysis of the region $u \le u_+$ of the next subsection. For that reason, we do not have a closed form for the trajectories of that part, $\mathcal{E}_3^-$, of the equivocal manifold. We check that indeed $\tau = \tau^-$ will be minimizing the Hamiltonian.

Use again the theory of generalized adjoint equations. It yields here $\dot{n} = \alpha n$ so that n cannot change sign along a trajectory. And, as it has to be positive at the boundary, it will stay so, proving that τ^- is indeed minimizing in the Hamiltonian. (It can easily be seen that $\alpha = -\tau^-(1+C^-)/(q^+ - C^-) > 0$.)

Again, we remark that on the whole equivocal junction, we have $Q^\varepsilon \check{\mathcal{V}}_t \leq 0$, $\varepsilon = \pm$, a property needed in the viscosity solution analysis.

3.4 Focal manifold $\mathcal{F} = (\downarrow) \cap (\uparrow)$

We now investigate the only region left: $t \leq t_-$, $u \leq u_+$. There, neither sheet (τ^-) nor (K) exist to construct an equivocal junction on. We will have two jump manifolds, one of each sign, joining on a 2-D focal manifold $\mathcal{F}$. The theory of a similar manifold in the case of a vanilla option was introduced in [7]. A more general theory of higher-dimensional focal manifolds was developed in [10]. Here, we may just notice that this manifold has to be traversed by both τ^- and τ^+ trajectories. According to the analysis provided for $\mathcal{E}^+$ and $\mathcal{E}^-$ above, this shows that necessarily $Q^\varepsilon[\check{\mathcal{V}}_t + \mathcal{T}(\check{\mathcal{V}}_u u - \mathcal{S}\check{\mathcal{V}})] = 0$, for both $\varepsilon = \pm$, hence resulting in (22). *This proves assertion 3 of Theorem 3.1.*

We need to provide (22) with boundary values to uniquely specify $\mathcal{F}$. We notice that $\mathcal{E}_3^-$ satisfies the same set of coupled PDE's. We may therefore consider that we have a single set of PDE's to solve in the domain $t \leq t_-$, $\in [u_\ell(t), u_+ \exp(-\tau^-(t_- - t))]$. Notice that the coefficients of this linear vector PDE are continuous. We showed in [10] that the trajectories τ^- and τ^+ are its characteristic curves (or, more classically, the characteristic curves of an equivalent scalar second-order PDE). This is a Goursat problem. The solution will be specified if we give consistent boundary conditions on two such curves. Concerning the right-hand boundary, we have found $\mathcal{E}_1^-$ (in closed form) on its right, and by continuity this provides our boundary condition. It remains to find boundary conditions on the trajectory $u_\ell(t)$, which is our left-hand boundary.

On u_ℓ, if at any time, $\tau < \tau^+$, the state drifts outside of Λ, and the optimal strategy is to sell v at once at a cost $-C^- v$, and do nothing ($\xi = 0$) thereafter. For this strategy to drive the state to the admissible end states (provide a hedge), it is necessary that $Q^-\mathcal{V} \geq 0$. This must be maintained along the trajectory u_ℓ (which has $\tau = \tau^+$), and still allow the state to reach $\mathcal{E}^+$ at $t = t_-$. The limit trajectory thus satisfies $Q^-\mathcal{V} = 0$. Differentiating with respect to time, this yields $\xi = \tau^+ \check{v}(1+C^-)/(C^+ - C^-)$, and we can integrate that limit trajectory backward from the boundary of $\mathcal{E}^+$. (This is the same rule as in $\mathcal{E}^+$, thus everything is smooth.)

There remains to check the signs of the time components of the normals to the jump manifolds, as we know that their being non-negative insures that the controls τ used in the construction of the manifold do minimize the relevant Hamiltonian.

Let $\nu^\varepsilon = (n^\varepsilon \; p^\varepsilon \; C^\varepsilon \; 1)^t$ be the normal to the jump manifold where ε is the sign of the jump. We shall denote $\bar\varepsilon$ the opposite sign to ε. The components in time and v of the generalized adjoint equations applied to $\mathcal{F}$ give

$$\dot n^\varepsilon = \alpha^\varepsilon (n^\varepsilon - n^{\bar\varepsilon})\,,$$
$$0 = -\tau^\varepsilon(1 + C^\varepsilon) + \alpha^\varepsilon(C^\varepsilon - C^{\bar\varepsilon})\,,$$

thus

$$\alpha^\varepsilon = \varepsilon\tau^\varepsilon \frac{1 + C^\varepsilon}{C^+ - C^-} > 0\,.$$

The differential equations for n^+ and n^- are to be integrated backward (not on the same trajectories, though), so that by a standard inward field argument, if $n^{\bar\varepsilon}$ is positive, so remains n^ε.

We now check the initial conditions for these backward integrations. We have stressed that $\check{\mathcal{V}}_t$ and $\check{\mathcal{V}}_u$ will be continuous over the whole region of integration of our PDE, and noticeably at $u = u_+$. On $\mathcal{E}^-$, we have seen that necessarily $Q^+\check{\mathcal{V}}_t = 0$. The normal ν^+ has to be orthogonal the tangent vector to $\mathcal{F}$ (and $\mathcal{E}^-$), which yields $n^+ + Q^+\check{\mathcal{V}}_t = 0$, hence $n^+ = 0$ on $u = u_+$. Integrating backwards will indeed provide $n > 0$ in $\mathcal{F}$ as long as $n^- > 0$.

On the left boundary, we have both $H = 0$ on the incoming negative jump manifold, and that its normal ν^- is orthogonal to the trajectory we specified. This yields, respectively,

$$n^- + \tau^-[p^- u_\ell + (1 + C^-)\check v] = 0\,,$$
$$n^- + \tau^+ p^- u_\ell + \tau^-(1 + C^-)\check v = 0\,,$$

so that we get that $p^- = 0$ and $n^- = -\tau^-(1 + C^-)\check v > 0$.

On the "top" boundary, at $t = t_-$, $u \in [u_\ell(t_-), u_+]$, the jump manifolds join continuously on the jump manifold towards $\mathcal{E}^+$ for the positive jump, to (τ^-) for the negative one. All components of the normal other than the n time-component are therefore continuous, and the requirement that the Hamiltonian be zero provides the continuity of the first component.

Thus, both n^- and n^+ are positive everywhere.

It is a simple matter to recover the controls ξ^ε on $\mathcal{F}$ according to the trajectory τ^ε considered:

$$\xi^\varepsilon = \frac{\bar\varepsilon 1}{C^+ - C^-}[Q^{\bar\varepsilon}\check{\mathcal{V}}_t + \tau^\varepsilon Q^{\bar\varepsilon}(\check{\mathcal{V}}_u u - \mathcal{S}\check{\mathcal{V}})] = \frac{\bar\varepsilon 1}{C^+ - C^-}[\frac{\tau^{\bar\varepsilon} - \tau^\varepsilon}{\tau^{\bar\varepsilon}} Q^{\bar\varepsilon}\check{\mathcal{V}}_t].$$

The square bracket is nonpositive because, as we have seen, $Q^{\bar\varepsilon}\check{\mathcal{V}}_t \le 0$, so that ξ^ε indeed has the sign of ε as it should. There does not seem to be closed form formulas for that manifold.

3.5 Case $t_+ < t_-$

We only sketch some features of that case.

The dispersal manifold is similar to the previous one, holding for $t \geq t_-$, $u \in [u_\ell, \min\{u_-, u_r(t)\}]$. For $t \leq t_-$, a negative jump manifold joins on the sheet (K) with an equivocal junction, involving $\tau = \tau^-$ and the "singular" control $\xi = 0$. In the region "above" the trajectory τ^- through $(t_-, u_\ell(t_-))$, i.e., $u \geq u_\ell(t_-)\exp[-\tau^-(t-t_-)]$, it is completely similar to the junction $\mathcal{E}_1^-$ of the previous case, its being given by the same formulas as $\mathcal{D}$ but with $\tilde{q}^-$ instead of q^-.

The region accounted for by the equivocal manifold $\mathcal{E}^+$ is empty. But a new case arises "below" the separating trajectory τ^-, as junction trajectories, still built with $\tau = \tau^-$, $\xi = 0$, reach the left boundary of Λ, i.e., $u = u_\ell(t)$ before time t_-. We need therefore a boundary condition for $\mathcal{E}^-$ in that region, generating a part $\mathcal{E}_2^-$ of the equivocal manifold that comes between $\mathcal{E}_1^-$ and $\mathcal{E}_3^-$.

On the boundary u_ℓ, two conditions must hold. On the one hand, we must insure that $w(T) \geq 0$ even if $\tau = \tau^-$ up to time T. Since we are at $t < t_-$, the trader should sell its stocks as soon as the state drifts off Λ, hence making a negative jump in v at a cost $-C^-v$. We must therefore have $w + C^-v \geq 0$. On the other hand, we want to be on the sheet (K). The limiting states are thus in $\{Q^-\mathcal{V} = 0\} \cap \{Q^+\mathcal{V} = D\}$, i.e., given by formulas similar to that of $\mathcal{D}$, but this time with $q^- = C^- : u = u_\ell(s)$, $s \in [t_-, t_+]$,

$$\check{v}(s,u) = \frac{D}{q^+(u) - C^-}\,, \quad \check{w}(s,u) = -C^-\check{v}(s,u)\,.$$

This is indeed a τ dispersal manifold. For $\tau = \tau^+$, the trader responds with $\xi = 0$, the state leaves the above manifold on the sheet (K), and we check that indeed it remains above the sheet $Q^-\mathcal{V} = 0$ since $\mathrm{d}(Q^-\mathcal{V})/\mathrm{d}t = (1 + C^-)\tau^+v > 0$.

From this 1-D manifold of "terminal" conditions, we integrate backwards the 2-D equivocal junction with $\tau = \tau^-$, $\xi = 0$. This whole construction can be explicitly performed with closed form formulas, allowing one to check the no-leakage condition (that τ^- actually is the minimizing control in the Hamiltonian):

$$\begin{array}{rcl}\tilde{u}(t,s) & = & u_l(s)e^{\tau^-(t-s)} \Longrightarrow s = \frac{\tau^+T - \tau^-t + \ln(\frac{u}{K})}{\tau^+ - \tau^-} \\ \check{v}(t, \tilde{u}(t,s)) = \tilde{v}(t,s) & = & \frac{D}{q^+(\tilde{u}(t,s)) - \tilde{q}^-(t,s)} \\ \check{w}(t, \tilde{u}(t,s)) = \tilde{w}(t,s) & = & -\tilde{q}^-(t,s)\tilde{v}(t,s)\end{array}$$

with $\tilde{q}^-(t,s) = (1 + C^-)e^{\tau^-(s-t)} - 1 < C^-$ since $\tau^-(s-t) < 0$.

The other regions bear a close resemblance to the previous case.

4 Conclusion

We have a complete description of the field of optimal trajectories, via a trajectory-wise description of the singular surfaces. The required sign checks are provided by the generalized adjoint equations. It proves the following representation theorem.

Theorem 4.1. *The value function is everywhere given by* (23)*, or equivalently:*

$$W(t,u,v) = Q^{\varepsilon}\check{\mathcal{V}} - q^{\varepsilon}v\,, \quad \varepsilon = \mathrm{sign}(\check{v}(t,u) - v)\,,$$

where $\check{\mathcal{V}}$ satisfies the pair of coupled linear PDEs (22)*, —or $\check{\mathcal{V}}_t = 0 = \check{\mathcal{V}}_u u - \mathcal{S}\check{\mathcal{V}}$ in the region where $\mathcal{T}$ is not defined because $q^+ = q^-$— with appropriate boundary values as discussed above.*

Concerning $\mathcal{F}$ and $\mathcal{E}_3^-$, we know of no other way to actually compute them than integrating these PDEs.

In a forthcoming paper, we shall show through a detailed analysis of the discontinuities of the Value function and of its gradient that it actually is a viscosity solution of the corresponding Isaacs quasi-variational inequality.

References

[1] J.P. Aubin, D. Pujal, and P. Saint-Pierre. Dynamic management of portfolios with transaction costs under tychastic uncertainty. In M. Breton and H. Ben-Ameur, editors, *Numerical Methods in Finance*, pp 59–89. Springer, New York, 2005.

[2] P. Bernhard. Singular surfaces in differential games, an introduction. In P. Haggerdon, G.J. Olsder, and H. Knoboloch, editors, *Differential games and Applications*, Vol 3 of *Lecture Notes in Information and Control Sciences*, pp 1–33. Springer, Berlin, 1977.

[3] P. Bernhard. Une approche déterministe de l'évaluation d'options. In José-Luis Menaldi, Edmundo Rofman, and Agnès Sulem, editors, *Optimal Control and Partial Differential Equations*, Vol in honor of Professor Alain Bensoussan's 60th birthday, pp 511–520. IOS Press, 2001.

[4] P. Bernhard. A robust control approach to option pricing. In M. Salmon, editor, *Applications of Robust Decision Theory and Ambiguity in Finance*. City University Press, London, 2003.

[5] P. Bernhard. The robust control approach to option pricing and interval models: an overview. In M. Breton and H. Ben-Ameur, editors, *Numerical Methods in Finance*, pp 91–108. Springer, New York, 2005.

[6] P. Bernhard. A robust control approach to option pricing including transaction costs. In A.S Novak and K. Szajowski, editors, *Advances in Dynamic Games*, Vol 7 of *Annals of the ISDG*, pp 391–416. Birkäuser, Boston, 2005.

[7] P. Bernhard, N. El Farouq, and S. Thiery. An impulsive differential game arising in finance with interesting singularities. In A. Haurie, S. Muto, L.A. Petrosjan, and T.E.S Raghavan, editors, *Advances in Dynamic Games* , Vol 8 of *Annals of the ISDG*, pp 335–363. Springer, New York, 2006. (Also in

10th ISDG International Symposium on Dynamic Games and Applications, Saint-Petersburg, 2002).

[8] P. Bernhard, N. El Farouq, and S. Thiery. Robust control approach to option pricing: a representation theorem and fast algorithm, *SIAM Journal on Control and Optimization*, Vol 46(6), pp 2280–2302, 2007.

[9] V. Kolokoltsov. Nonexpansive maps and option pricing theory. *Kybernetica*, Vol 34(6), pp 713–724, 1998.

[10] A. Melikyan and P. Bernhard. Geometry of optimal trajectories around a focal singular surface in differential games. *Applied Mathematics and Optimization*, Vol 52, pp 22–37, 2005.

[11] D. Pujal. *Evaluation et gestion dynamiques de portefeuilles*. PhD thesis, University Paris 9 Dauphine, 2000.

[12] B. Roorda, J. Engwerda, and H. Schumacher. Performance of hedging strategies in interval models. *Kybernetika*, Vol 41, pp 575–592, 2005.

[13] G. Shafer and V. Vovk. *Probability and finance. It's only a game!* Wiley Interscience, New York, 2001.

[14] S. Thiery. *Évaluation d'options vanilles et digitales dans le modèle de marché à intervalles*, PhD thesis, Université de Nice-Sophia Antipolis, to appear.

A Search Game with a Strategy of Energy Supply for Target

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

Keiko Ikeda
Bureau of Defense Policy
Ministry of Defense
Tokyo, Japan

Abstract

This paper deals with a two-person zero-sum game called the search allocation game (SAG). A searcher distributes his searching resources in a search space to detect a target while the target moves around in the search space to evade the searcher. The target movement is subject to energy constraints, which means that the target spends some energy to move and he cannot move to other places on exhaustion of his energy. However, the target can supply his energy at the risk of letting himself more likely to be detected by the searcher. A strategy of the searcher is a distribution of his searching resources and a target strategy is to select a path running in the search space and make a decision of energy supply. We propose two linear programming formulations to solve the SAG with energy supply. We also analyze some characteristics of optimal strategies by some numerical examples.

Key words. Search game, two-person zero-sum game, linear programming, dynamic programming.

AMS Subject Classifications. Primary 90B40; Secondary 91A05, 90C05.

1 Introduction

In Search Theory, two models have been mainly studied so far for search games with moving targets: *search-and-evasion game* and *search allocation game* (SAG) [6]. This paper deals with the SAG. The SAG is a two-person zero-sum game, in which a searcher and a target take part. The searcher distributes his searching resources in a search space to detect the target. On the other hand, the target moves in the search space to avoid the searcher.

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_16,

The search problem has many applications such as search-and-rescue activities and military operations in the ocean. At first, the research on search problems started from one-sided optimization problems. Koopman [17] got together the results of the naval Operations Research activities of U.S. Navy in World War II. He studied so-called datum search, where a target took a diffusive motion from an exposed point on a randomly selected course. Meinardi [18] models the datum search as a search game. He considers a discrete model, in which the search is executed in a discrete space at discrete time points. To solve the game, he tries to manipulate the target transition so as to make the probability distribution of the target as uniform as possible in the space at all times. That is why his method is difficult to be applied to other search problems. His model is one of the search-and-evasion games. The direct application of the datum search model could be military operations such as anti-submarine warfare (ASW). Danskin [3] deals with a search game of ASW, where a submarine selects a course and speed at the beginning of the search while an ASW helicopter chooses a sequence of points for dipping a sonar, and he finds an equilibrium point of the game. Baston and Bostock [1] and Garnaev [5] discuss games to determine the best points of hiding a submarine and throwing down depth charges by an ASW airplane in a one-dimensional discrete space. Washburn's work [21] is about a multi-stage game with target's and searcher's discrete motions, where both players have no restriction on their motion and a payoff is the total traveling distance until a coincidence of their positions. Nakai [19] deals with an interesting model in the sense that a safety zone is set up for the target. His model is also a multi-stage game with the payoff of the detection probability of the target. Kikuta [15,16] studies a game with the payoff of traveling cost. Eagle

In those studies, the searcher's strategy is to choose his search paths. But it could be a distribution of searching resources in a search space, especially in the case that the searcher can move faster than the target and can move wherever he likes. Such a game is called SAG. For the SAG, a basic problem is to determine a hiding point for a stationary target and a distribution plan of searching resources of searchers (see Garnaev [6]). Nakai [20] and Iida et al. [13] show research on such games with stationary targets. Concerning moving target games, there are Hohzaki and Iida's papers [14,8,10]. Hohzaki and Iida [9] propose a numerical method to solve more generalized game, where it is just required that a payoff is concave for searcher's strategy and linear for target's strategy.

In most of the studies on search games outlined above, authors set comparatively simple constraints on target motion. That is why optimal solutions keep a kind of uniformity in the problems and then they are easy to be estimated. In Washburn and Hohzaki [22,12], they consider a datum search game with energy constraint in a continuous search space. The energy constraint helps the problem to be more practical but carries off uniformity from optimal solutions. They could not succeed to derive an optimal solution because the continuous space is more difficult to deal with than the discrete one for optimization problems, but they propose an estimation method for lower and upper bounds on the value of the game. In a discrete space, Hohzaki et al. [11] propose an exact method to solve a SAG with energy constraint. Furthermore, Hohzaki [7] elucidates a relation between two types of SAGs defined on a continuous space and a discrete space.

In Washburn and Hohzaki's papers [22,12], they handle the target energy as a state of the target. Dambreville and Le Cadre [2] also take account of a state space of the target in their model. However, not to mention cars or ships, they can usually supply the vehicles with fuel or energy. The energy is crucial to the mobility of the target and the mobility gives the target the flexibility to avoid the searcher. This paper deals with the supply of energy as one of target strategies.

In the next section, we describe a SAG model with energy supply. In Sec. 3, we propose a method to solve the SAG by a linear programming problem and give another linear programming formulation to cope with a large size of problems. We take some numerical examples to analyze characteristics of optimal strategies of players in Sec. 4.

2 SAG Model with Energy Supply

Here we define a SAG in which a searcher and target participate, in a discrete search space.

A1. A search space consists of a discrete geographic cell space and a time space. The geographic cell space is denoted by $\boldsymbol{K} = \{1, \cdots, K\}$ and the time space by $\boldsymbol{T} = \{1, \cdots, T\}$.

A2. A target moves in the search space. However, his motion is subject to the following conditions. The target starts from one of cells $S_0 \subseteq \boldsymbol{K}$. From a cell i at time t, he can move only to a set of cells $N(i,t)$ at the next time $t+1$. He expends energy $\mu(i,j)$ to move from cell i to j so that he cannot move farther than cells that his current energy allows him to go to. He has initial energy e_0. On exhaustion of his energy, he must stay in his current cell and cannot move to other cells. He can supply his energy by $\delta > 0$ during a time period. He chooses one of two options concerning the supply of energy: supply or no-supply. If he takes a strategy of supply, his energy increases by δ at the end of a relevant time point or at the beginning of the next time. This strategy, which we call energy-supply strategy, affects the likelihood that the target is detected, which is explained in detail later in Assumption (A (4)).

Let us denote a target path by $\omega = \{\omega(t), t \in \boldsymbol{T}\}$, where $\omega(t)$ indicates a target position at time t, and an energy-supply strategy by $\sigma = \{\sigma(t), t \in \boldsymbol{T}\}$, where $\sigma(t)$ is 0 for a no-supply strategy and 1 for a supply strategy at time t.

A3. To detect the target, a searcher begins to distribute his searching resources from time τ and then we denote a time period available for the searching by $\widehat{\boldsymbol{T}} = \{\tau, \tau+1, \cdots, T\}$. The searcher can use an amount of searching resources $\Phi(t)$ at most at time $t \in \widehat{\boldsymbol{T}}$. The available resources are continuously divisible.

We denote a distribution strategy of searching resources by $\varphi = \{\varphi(i,t),\ i \in \boldsymbol{K},\ t \in \boldsymbol{T}\}$, where $\varphi(i,t)$ is the nonnegative amount of resources to be distributed in cell i at time t by the searcher.

A4. For a target strategy $\rho = (\omega, \sigma)$ and a searcher strategy φ, the searcher can detect the target with probability $1 - \exp(-g(\varphi, \rho))$, where $g(\varphi, \rho)$ is a weighted amount of effective resources distributed along path ω and its exact expression will be provided later. The weight indicates effectiveness of unit searching resource in cell i over the detection of the target. The weight depends on an energy-supply strategy of the target there. It is α_i for the no-supply strategy and β_i for the supply-strategy, while $\alpha_i < \beta_i$ usually holds.

On detection of the target, the searcher gets reward 1 but the target loses the same. A payoff of the game is assumed to be given by the searcher's reward.

In Assumption (A (2)), initial cells of the target S_0 could be $\boldsymbol{K}$ or a point. If it consists of a point, we can model a datum search game, where the target starts from an exposed point. $N(i,t)$ is a representative for practical geographical constraints.

As a path strategy, the target chooses one path, which is a mapping $\omega(t)$ from $\boldsymbol{T}$ to $\boldsymbol{K}$. He also chooses his energy-supply strategy $\sigma(t)$ each time, which is a mapping from $\boldsymbol{T}$ to $\boldsymbol{S} \equiv \{0,1\}$. A residual energy of the target at time t, $e(t)$, is given deterministically by his path strategy and energy-supply strategy. On a path ω, he expends energy $\mu(\omega(t), \omega(t+1))$ after moving from cell $\omega(t)$ to $\omega(t+1)$, but he can increase his energy by δ by a supply strategy. His energy $e(t+1)$ becomes $e(t+1) = e(t) - \mu(\omega(t), \omega(t+1)) + \delta\sigma(t)$ at the beginning of time point $t+1$. We represent a combination of a path strategy ω and an energy-supply strategy σ by $\rho = (\omega, \sigma) = \{(\omega(t), \sigma(t)), t \in \boldsymbol{T}\}$ and call a combination a routing strategy of the target. Let us denote a feasibility region of routing strategy $\rho = (\omega, \sigma)$ by $\boldsymbol{P}$, which is given by the following conditions.

- Initial position: $\omega(0) \in S_0$
- Geographical constraint: $\omega(t+1) \in N(\omega(t), t),\ t = 1, \cdots, T-1$
- Initial energy: $e(1) = e_0$
- Conservation law of energy: $e(t+1) = e(t) - \mu(\omega(t), \omega(t+1)) + \delta\sigma(t),\ t = 1, \cdots, T-1$
- Moving energy constraint: $\mu(\omega(t), \omega(t+1)) \le e(t),\ t = 1, \cdots, T-1$

From Assumption (A (3)), a feasible region of searcher's strategy φ is given by:

$$\Psi = \{\varphi \mid \sum_{i \in \boldsymbol{K}} \varphi(i,t) \le \Phi(t),\ \varphi(i,t) \ge 0,\ i \in \boldsymbol{K},\ t \in \widehat{\boldsymbol{T}}\}. \tag{1}$$

This problem is a two-person zero-sum game with the payoff of the detection probability of the target because the expected reward of the searcher is just the detection probability. In the game, the searcher is a maximizer and the target is a minimizer.

Let us formulate a payoff of the game or a function $g(\cdot)$. Assume that a target takes a routing strategy $\rho = (\omega, \sigma) \in \boldsymbol{P}$ and a searcher takes a distribution plan of searching resources φ. At time t, the target is in cell $\omega(t)$, where the searcher scatters $\varphi(\omega(t), t)$ searching resources. The effectiveness of unit resource is $\alpha_{\omega(t)}$ for a no-supply strategy of the target and $\beta_{\omega(t)}$ for a supply strategy. Now we can estimate the weighted amount of resources over path ω by:

$$g(\varphi, \rho) = \sum_{t \in \widehat{\boldsymbol{T}}} \varphi(\omega(t), t)\{\alpha_{\omega(t)}(1 - \sigma(t)) + \beta_{\omega(t)}\sigma(t)\}. \tag{2}$$

From Assumption (A (4)), a payoff is defined by:

$$R(\varphi, \rho) = 1 - \exp\left(-\sum_{t \in \widehat{\boldsymbol{T}}} \varphi(\omega(t), t)\{\alpha_{\omega(t)}(1 - \sigma(t)) + \beta_{\omega(t)}\sigma(t)\}\right). \tag{3}$$

Here, let us take a mixed strategy for the target, $\pi = \{\pi(\rho),\ \rho \in \boldsymbol{P}\}$, where $\pi(\rho)$ is the probability that the target chooses a routing strategy ρ. A feasible region of π is

$$\Pi = \{\pi(\rho) \mid \sum_{\rho \in \boldsymbol{P}} \pi(\rho) = 1,\ \pi(\rho) \geq 0,\ \rho \in \boldsymbol{P}\}\ . \tag{4}$$

For a pure strategy of the searcher φ and a mixed strategy of the target π, an expected payoff is given by $R(\varphi, \pi) = \sum_\rho \pi(\rho) R(\varphi, \rho)$. Because the expected payoff is linear for π and strictly concave for φ, we already know that the game has an equilibrium point, that is, a minimax value of the expected payoff coincides with a maximin value (See Hohzaki [9]).

3 Equilibrium Point

Here we propose two linear programming formulations to solve the game and derive optimal strategies of the searcher and the target.

3.1 Formulation using routing strategy of target

From here, we are going to find the value of the game by solving a maximin optimization problem and derive an equilibrium point for strategies φ and π. Noting $\sum_{\rho \in \boldsymbol{P}} \pi(\rho) = 1$, we can transform the problem $\max_\varphi \min_\pi R(\varphi, \pi)$ as follows:

$$\begin{aligned} \max_{\varphi \in \Psi} \min_{\pi \in \Pi} R(\varphi, \pi) &= \max_{\varphi \in \Psi} \min_{\pi \in \Pi} \sum_\rho \pi(\rho) R(\varphi, \rho) = \max_{\varphi \in \Psi} \min_{\rho \in \boldsymbol{P}} R(\varphi, \rho) \\ &= \max_{\varphi \in \Psi, \zeta} \{\zeta \mid 1 - \exp(-g(\varphi, \rho)) \geq \zeta,\ \rho \in \boldsymbol{P}\}\ . \end{aligned}$$

A replacement of $\eta = \ln(1/(1-\zeta))$ changes the above expression to the following.

$$
\begin{aligned}
&= \max_{\varphi\in\Psi,\eta} \{1-\exp(-\eta) \mid g(\varphi,\rho) \geq \eta,\ \rho \in \boldsymbol{P}\} \\
&= 1-\exp(-\max_{\varphi\in\Psi,\eta} \{\eta \mid g(\varphi,\rho) \geq \eta,\ \rho \in \boldsymbol{P}\}) \qquad (5)\\
&= 1-\exp\left(-\max_{\varphi\in\Psi,\eta}\right. \\
&\left.\left\{\eta \;\middle|\; \sum_{t\in\widehat{\boldsymbol{T}}} \varphi(\omega(t),t)\{\alpha_{\omega(t)}(1-\sigma(t))+\beta_{\omega(t)}\sigma(t)\} \geq \eta,\ (\omega,\sigma)\in\boldsymbol{P}\right\}\right).
\end{aligned}
$$

The above problem is equivalent to the following linear programming problem:

$$
\begin{aligned}
P^S \ :\quad & \max_{\varphi,\eta}\ \eta \\
s.t. \quad & \sum_{t\in\widehat{\boldsymbol{T}}} \varphi(\omega(t),t)\{\alpha_{\omega(t)}(1-\sigma(t))+\beta_{\omega(t)}\sigma(t)\} \geq \eta,\ (\omega,\sigma)\in\boldsymbol{P} \qquad (6)\\
& \sum_{i\in\boldsymbol{K}} \varphi(i,t) \leq \Phi(t),\ t\in\widehat{\boldsymbol{T}} \qquad (7)\\
& \varphi(i,t) \geq 0,\ i\in\boldsymbol{K},\ t\in\widehat{\boldsymbol{T}}\,. \qquad (8)
\end{aligned}
$$

Using an optimal value of the above problem, η^*, we can calculate the maximin value of the original game or the value of the game by $1-\exp(-\eta^*)$. At the same time, a solution of problem (P^S) gives us an optimal searcher strategy φ^*.

Since problem (P^S) is nothing but the maximin problem with a linear form of expected payoff $R(\varphi,\pi) = \sum_\rho \pi(\rho)g(\varphi,\rho)$ from Eq. (5), we adopt this function $R(\varphi,\pi)$ as a new expected payoff from now on. Let us turn our attention to a minimax problem with this expected payoff. We transform $R(\varphi,\pi)$ first, like this:

$$
\begin{aligned}
R(\varphi,\pi) &= \sum_{\rho=(\omega,\sigma)\in\boldsymbol{P}} \pi(\rho) \sum_{t\in\widehat{\boldsymbol{T}}} \varphi(\omega(t),t)\{\alpha_{\omega(t)}(1-\sigma(t))+\beta_{\omega(t)}\sigma(t)\} \\
&= \sum_{t\in\widehat{\boldsymbol{T}}} \sum_{i\in\boldsymbol{K}} \sum_{\rho=(\omega,\sigma)\in\boldsymbol{P}} \delta_{i\omega(t)}\pi(\rho)\varphi(i,t)\{\alpha_i(1-\sigma(t))+\beta_i\sigma(t)\} \\
&= \sum_{t\in\widehat{\boldsymbol{T}}} \sum_{i\in\boldsymbol{K}} \sum_{\rho=(\omega,\sigma)\in\boldsymbol{P}_{it}} \pi(\rho)\varphi(i,t)\{\alpha_i(1-\sigma(t))+\beta_i\sigma(t)\}, \qquad (9)
\end{aligned}
$$

where δ_{ij} is the Kronecker's delta and $\boldsymbol{P}_{it}$ is a set of routing strategies with paths running through cell i at time t. $\boldsymbol{P}_{it}$ is defined by $\boldsymbol{P}_{it} \equiv \{\rho = (\omega,\sigma)\in\boldsymbol{P} \mid \omega(t) = i\}$.

Noting $\sum_{i\in \boldsymbol{K}} \varphi(i,t) \le \Phi(t)$, we can transform a maximization problem $\max_\varphi R(\varphi,\pi)$ with respect to φ, as follows:

$$\max_{\varphi\in\Psi} R(\varphi,\pi) \\ = \sum_{t\in\widehat{\boldsymbol{T}}} \Phi(t) \max_{i\in\boldsymbol{K}} \sum_{\rho=(\omega,\sigma)\in \boldsymbol{P}_{it}} \pi(\rho)\{\alpha_i(1-\sigma(t)) + \beta_i\sigma(t)\}. \tag{10}$$

By introducing variable $\nu(t)$ as a substitute for $\max_i \sum_{\rho=(\omega,\sigma)\in\boldsymbol{P}_{it}} \pi(\rho)\,\{\alpha_i(1-\sigma(t)) + \beta_i\sigma(t)\}$, we obtain a minimax value of $R(\varphi,\pi)$ by the following linear programming formulation:

$$P^T \; : \; \min_{\pi,\nu} \sum_{t\in\widehat{\boldsymbol{T}}} \Phi(t)\nu(t)$$

$$s.t. \sum_{\rho=(\omega,\sigma)\in\boldsymbol{P}_{it}} \pi(\rho)\{\alpha_i(1-\sigma(t)) + \beta_i\sigma(t)\} \le \nu(t),\; t\in\widehat{\boldsymbol{T}},\; i\in\boldsymbol{K} \tag{11}$$

$$\sum_{\rho\in\boldsymbol{P}} \pi(\rho) = 1 \tag{12}$$

$$\pi(\rho) \ge 0,\; \rho\in\boldsymbol{P}\;. \tag{13}$$

We can obtain an optimal mixed strategy of the target, π^*, by solving problem (P^T). On the other hand, an optimal searcher's strategy, φ^*, is given by problem (P^S), as we mentioned before. Both problems give us the value of the game, of course. Practically, we can easily verify that (P^S) and (P^T) are dual each other and then both problems have an identical optimal value. We only need to solve one of these problems once in order to obtain optimal strategies of both players. Now we state our results.

Theorem 3.1. *The value of the game is given as an optimal value of problem* (P^S) *or* (P^T). *An optimal strategy of the searcher,* φ^*, *is given as an optimal solution of* (P^S) *or optimal dual variables corresponding to condition* (11) *in* (P^T). *An optimal strategy of the target,* π^*, *is given as an optimal solution of* (P^T) *or optimal dual variables corresponding to condition* (6) *in* (P^S).

3.2 Formulation using transition probability of target

As seen from conditions (6) or (13), formulations (P^S) and (P^T) enumerate all routing strategies of the target $\boldsymbol{P}$. However, if there is no constraint on the target motion, we can count the number of strategies up to $(2|\boldsymbol{K}|)^{|\boldsymbol{T}|}$ at most in search space $\boldsymbol{K}\times\boldsymbol{T}$ so that we doubt a computation by (P^S) or (P^T) in feasible computation time although they are linear programming problems. That is why we are going to develop an alternative formulation to cope with the large size of the search space.

For simplicity, we assume that energy consumption function $\mu(i,j)$ is integer-valued, and initial energy e_0 and the increase of energy δ by a supply strategy are also integers in Assumption (A (2)). Now we can denote a set of energy states of the target by $\boldsymbol{E} = \{0, \cdots, e_0, \cdots, e_0 + T\delta\}$. To represent a state of the target, we can use a combination of four elements, say (i,e,s,t), which indicates that the target is in cell i with residual energy e under a supply $(s=1)$ or no-supply $(s=0)$ strategy at time t. The mixture of target routes, $\pi(\rho),\ \rho \in \boldsymbol{P}$, generates a probability distribution of the state (i,e,s,t). We denote the probability that the target is in state (i,e,s,t) by $q(i,e,s,t)$. We also define the transition probability, that the target in state (i,e,s,t) moves to cell j under an energy-supply strategy r at the next time $t+1$, by $v(i,e,s,t,j,r)$. In this subsection, we take variables $q(\cdot)$ and $v(\cdot)$ as a new target strategy while $\varphi(\cdot)$ is still a searcher's strategy. We define other notation. $C(i,e,t) = \{j \in N(i,t) | \mu(i,j) \le e\}$ is a set of cells, to which the target in state (i,e,s,t) can move at the next time. $C^*(r,i,e,t) = \{j \in \boldsymbol{K} | i \in C(j, e + \mu(j,i) - \delta r, t-1)\}$ is a set of cells, from which the target under an energy-supply strategy r at time $t-1$ moves into state (i,e,s,t).

Let us review the expression of the expected payoff (9) and transform it further:

$$R(\varphi,\pi) = \sum_{t \in \hat{\boldsymbol{T}}} \sum_{i \in \boldsymbol{K}} \varphi(i,t) \sum_{\rho=(\omega,\sigma) \in \boldsymbol{P}_{it}} \pi(\rho)\{\alpha_i(1-\sigma(t)) + \beta_i \sigma(t)\}$$
$$= \sum_{t \in \hat{\boldsymbol{T}}} \sum_{i \in \boldsymbol{K}} \varphi(i,t) \left(\alpha_i \sum_{\rho \in \boldsymbol{P}_{it,\sigma(t)=0}} \pi(\rho) + \beta_i \sum_{\rho \in \boldsymbol{P}_{it,\sigma(t)=1}} \pi(\rho) \right). \quad (14)$$

In the last expression, $\sum_{\rho \in \boldsymbol{P}_{it,\sigma(t)=0}} \pi(\rho)$ or $\sum_{\rho \in \boldsymbol{P}_{it,\sigma(t)=1}} \pi(\rho)$ means the existence probability that the target is in cell i at time t under a no-supply strategy $s=0$ or under a supply strategy $s=1$, respectively. We can replace these probabilities by $\sum_{e \in \boldsymbol{E}} q(i,e,s,t)$ with $s=0$ or $s=1$ using new target strategy $q(i,e,s,t)$ to obtain another expression for the expected payoff:

$$R(\varphi,\pi) = \sum_{t \in \hat{\boldsymbol{T}}} \sum_{i \in \boldsymbol{K}} \varphi(i,t) \left(\alpha_i \sum_{e \in \boldsymbol{E}} q(i,e,0,t) + \beta_i \sum_{e \in \boldsymbol{E}} q(i,e,1,t) \right). \quad (15)$$

Now we are going to solve a minimax optimization for the expected payoff with new target strategy $q(\cdot)$ and $v(\cdot)$. First of all, let us define a value $h(t)$ as a maximum payoff rewarded by an optimal distribution of searching resources after time t. From Eq. (14), a part of payoff $\varphi(i,t)\alpha_i \sum_{e \in \boldsymbol{E}} q(i,e,0,t)$ is locally produced in cell i at time t for $s=0$ or $\varphi(i,t)\beta_i \sum_{e \in \boldsymbol{E}} q(i,e,1,t)$ for $s=1$. We can see that the following dynamic programming formulation holds between $h(t)$ and $h(t+1)$:

$$\begin{aligned} & h(t) \\ & = \max_{\varphi \in \Psi} \left\{ \sum_{i \in \boldsymbol{K}} \varphi(i,t) \left(\alpha_i \sum_{e \in \boldsymbol{E}} q(i,e,0,t) + \beta_i \sum_{e \in \boldsymbol{E}} q(i,e,1,t) \right) + h(t+1) \right\} \\ & = \Phi(t) \max_{i \in \boldsymbol{K}} \sum_{e \in \boldsymbol{E}} (\alpha_i q(i,e,0,t) + \beta_i q(i,e,1,t)) + h(t+1). \end{aligned} \tag{16}$$

In a brace in the first line, the first term indicates the expected payoff yielded at time t and the second one a maximal payoff expected after $t+1$. An expression in the second line is derived from a feasibility constraint $\sum_{i \in \boldsymbol{K}} \varphi(i,t) \leq \Phi(t)$. Equation (16) leads us to an inequality $h(t) \geq \Phi(t) \sum_{e \in \boldsymbol{E}} (\alpha_i q(i,e,0,t) + \beta_i q(i,e,1,t)) + h(t+1)$ for any cell i. $h(\tau)$ is a maximum expected payoff over the whole search, which is just what the target desires to minimize. We can formulate a linear programming problem for optimal target strategy $q^*,\ v^*$ and the value of the game:

$$(P_m^T) \quad \min \ h(\tau)$$

$$s.t. \ \ h(t) \geq \Phi(t) \sum_{e \in \boldsymbol{E}} (\alpha_i q(i,e,0,t) + \beta_i q(i,e,1,t)) + h(t+1), \quad i \in \boldsymbol{K},\ t = \tau, \cdots, T-1 \tag{17}$$

$$h(T) \geq \Phi(T) \sum_{e \in \boldsymbol{E}} (\alpha_i q(i,e,0,T) + \beta_i q(i,e,1,T)),\ i \in \boldsymbol{K} \tag{18}$$

$$q(i,e,s,t) = \sum_{r \in \boldsymbol{S}} \sum_{j \in C(i,e,t)} v(i,e,s,t,j,r), \quad i \in \boldsymbol{K},\ e \in \boldsymbol{E},\ s \in \boldsymbol{S},\ t = 1, \cdots, T-1 \tag{19}$$

$$q(i,e,s,t) = \sum_{r \in \boldsymbol{S}} \sum_{j \in C^*(r,i,e,t)} v(j, e+\mu(j,i) - \delta r, r, t-1, i, s), \quad i \in \boldsymbol{K},\ e \in \boldsymbol{E},\ s \in \boldsymbol{S},\ t = 2, \cdots, T \tag{20}$$

$$\sum_{i \in S_0} \sum_{s \in \boldsymbol{S}} q(i, e_0, s, 1) = 1 \tag{21}$$

$$\sum_{i \in \boldsymbol{K}} \sum_{e \in \boldsymbol{E}} \sum_{s \in \boldsymbol{S}} q(i,e,s,t) = 1,\ t \in \boldsymbol{T} \tag{22}$$

$$v(i,e,s,t,j,r) \geq 0,\ i,j \in \boldsymbol{K}, e \in \boldsymbol{E}, s,r \in \boldsymbol{S}, t = 1, \cdots, T-1. \tag{23}$$

Conditions (17) and (18) come from the recursive Eq. (16). Conditions (19), (20), and (22), represent the so-called conservation law of probability flows $q(\cdot)$ and $v(\cdot)$. Equation (21) is derived from initial conditions of the target: initial cells S_0 and initial energy e_0. To solve this linear programming problem, we need the order $O(4|\boldsymbol{K}|^2|\boldsymbol{T}||\boldsymbol{E}|)$ of computer memories for variable $v(i,e,s,t,j,r)$ at most.

Our next job is to obtain an optimal strategy of the searcher. We also make use of a dynamic programming equation to formulate another linear programming problem for an optimal searcher strategy. Let us define a value $w(i,e,s,t)$ as a minimum expected payoff by an optimal route of the target starting from state (i,e,s,t) given a searcher strategy $\varphi(i,t)$. In cell i at time $t(\geq \tau)$, an energy-supply strategy $s=0$ or $s=1$ brings out a local payoff $\alpha_i\varphi(i,t)$ or $\beta_i\varphi(i,t)$, respectively. From the definition of $w(i,e,s,t)$, we have:

$$w(i,e,s,t) = \min_{j\in C(i,e,t)} \min_{r\in \boldsymbol{S}} \{(\alpha_i(1-s)+\beta_i s)\varphi(i,t) + w(j, e-\mu(i,j)+\delta s, r, t+1)\} \quad (24)$$

after the search starts at time τ. At terminal time T, there occurs a local payoff $w(i,e,s,T) = (\alpha_i(1-s)+\beta_i s)\varphi(i,T)$. At time $t=1,\cdots,\tau-1$ before τ, any payoff never occurs and then we have a recursive equation

$$w(i,e,s,t) = \min_{j\in C(i,e,t)} \min_{r\in \boldsymbol{S}} w(j, e-\mu(i,j)+\delta s, r, t+1). \quad (25)$$

Over the whole search, the target can minimize the expected payoff down to $\min_{i\in S_0}\min_{s\in \boldsymbol{S}} w(i,e_0,s,1)$. On the other hand, the searcher wishes to maximize the minimum payoff and his optimal strategy produces the maximin value of the expected payoff. The above discussion validates the following linear programming formulation.

$$\begin{aligned}
&(P_m^S) \quad \max\ \xi \\
&s.t.\ \ w(i,e_0,s,1) \geq \xi,\ i\in S_0,\ s\in \boldsymbol{S} \qquad (26)\\
&\quad w(i,e,s,1) = 0,\ i\in \boldsymbol{K},\ e\in \boldsymbol{E}\backslash\{e_0\},\ s\in \boldsymbol{S}\\
&\quad w(i,e,s,1) = 0,\ i\in \boldsymbol{K}\backslash S_0,\ e\in \boldsymbol{E},\ s\in \boldsymbol{S}\\
&\quad w(i,e,s,t) \leq w(j,e-\mu(i,j)+\delta s, r, t+1),\\
&\qquad\qquad i\in \boldsymbol{K},\ j\in C(i,e,t),\ e\in \boldsymbol{E},\ s,r\in \boldsymbol{S},\ t=1,\cdots,\tau-1\\
&\quad w(i,e,0,t) \leq \alpha_i\varphi(i,t) + w(j,e-\mu(i,j),r,t+1),\\
&\qquad\qquad i\in \boldsymbol{K},\ j\in C(i,e,t),\ e\in \boldsymbol{E},\ r\in \boldsymbol{S},\ t=\tau,\cdots,T-1\\
&\quad w(i,e,1,t) \leq \beta_i\varphi(i,t) + w(j,e-\mu(i,j)+\delta,r,t+1),\\
&\qquad\qquad i\in \boldsymbol{K},\ j\in C(i,e,t),\ e\in \boldsymbol{E},\ r\in \boldsymbol{S},\ t=\tau,\cdots,T-1\\
&\quad w(i,e,0,T) = \alpha_i\varphi(i,T),\ i\in \boldsymbol{K},\ e\in \boldsymbol{E}\\
&\quad w(i,e,1,T) = \beta_i\varphi(i,T),\ i\in \boldsymbol{K},\ e\in \boldsymbol{E}\\
&\quad \sum_{i\in \boldsymbol{K}} \varphi(i,t) = \Phi(t),\ t=\tau,\cdots,T\\
&\quad \varphi(i,t) \geq 0,\ i\in \boldsymbol{K},\ t=\tau,\cdots,T.
\end{aligned}$$

Even though $w(j,e_0,s,1) > \xi$ for some $j\in S_0$, we can change the variable to $w(j,e_0,s,1)=\xi$ without affecting an optimal value ξ. That is why we can replace inequality (26) with $w(i, e_0, s, 1) = \xi$. Anyway, we ought to note that the order $O(2|\boldsymbol{K}||\boldsymbol{T}||\boldsymbol{E}|)$ of memory size is required to solve problem (P_m^S). This order is less than problem (P_m^T).

From now, let us prove a duality between problems (P_m^S) and (P_m^T). The duality would imply that we only need to solve one of the problems once to obtain optimal strategies of the searcher and the target. Let us generate a dual problem corresponding to problem (P_m^T). We assign dual variables $\eta(i,t)$ to condition (17) and $\eta(i,T)$ to condition (18). After setting dual variables $y(i,e,s,t),\ z(i,e,s,t),\ \lambda$ and $\nu(t)$ corresponding to (19), (20), (21), and (22), respectively, we have the following dual problem:

$$(D_1^T)\ \ \max\ \lambda + \sum_{t=1}^{T} \nu(t)$$

$$s.t.\ \ \lambda + y(i, e_0, s, 1) + \nu(1) = 0,\ i \in S_0,\ s \in \boldsymbol{S}$$

$$y(i,e,s,1) + \nu(1) = 0,\ i \in \boldsymbol{K},\ e \in \boldsymbol{E}\backslash\{e_0\},\ s \in \boldsymbol{S}$$

$$y(i,e,s,1) + \nu(1) = 0,\ i \in \boldsymbol{K}\backslash S_0,\ e \in \boldsymbol{E},\ s \in \boldsymbol{S}$$

$$z(i,e,s,t) + y(i,e,s,t) + \nu(t) = 0,\quad i \in \boldsymbol{K},\ e \in \boldsymbol{E},\ s \in \boldsymbol{S},\ t = 2,\cdots,\tau-1 \tag{27}$$

$$z(i,e,0,t) + y(i,e,0,t) + \nu(t) - \alpha_i \Phi(t)\eta(i,t) = 0,\quad i \in \boldsymbol{K},\ e \in \boldsymbol{E},\ t = \tau,\cdots,T-1 \tag{28}$$

$$z(i,e,1,t) + y(i,e,1,t) + \nu(t) - \beta_i \Phi(t)\eta(i,t) = 0,\quad i \in \boldsymbol{K},\ e \in \boldsymbol{E},\ t = \tau,\cdots,T-1 \tag{29}$$

$$z(i,e,0,T) + \nu(T) - \alpha_i \Phi(T)\eta(i,T) = 0,\ i \in \boldsymbol{K},\ e \in \boldsymbol{E}$$

$$z(i,e,1,T) + \nu(T) - \beta_i \Phi(T)\eta(i,T) = 0,\ i \in \boldsymbol{K},\ e \in \boldsymbol{E}$$

$$-y(i,e,s,t) - z(j, e-\mu(i,j)+s\delta, r, t+1) \le 0,\quad i \in \boldsymbol{K},\ e \in \boldsymbol{E},\ s,r \in \boldsymbol{S},\ t = 1,\cdots,T-1,\ j \in C(i,e,t)$$

$$\sum_{i \in \boldsymbol{K}} \eta(i,\tau) = 1 \tag{30}$$

$$-\sum_{i \in \boldsymbol{K}} \eta(i,t) + \sum_{i \in \boldsymbol{K}} \eta(i,t+1) = 0,\ t = \tau,\cdots,T-1 \tag{31}$$

$$\eta(i,t) \ge 0,\ i \in \boldsymbol{K},\ t = \tau,\cdots,T.$$

There is redundancy between variables y and z, as seen from a transformation $y(i,e,s,t) = -z(i,e,s,t) - \nu(t)$ of Eq. (27). After carefully replacing $y(i,e,s,t)$

with $z(i,e,s,t)$ by Eqs. (27), (28), and (29), that is,

$$
\begin{aligned}
y(i,e,s,t) &= -z(i,e,s,t)-\nu(t),\ i\in\boldsymbol{K}, e\in\boldsymbol{E}, s\in\boldsymbol{S}, t=2,\cdots,\tau-1\\
y(i,e,0,t) &= -z(i,e,0,t)-\nu(t)+\alpha_i\Phi(t)\eta(i,t),\\
&\qquad\qquad i\in\boldsymbol{K},\ e\in\boldsymbol{E},\ t=\tau,\cdots,T-1\\
y(i,e,1,t) &= -z(i,e,1,t)-\nu(t)+\beta_i\Phi(t)\eta(i,t),\\
&\qquad\qquad i\in\boldsymbol{K},\ e\in\boldsymbol{E},\ t=\tau,\cdots,T-1,
\end{aligned}
$$

and putting together Eqs. (30) and (31), we have another formulation. There we insert a newly defined variable $z(i,e,s,1)\equiv -y(i,e,s,1)-\nu(1)$, which has not been defined yet in problem (D_1^T).

$$
\begin{aligned}
&(D_2^T)\ \ \max\ \lambda+\sum_{t=1}^{T}\nu(t)\\
&s.t\ \ z(i,e_0,s,1)=\lambda,\ i\in S_0,\ s\in\boldsymbol{S}\\
&\qquad z(i,e,s,1)=0,\ i\in\boldsymbol{K},\ e\in\boldsymbol{E}\backslash\{e_0\},\ s\in\boldsymbol{S}\\
&\qquad z(i,e,s,1)=0,\ i\in\boldsymbol{K}\backslash S_0,\ e\in\boldsymbol{E},\ s\in\boldsymbol{S}\\
&\qquad z(i,e,s,t)+\nu(t)\le z(j,e-\mu(i,j)+s\delta,r,t+1),\\
&\qquad\qquad i\in\boldsymbol{K},\ e\in\boldsymbol{E},\ s,r\in\boldsymbol{S},\ t=1,\cdots,\tau-1, j\in C(i,e,t)\\
&\qquad z(i,e,0,t)+\nu(t)\le \alpha_i\Phi(t)\eta(i,t)+z(j,e-\mu(i,j),r,t+1)\\
&\qquad\qquad i\in\boldsymbol{K},\ e\in\boldsymbol{E},\ t=\tau,\cdots,T-1,\ j\in C(i,e,t)\\
&\qquad z(i,e,1,t)+\nu(t)\le \beta_i\Phi(t)\eta(i,t)+z(j,e-\mu(i,j)+\delta,r,t+1)\\
&\qquad\qquad i\in\boldsymbol{K},\ e\in\boldsymbol{E},\ t=\tau,\cdots,T-1,\ j\in C(i,e,t)\\
&\qquad z(i,e,0,T)+\nu(T)=\alpha_i\Phi(T)\eta(i,T),\ i\in\boldsymbol{K},\ e\in\boldsymbol{E}\\
&\qquad z(i,e,1,T)+\nu(T)=\beta_i\Phi(T)\eta(i,T),\ i\in\boldsymbol{K},\ e\in\boldsymbol{E}\\
&\qquad \sum_{i\in\boldsymbol{K}}\eta(i,t)=1,\ t=\tau,\cdots,T,\\
&\qquad \eta(i,t)\ge 0,\ i\in\boldsymbol{K},\ t=\tau,\cdots,T.
\end{aligned}
$$

Using $w(i,e,s,t)\equiv z(i,e,s,t)+\sum_{k=t}^{T}\nu(k)$ instead of $z(i,e,s,t)$, we modify the problem further.

$$
\begin{aligned}
&(D_3^T)\ \ \max\ \lambda+\sum_{t=1}^{T}\nu(t)\\
&s.t.\ \ w(i,e_0,s,1)=\lambda+\sum_{t=1}^{T}\nu(t),\ i\in S_0,\ s\in\boldsymbol{S}\\
&\qquad w(i,e,s,1)=\sum_{t=1}^{T}\nu(t),\ i\in\boldsymbol{K},\ e\in\boldsymbol{E}\backslash\{e_0\},\ s\in\boldsymbol{S} \qquad (32)
\end{aligned}
$$

$$w(i,e,s,1) = \sum_{t=1}^{T} \nu(t),\ i \in \boldsymbol{K} \backslash S_0,\ e \in \boldsymbol{E},\ s \in \boldsymbol{S} \tag{33}$$

$$w(i,e,s,t) \le w(j, e-\mu(i,j)+s\delta, r, t+1),$$
$$i \in \boldsymbol{K}, e \in \boldsymbol{E}, s, r \in \boldsymbol{S}, t = 1, \cdots, \tau-1, j \in C(i,e,t) \tag{34}$$

$$w(i,e,0,t) \le \alpha_i \Phi(t)\eta(i,t) + w(j, e-\mu(i,j), r, t+1),$$
$$i \in \boldsymbol{K},\ e \in \boldsymbol{E},\ t = \tau, \cdots, T-1,\ j \in C(i,e,t) \tag{35}$$

$$w(i,e,1,t) \le \beta_i \Phi(t)\eta(i,t) + w(j, e-\mu(i,j)+\delta, r, t+1),$$
$$i \in \boldsymbol{K},\ e \in \boldsymbol{E},\ t = \tau, \cdots, T-1,\ j \in C(i,e,t) \tag{36}$$

$$w(i,e,0,T) = \alpha_i \Phi(T)\eta(i,T),\ i \in \boldsymbol{K},\ e \in \boldsymbol{E} \tag{37}$$

$$w(i,e,1,T) = \beta_i \Phi(T)\eta(i,T),\ i \in \boldsymbol{K},\ e \in \boldsymbol{E} \tag{38}$$

$$\sum_{i \in \boldsymbol{K}} \eta(i,t) = 1,\ t = \tau, \cdots, T \tag{39}$$

$$\eta(i,t) \ge 0,\ i \in \boldsymbol{K},\ t = \tau, \cdots, T.$$

We can prove that the following formulation gives the same optimal value as problem (D_3^T):

$$(D_4^T)\ \ \max\ \xi$$
$$s.t.\ \ w(i,e_0,s,1) = \xi,\ i \in S_0,\ s \in \boldsymbol{S}$$
$$w(i,e,s,1) = 0,\ i \in \boldsymbol{K},\ e \in \boldsymbol{E} \backslash \{e_0\},\ s \in \boldsymbol{S}$$
$$w(i,e,s,1) = 0,\ i \in \boldsymbol{K} \backslash S_0,\ e \in \boldsymbol{E},\ s \in \boldsymbol{S}$$
$$w(i,e,s,t) \le w(j, e-\mu(i,j)+s\delta, r, t+1),$$
$$i \in \boldsymbol{K},\ e \in \boldsymbol{E},\ s, r \in \boldsymbol{S},\ t = 1, \cdots, \tau-1,\ j \in C(i,e,t)$$
$$w(i,e,0,t) \le \alpha_i \Phi(t)\eta(i,t) + w(j, e-\mu(i,j), r, t+1),$$
$$i \in \boldsymbol{K},\ e \in \boldsymbol{E},\ t = \tau, \cdots, T-1,\ j \in C(i,e,t)$$
$$w(i,e,1,t) \le \beta_i \Phi(t)\eta(i,t) + w(j, e-\mu(i,j)+\delta, r, t+1),$$
$$i \in \boldsymbol{K},\ e \in \boldsymbol{E},\ t = \tau, \cdots, T-1,\ j \in C(i,e,t)$$
$$w(i,e,0,T) = \alpha_i \Phi(T)\eta(i,T),\ i \in \boldsymbol{K},\ e \in \boldsymbol{E}$$
$$w(i,e,1,T) = \beta_i \Phi(T)\eta(i,T),\ i \in \boldsymbol{K},\ e \in \boldsymbol{E}$$
$$\sum_{i \in \boldsymbol{K}} \eta(i,t) = 1,\ t = \tau, \cdots, T$$
$$\eta(i,t) \ge 0,\ i \in \boldsymbol{K},\ t = \tau, \cdots, T.$$

Since (D_4^T) is given by adding one more condition $\sum_{t=1}^{m} \nu(t) = 0$ to (D_3^T), an optimal value of (D_3^T) is equal to or larger than (D_4^T). At the same time, we can say that a value of (D_4^T) is equal to or larger than (D_3^T) by the following reason. Variable $w(i,e,s,T)$ is nonnegative from Eqs. (37) and (38). Since problem (D_3^T) is to maximize $w(i,e_0,s,1)$, variable $w(i,e,s,t)$ varies as large as possible while satisfying inequalities (34), (35), and (36). In the result, the values of expressions (32) and (33) become nonnegative. Given an optimal solution $w^*(i,e,s,t),\ \nu^*(t),\ \lambda^*$ and $\eta^*(i,t)$ for (D_3^T), we can obtain the same objective value even though we set $\nu(t)=0$ for $t\in\boldsymbol{T},\ \lambda=\lambda^*+\sum_t \nu^*(t)$ and then change Eqs. (32) and (33) to $w(i,e,s,1)=0$ for either of $i\in\boldsymbol{K}\backslash S_0$ or $e\in\boldsymbol{E}\backslash\{e_0\}$ remaining other variables unchanged. Therefore, an optimal value of problem (D_4^T) is equal to or lager than (D_3^T). We conclude that problem (D_3^T) is equivalent to (D_4^T), which is the same as (P_m^S). We need to remember a remark stated just after formulation (P_m^S), though.

Now we have completed the proof for the duality between problems (P_m^S) and (P_m^T). At the same time, we can see that an optimal strategy $\varphi(i,t)$ of the searcher is given as $\Phi(t)\eta(i,t)$ of problem (D_4^T) when we compare problems (D_4^T) and (P_m^S). Summing up the discussion so far, we obtain the following theorem.

Theorem 3.2. *The value of the game is given as an optimal value of problem* (P_m^T) *or* (P_m^S)*. An optimal strategy of the searcher,* φ^**, is given by an optimal solution of* (P_m^S) *or optimal dual variables corresponding to conditions* (17) *and* (18) *in* (P_m^T)*. An optimal strategy of the target,* $q(\cdot),\ v(\cdot)$*, is given by an optimal solution of* (P_m^T).

4 Numerical Examples

Here, let us consider some examples of our search game, where a search space is $\boldsymbol{K}\times\boldsymbol{T}=\{1,\cdots,9\}\times\{1,\cdots,7\}$. A target starts from cell $S_0=\{1\}$ at time $t=1$ and a searcher begins his search at time $\tau=2$ to chase the target. Cells are aligned in the order of $1,2,\cdots,9$ and the target can move to 3-neighbored cells from his current cell, that is, $N(i,t)=\{i-3,i-2,\cdots,i,\cdots,i+3\}\cap\boldsymbol{K}$. He has initial energy $e_0=4$ and expends energy $\mu(i,j)=|i-j|^2$ to move from cell i to j. He has to make a decision about no-supply $(s=0)$ or supply $(s=1)$ as well as a moving path to follow. For a no-supply strategy $s=0$, effectiveness of unit searching resource on target detection is assumed to be the same for all cells or $\alpha_i=1$. For $s=1$, effectiveness β_i is also the same for all cells, but we change it in some cases to analyze its influence upon optimal strategies of players. By a supply strategy $s=1$, the target can increase his current energy by δ. We set $\delta=1$ in many cases but change it in Case 4.

As seen from Eq. (15), $\alpha_i\sum_{e\in\boldsymbol{E}} q(i,e,0,t)+\beta_i\sum_{e\in\boldsymbol{E}} q(i,e,1,t)$ is a contribution of unit searching resource in cell i at time t on the payoff of the game. Then the number could be an indicator, from which the searcher can judge attractive spots to scatter his resources. We call the number "strength of detection" in cell i at time t. From the same Eq. (15), the target can regard $\varphi(i,t)$ as an indicator to recognize his unfavorable spots by and he ought to move so as to form smaller strength of detection for places with larger $\varphi(i,t)$.

(1) Case 1 ($\beta=1$).

Let us consider an unusual case with $\beta_i = 1$, where the detectability of the target never depends on his decision making of $s = 0$ or $s = 1$. Table 1-a, with a coordinate of time points and cell numbers, shows an optimal distribution of the target $\sum_{e \in \boldsymbol{E}} \sum_{s \in \boldsymbol{S}} q(i, e, s, t)$. To discriminate between two energy-supply strategies, we place $\sum_{e \in \boldsymbol{E}} q(i, e, 1, t)$ for $s = 1$ in upper positions and $\sum_{e \in \boldsymbol{E}} q(i, e, 0, t)$ for $s = 0$ in lower positions in Table 1-b. Table 1-c shows an optimal distribution of searching resources for the searcher.

As we see in Table 1-a, the target gradually expands his possible area and makes his distribution uniform in the area. Corresponding to the target strategy, the searcher selects a uniform distribution of his searching resources, as seen in Table 1-c. The strength of detection is the same as Table 1-a because of $\alpha_i = \beta_i = 1$. In this case, a supply strategy $s = 1$ never brings the target any demerit, and then the target secures the flexibility of moving by taking a supply strategy $s = 1$ more often than $s = 0$, as seen in Table 1-b, to expand his possible area and make a perfectly uniform distribution of his existence or strength of detection.

We can say that "expansion of possible area" and "uniformity of strength of detection" are key pointers for an optimal target strategy. The former strategy forces the searcher to scatter his resources over a wider area and makes the scattered resources much thinner. The latter disturbs an efficient search, which the searcher can achieve by focusing searching resources on cells with high density of the target distribution. Consequently, both strategies yield smaller payoff.

Table 1-a: Optimal distribution of target (Case 1: $\beta = 1.0$)

9	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0.125
7	0	0	0	0	0	0.143	0.125
6	0	0	0	0	0.167	0.143	0.125
5	0	0	0	0.2	0.167	0.143	0.125
4	0	0	0.25	0.2	0.167	0.143	0.125
3	0	0.333	0.25	0.2	0.167	0.143	0.125
2	0	0.333	0.25	0.2	0.167	0.143	0.125
1	1	0.333	0.25	0.2	0.167	0.143	0.125
cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7

(2) Case 2 ($\beta = 1.5$).

In the case of $\beta = 1.5$, an optimal distribution of the target, an optimal selection of the energy-supply strategy and the strength of detection are shown in Tables 2-a, 2-b, and 2-c. Table 2-d illustrates an optimal distribution of searching resources for the searcher. From Table 2-b, the number of times the target takes a supply strategy $s = 1$ in the search dramatically decreases compared to Table 1-b and the supply is executed near the boundary of the target possible area, $(t, i) = (1, 1), (3, 3), (3, 4), (4, 5), (5, 6), (6, 6), (6, 7), (7, 8)$.

Table 1-b: Optimal strategy of energy supply (Case 1: $\beta = 1.0$)

cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7
	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0
	0	0	0	0	0	0	0.063
8	0	0	0	0	0	0	0.063
	0	0	0	0	0	0.101	0.063
7	0	0	0	0	0	0.042	0.063
	0	0	0	0	0.129	0.085	0.063
6	0	0	0	0	0.037	0.058	0.063
	0	0	0	0.161	0.108	0.080	0.063
5	0	0	0	0.039	0.059	0.063	0.063
	0	0	0.203	0.137	0.100	0.077	0.063
4	0	0	0.047	0.063	0.067	0.066	0.063
	0	0.276	0.176	0.125	0.095	0.076	0.063
3	0	0.058	0.074	0.075	0.071	0.067	0.063
	0	0.233	0.163	0.119	0.093	0.075	0.063
2	0	0.101	0.087	0.081	0.074	0.068	0.063
	0.745	0.238	0.163	0.122	0.094	0.074	0.063
1	0.255	0.096	0.087	0.078	0.073	0.069	0.063

Table 1-c: Optimal distribution of searching resources (Case 1: $\beta = 1.0$)

cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7
9		0	0	0	0	0	0
8		0	0	0	0	0	0.125
7		0	0	0	0	0.143	0.125
6		0	0	0	0.167	0.143	0.125
5		0	0	0.2	0.167	0.143	0.125
4		0	0.25	0.2	0.167	0.143	0.125
3		0.333	0.25	0.2	0.167	0.143	0.125
2		0.333	0.25	0.2	0.167	0.143	0.125
1		0.333	0.25	0.2	0.167	0.143	0.125

Table 2-a: Optimal distribution of target (Case 2: $\beta = 1.5$)

9	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0.081
7	0	0	0	0	0	0.115	0.131
6	0	0	0	0	0.118	0.108	0.131
5	0	0	0	0.153	0.176	0.155	0.131
4	0	0	0.194	0.212	0.176	0.155	0.131
3	0	0.333	0.268	0.212	0.176	0.155	0.131
2	0	0.333	0.269	0.212	0.176	0.155	0.131
1	1	0.333	0.269	0.212	0.176	0.155	0.131
cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7

Table 2-b: Optimal strategy of energy supply (Case 2: $\beta = 1.5$)

	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0
	0	0	0	0	0	0	0.030
8	0	0	0	0	0	0	0.051
	0	0	0	0	0	0.081	0
7	0	0	0	0	0	0.034	0.131
	0	0	0	0	0.118	0.094	0
6	0	0	0	0	0	0.014	0.131
	0	0	0	0.118	0	0	0
5	0	0	0	0.035	0.176	0.155	0.131
	0	0	0.150	0	0	0	0
4	0	0	0.044	0.212	0.176	0.155	0.131
	0	0	0.003	0	0	0	0
3	0	0.333	0.265	0.212	0.176	0.155	0.131
	0	0	0	0	0	0	0
2	0	0.333	0.269	0.212	0.176	0.155	0.131
	0.75	0	0	0	0	0	0
1	0.25	0.333	0.269	0.212	0.176	0.155	0.131
cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7

Table 2-c: Strength of detection (Case 2: $\beta = 1.5$)

9	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0.095
7	0	0	0	0	0	0.155	0.131
6	0	0	0	0	0.176	0.155	0.131
5	0	0	0	0.212	0.176	0.155	0.131
4	0	0	0.269	0.212	0.176	0.155	0.131
3	0	0.333	0.269	0.212	0.176	0.155	0.131
2	0	0.333	0.269	0.212	0.176	0.155	0.131
1	1.375	0.333	0.269	0.212	0.176	0.155	0.131
cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7

Table 2-d: Optimal distribution of searching resources (Case 2: $\beta = 1.5$)

9		0	0	0	0	0	0
8		0	0	0	0	0	0
7		0	0	0	0	0.148	0.074
6		0	0	0	0.138	0.148	0.148
5		0	0	0.093	0.172	0.148	0.156
4		0	0.25	0.227	0.172	0.139	0.156
3		0.333	0.25	0.227	0.172	0.139	0.156
2		0.333	0.25	0.227	0.172	0.139	0.156
1		0.333	0.25	0.227	0.172	0.139	0.156
cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7

The energy supply at $t = 7$ is no use or rather harmful for the target since the search terminates at the time. However, we can neglect the effect of the energy supply on the payoff because the searcher does not allocate any searching resource there. As seen by comparing probabilities of choosing $s = 1$ each other, the probability is larger at earlier time when the target possible area does not expand large enough. In the above observations, we can find an intention of the target to expand his possible area as large as possible. The target keeps his existence probabilities smaller in the cells where he refuels, and then constructs a uniform distribution of the strength of detection almost everywhere except $(t, i) = (7, 8)$ in Table 2-c. The table indicates another intention of the target to make the strength of detection as uniform as possible.

In Table 2-d of an optimal searcher's strategy, although uniformity is a basic feature especially in interior areas, there are some perturbation such as less distributed resources at $(t, i) = (4, 5), (5, 6), (7, 6), (7, 7)$ and more resources at $(t, i) = (6, 5), (6, 6), (6, 7)$. The searcher allocates more searching resources around the boundary area at time 6 to catch the target running far from cell 1 and then the probabilities of the target are smaller at $(t, i) = (6, 6), (6, 7)$ in Table 2-a. It causes less value 0.081 or 0.095 for the target probability or the strength of detection, respectively, in cell 8 at time 7. That is why the searcher can save his searching resources there and use them in other interior cells.

(3) Case 3 ($\beta = 4$).

We set $\beta = 4$ to make the demerit of a strategy $s = 1$ larger than Case 2 for the target. Tables 3-a, 3-b, 3-c, and 3-d show an optimal distribution of the target, an optimal selection of the energy-supply strategy, the strength of detection and an optimal distribution of searching resources. We can explain qualitative features appeared in those tables in a similar manner to Case 2. However, let us compare Table 3 with Table 2 to make some differences between them clear. In Table 3-b, points where the target takes a supply strategy $s = 1$ are more restricted to boundary areas $(t, i) = (1, 1), (3, 4), (4, 5), (5, 6), (6, 7), (7, 7), (7, 8)$ than in Table 2-b. The energy supply at $(t, i) = (7, 7), (7, 8)$ is negligible from the view of the payoff of the game by the same reason as in Case 2. The probabilities that the target chooses $s = 1$ are smaller than in Table 2-b. These smaller probabilities make the strength of detection smaller at $(t, i) = (4, 5), (6, 7), (7, 7), (7, 8)$ in Table 3-c. We see a uniform distribution in the table as its basic property, though.

In Table 3-d, we can see less resources allocated at $(t, i) = (3, 4)$, $(5, 5)$, $(5, 6), (6, 6)$ and more resources at $(t, i) = (4, 4), (7, 6)$ while the table has uniformity in interior areas. As explained in Case 2, this feature indicates a wise strategy of the searcher taking account of the continuity of target motion in the search space. In Table 3-c, the searcher does not distribute any searching resources at four points where the strength of detection is smaller than other cells.

(4) Case 4 ($\beta = 1.5,\ \delta = 4$):

In Case 2, the target increases his energy by $\delta = 1$ each time he takes a supply strategy $s = 1$. Here, we set $\delta = 4$ as supply energy, by which the target can move farther. Tables 4-a, 4-b, 4-c, and 4-d show optimal strategies of players. The target has the high mobility to reach cell 13 at time 7. The energy supply brings some disadvantage for the target, but it

Table 3-a: Optimal distribution of target (Case 3: $\beta = 4$)

9	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0.013
7	0	0	0	0	0	0.048	0.035
6	0	0	0	0	0.048	0.159	0.159
5	0	0	0	0.060	0.190	0.159	0.159
4	0	0	0.116	0.235	0.190	0.159	0.159
3	0	0.333	0.295	0.235	0.190	0.159	0.159
2	0	0.333	0.295	0.235	0.190	0.159	0.159
1	1	0.333	0.295	0.235	0.190	0.159	0.159
cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7

Table 3-b: Optimal strategy of energy supply (Case 3: $\beta = 4$)

	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0
	0	0	0	0	0	0	0.006
8	0	0	0	0	0	0	0.007
	0	0	0	0	0	0.024	0.014
7	0	0	0	0	0	0.023	0.021
	0	0	0	0	0.048	0	0
6	0	0	0	0	0	0.159	0.159
	0	0	0	0.052	0	0	0
5	0	0	0	0.008	0.190	0.159	0.159
	0	0	0.060	0	0	0	0
4	0	0	0.057	0.235	0.190	0.159	0.159
	0	0	0	0	0	0	0
3	0	0.333	0.295	0.235	0.190	0.159	0.159
	0	0	0	0	0	0	0
2	0	0.333	0.295	0.235	0.190	0.159	0.159
	0.708	0	0	0	0	0	0
1	0.292	0.333	0.295	0.235	0.190	0.159	0.159
cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7

Table 3-c: Strength of detection (Case 3: $\beta = 4$)

9	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0.030
7	0	0	0	0	0	0.121	0.078
6	0	0	0	0	0.190	0.159	0.159
5	0	0	0	0.215	0.190	0.159	0.159
4	0	0	0.295	0.235	0.190	0.159	0.159
3	0	0.333	0.295	0.235	0.190	0.159	0.159
2	0	0.333	0.295	0.235	0.190	0.159	0.159
1	3.123	0.333	0.295	0.235	0.190	0.159	0.159
cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7

gives the target higher mobility at the same time. Due to the advantage, the target tends to adopt the supply strategy more often than Tables 2-b and 3-b, as seen in Table 4-b. In Cases 2 and 3, the target refuels his energy only around boundary areas, but in this case, he does even in interior areas. The energy supply around the boundary area is useful for the expansion of the target possible area and the supply in the interior area is also helpful for flattening the distribution of the target in the interior area.

Table 3-d: Optimal distribution of searching resources (Case 3: $\beta = 4$)

9		0	0	0	0	0	0
8		0	0	0	0	0	0
7		0	0	0	0	0	0
6		0	0	0	0.103	0.151	0.182
5		0	0	0	0.079	0.170	0.164
4		0	0.156	0.344	0.204	0.170	0.164
3		0.333	0.281	0.219	0.204	0.170	0.164
2		0.333	0.281	0.219	0.204	0.170	0.164
1		0.333	0.281	0.219	0.204	0.170	0.164
cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7

For example, at time 6, the target supplies energy in interior cells $6, 7$ as well as in boundary cells $9, 10, 11$, and the supply helps the distribution of the target uniform widely at time 7. The strength of detection has a uniform distribution almost everywhere in Table 4-c except cells $12, 13$ at time 7. Table 4-d shows similar features to other cases that the searcher distributes his searching resources in a uniform fashion in interior areas but disturbs the fashion near boundary areas. It is difficult to give the fashion any exact quantitative explanation, but the searcher can save his resources for cells $12, 13$ at time 7 due to smaller strength of detection there.

As the last analysis by numerical examples, let us compare the values of the game, which are calculated by Eq. (15), in Cases 1–4. The values are $1.22,\ 1.28,$ and 1.37 for Cases 1, 2, and 3, respectively, because large β_i works advantageously for the searcher and disadvantageously for the target. In Case 4 with $\delta = 4$, the high mobility of the target pushes the value of the game down to 1.02.

Table 4-a: Optimal distribution of target (Case 4: $\beta = 1.5,\ \delta = 4$)

13	0	0	0	0	0	0	0.031
12	0	0	0	0	0	0	0.050
11	0	0	0	0	0	0.073	0.083
10	0	0	0	0	0	0.082	0.083
9	0	0	0	0	0.086	0.093	0.083
8	0	0	0	0	0.092	0.098	0.083
7	0	0	0	0.106	0.093	0.075	0.083
6	0	0	0	0.122	0.090	0.088	0.083
5	0	0	0.151	0.138	0.128	0.098	0.083
4	0	0	0.193	0.158	0.128	0.098	0.083
3	0	0.333	0.211	0.159	0.128	0.098	0.083
2	0	0.333	0.223	0.159	0.128	0.098	0.083
1	1	0.333	0.223	0.159	0.128	0.098	0.083
cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7

Table 4-b: Optimal strategy of energy supply (Case 4: $\beta = 1.5,\ \delta = 4$)

	0	0	0	0	0	0	0.013
13	0	0	0	0	0	0	0.018
	0	0	0	0	0	0	0.022
12	0	0	0	0	0	0	0.029
	0	0	0	0	0	0.049	0
11	0	0	0	0	0	0.024	0.083
	0	0	0	0	0	0.032	0
10	0	0	0	0	0	0.050	0.083
	0	0	0	0	0.085	0.010	0
9	0	0	0	0	0.001	0.083	0.083
	0	0	0	0	0.071	0	0
8	0	0	0	0	0.021	0.098	0.083
	0	0	0	0.105	0.071	0.047	0
7	0	0	0	0.001	0.022	0.027	0.083
	0	0	0	0.073	0.076	0.021	0
6	0	0	0	0.049	0.013	0.066	0.083
	0	0	0.145	0.041	0	0	0
5	0	0	0.006	0.097	0.128	0.098	0.083
	0	0	0.060	0.001	0	0	0
4	0	0	0.132	0.157	0.128	0.098	0.083
	0	0	0.023	0	0	0	0
3	0	0.333	0.188	0.159	0.128	0.098	0.083
	0	0	0	0	0	0	0
2	0	0.333	0.223	0.159	0.128	0.098	0.083
	0.894	0	0	0	0	0	0
1	0.106	0.333	0.223	0.159	0.128	0.098	0.083
cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7

Table 4-c: Strength of detection (Case 4: $\beta = 1.5,\ \delta = 4$)

13	0	0	0	0	0	0	0.038
12	0	0	0	0	0	0	0.061
11	0	0	0	0	0	0.098	0.083
10	0	0	0	0	0	0.098	0.083
9	0	0	0	0	0.128	0.098	0.083
8	0	0	0	0	0.128	0.098	0.083
7	0	0	0	0.159	0.128	0.098	0.083
6	0	0	0	0.159	0.128	0.098	0.083
5	0	0	0.223	0.159	0.128	0.098	0.083
4	0	0	0.223	0.159	0.128	0.098	0.083
3	0	0.333	0.223	0.159	0.128	0.098	0.083
2	0	0.333	0.223	0.159	0.128	0.098	0.083
1	1.447	0.333	0.223	0.159	0.128	0.098	0.083
cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7

Table 4-d: Optimal distribution of searching resources (Case 4: $\beta = 1.5,\ \delta = 4$)

13		0	0	0	0	0	0
12		0	0	0	0	0	0
11		0	0	0	0	0.043	0.022
10		0	0	0	0	0.043	0.052
9		0	0	0	0.092	0.060	0.061
8		0	0	0	0.092	0.102	0.061
7		0	0	0.125	0.103	0.107	0.115
6		0	0	0.125	0.119	0.107	0.115
5		0	0.2	0.15	0.119	0.107	0.115
4		0	0.2	0.15	0.119	0.107	0.115
3		0.333	0.2	0.15	0.119	0.107	0.115
2		0.333	0.2	0.15	0.119	0.107	0.115
1		0.333	0.2	0.15	0.119	0.107	0.115
cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7

5 Conclusions

In this paper, we deal with a two-person zero-sum game called search allocation game (SAG), where a searcher distributes searching resources to detect a target and the target moves to evade the searcher. Each time the target moves in the search space, he can refuel his energy for his high-mobility at the risk of being more detectable. This study comes from a SAG with target energy, which Washburn and Hohzaki [22,12] modeled first. They pay attention to the target energy or the target mobility from the practical point of view. Through this study, we elucidate what an important role the energy or the mobility plays for search games.

In this paper, we formulate the game as a linear programming problem based on target paths. We also derive another linear programming formulation based on existence probability and transition probability of the target to cope with large size of the game. We obtain some analytical results about these formulations, e.g., duality. By some numerical examples, we show in detail that two players cleverly play the game each other. For an optimal target strategy, it is vital that the target expands his possible area and constructs a uniform distribution of his existence. For an optimal strategy of the searcher, it is important that the searcher restricts the target motion and focuses his searching resources on some restricted areas in an effective way. We can make it clear how wisely the target utilizes the energy-supply strategy to materialize the above vital situations and how wisely the searcher allocates his searching resources in the space. We expect this study or our proposed methods to be applied to practical search problem or more complicated search situations.

References

[1] Baston V.J., Bostock F.A., *A one-dimensional helicopter-submarine game*, Naval Research Logistics, **36** (1989), 479–490.

[2] Dambreville F., Le Cadre J.P., *Search game for a moving target with dynamically generated informations*, Proceedings of the 5th International Conference on Information Fusion (FUSION'2002), (2002), 243–250.

[3] Danskin J.M., *A helicopter versus submarine search game*, Operations Research, **16** (1968), 509–517.

[4] Eagle J.N., Washburn A.R., *Cumulative search-evasion games*, Naval Research Logistics, **38** (1991), 495–510.

[5] Garnaev A.Y., *A remark on a helicopter-submarine game*, Naval Research Logistics, **40** (1993), 745–753.

[6] Garnaev A.Y., *Search Games and Other Applications of Game Theory*, Springer-Verlag, Tokyo, (2000).

[7] Hohzaki R., *Search allocation game*, European Journal of Operational Research, **172** (2006), 101–119.

[8] Hohzaki R., Iida K., *A search game with reward criterion*, Journal of the Operations Research Society of Japan, **41** (1998), 629–642.

[9] Hohzaki R., Iida K., *A solution for a two-person zero-sum game with a concave payoff function*, in W. Takahasi and T. Tanaka (ed.), *Nonlinear Analysis and Convex Analysis*, World Science Publishing Co., London, (1999), 157–166.

[10] Hohzaki R., Iida K., *A search game when a search path is given*, European Journal of Operational Research, **124** (2000), 114–124.

[11] Hohzaki R., Iida K., Komiya T., *Discrete search allocation game with energy constraints*, Journal of the Operations Research Society of Japan, **45** (2002), 93–108.

[12] Hohzaki R, Washburn A., *An approximation for a continuous datum search game with energy constraint*, Journal of the Operations Research Society of Japan, **46** (2003), 306–318.

[13] Iida K., Hohzaki R., Sato K, *Hide-and-search game with the risk criterion*, Journal of the Operations Research Society of Japan, **37** (1994), 287–296.

[14] Iida K., Hohzaki R., Furui S., *A search game for a mobile target with the conditionally deterministic motion defined by paths*, Journal of the Operations Research Society of Japan, **39** (1996), 501–511.

[15] Kikuta K, *A hide and seek game with traveling cost*, Journal of the Operations

Research Society of Japan, **33** (1990), 168–187.

[16] Kikuta K., *A search game with traveling cost*, Journal of the Operations Research Society of Japan, **34** (1991), 365–382.

[17] Koopman B.O., *The theory of search III: the optimum distribution of searching effort*, Operations Research, **5** (1957), 613–626.

[18] Meinardi J.J., *A sequentially compounded search game*, in *Theory of Games: Technique and Applications*, The English Universities Press, London, (1964), 285–299.

[19] Nakai T., *A sequential evasion-search game with a goal*, Journal of the Operations Research Society of Japan, **29** (1986), 113–122.

[20] Nakai T., *Search models with continuous effort under various criteria*, Journal of the Operations Research Society of Japan, **31** (1988), 335–351.

[21] Washburn A.R., *Search-evasion game in a fixed region*, Operations Research, **28** (1980), 1290–1298.

[22] Washburn A.R., Hohzaki R., *The diesel submarine flaming datum problem*, Military Operations Research, **4** (2001), 19–30.

The LQG Game Against Nature

Meir Pachter
Department of Electrical and Computer Engineering
Air Force Institute of Technology
Wright-Patterson AFB, OH 45433, USA
meir.pachter@afit.edu

Abstract

A novel Linear Quadratic Gaussian (LQG) control paradigm where the realization of the random disturbance is known at decision time is introduced. A game of chance is played where Nature randomly chooses the disturbance signal and the player, having access to the random disturbance information, employs a feedback/stroboscopic strategy. The closed-form optimal solution is derived and is compared to the solution of the classical LQG problem. Furthermore, the Gaussian assumption is relaxed and the interesting game theoretic interpretation is analyzed.

Key words. LQG Game, Stochastic Game, Stochastic Dynamic Programming.

AMS Subject Classifications. Primary 91A25; Secondary 49N90.

1 Introduction

The following stochastic control problem is of interest. In support of the current COUNTER program [1], in reference [2] a simplified fixed wing Micro Air Vehicles (MAVs) based, Intelligence, Surveillance and Reconnaissance (ISR) scenario was considered. It envisaged a camera equipped MAV which is deployed to a battle space populated by n objects of interest, most of which are clutter objects. Only a small fraction of the objects of interest are Targets (Ts). Targets are characterized by a certain feature which clutter objects, a.k.a. False Targets (FTs), do not possess. The MAV sequentially inspects the n objects of interest. As the MAV flies over an object, the image of the inspected object is transmitted to a human operator for classification. The classification is received after a random operator delay, during which time the MAV continues flying toward the next object of interest. The statistics of the operator delay are known. When the received image is ambiguous and the feature characterizing a target is not discerned by the operator, revisiting the object and sending to the operator a new image, possibly from a more favorable aspect angle, might be conducive to classification, and consequently, an increase in the information about this object. At the same time, the MAV's endurance is limited and thus, there is a cost associated with turning around and revisiting an

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_17,

object. It is required to design a control strategy to maximize the information about the n objects of interest, subject to the endurance constraint. The inspection operation was analyzed in [2]–[4], where a n-stages discrete-time stochastic optimal control problem was posed. The solution of the stochastic optimal control problem yielded the optimal inspection strategy.

Stochastic control problems have dynamics which are driven by the decision variable and by a random input, often referred to as process noise. Specifically, the dynamics of the sequential inspection stochastic control problem are driven by the decision variable and by the operator's delay random variable. However, the sequential inspection stochastic optimal control problem is off the beaten path in that the realization of the random variable is known at decision time. Motivated by this insight, in this paper a novel and general Linear Quadratic Gaussian (LQG) optimal control problem which captures the essence of the sequential inspection optimization problem is formulated, where a game against Nature is played, and its closed form solution is given. Similar to the sequential inspection problem, the random variable is known at decision time, so that the information structure of the sequential inspection problem is preserved. Whereas in the sequential inspection problem the cost functional is linear in the decision variable, however the dynamics are nonlinear, in the LQG problem the dynamics are linear and the cost is quadratic, which affords a closed form solution. Thus, in this paper the novel closed form solution is compared to the solution of the classical LQG problem with the conventional information pattern. Hence, the role of the information structure is brought to the foreground. Furthermore, the Gaussian assumption is relaxed and the interesting game theoretic interpretation is analyzed.

The paper is organized as follows. In Sec. 2 a brief outline of the MAV sequential inspection problem is given. It provides a vehicle for introducing stochastic control problems where the realization of the random variable is known at decision time. A TV game show scenario which could serve as a canonical example of decision problems where the realization of the random variable is known at decision time is also discussed. Thus, the set of stochastic optimal control problems where the random variable is known at decision time is not empty, which motivates us to revisit the familiar LQG paradigm. In Sec. 3, an LQG optimal control problem akin to the sequential inspection problem is introduced. The novel LQG problem can be considered a dynamic game against Nature: at each time instant Nature randomly chooses the disturbance/process noise and the Player responds with a feedback/stroboscopic strategy [5], that is, at decision time the player is privy to Nature's choice. The main result of the paper, namely, a closed form solution of the novel LQG game against Nature, is given in Sec. 4. In Sec. 5, the Gaussian assumption is relaxed and the interesting game theoretic interpretation is analyzed. The case where the Player does not have complete state information is addressed in Sec. 6, followed by concluding remarks in Sec. 7.

2 Optimal Sequential Inspection

Consider the COUNTER [1] scenario, where a MAV is required to sequentially inspect a designated set of n objects of interest. Some objects are Ts and some objects of interest might be FTs. The flight time between two objects of interest is one unit of time and the endurance of the MAV is $k_0 \geq n$. Thus, the MAV's endurance reserve is $k_0 - n\ (> 0)$ and some objects of interest can be re-visited, in which case the average information about the set of objects of interest will increase. The human operator introduces a delay τ_i, $0 \leq \tau_i < 1$. Thus, the decision of whether to turn around and revisit object of interest i, $i = 1, ..., n$, is taken at time $0 \leq \tau_i < 1$ after object i has been overflown. The operator's delay τ_i is a random variable with a known probability density distribution function $f(\tau_i)$. The decision at stage i entails the MAV's choice of whether to continue to the $i+1$ object of interest, or turn around and revisit object i, that is, the decision variable at stage i is $u_i \in \{0, 1\}$. Hence, the endurance reserve at time i, c_i, satisfies the scalar stochastic difference equation

$$c_{i+1} = c_i - 2\tau_i u_i\ , \quad i = 1, ..., n-1\ , \quad c_1 = k_0 - n. \tag{1}$$

A decision $u_i = 1$ to turn around and revisit object i is feasible, provided that the current endurance reserve c_i is sufficient, that is $c_i \geq 2\tau_i$.

The following insights are helpful.

- Revisiting an object always brings about an increase in the average information about the object.
- All the objects are similar and therefore the increase in average information is constant for all the revisited objects. Hence, the average running cost when an object is revisited is constant and is not time dependent.
- When deciding on revisiting object i, the critical factor is the currently available endurance reserve c_i. The endurance cost of a revisit of object i is the random number $2\tau_i$, The scalar state variable is the remaining endurance c_i and the dynamics are given by Eq. (1).
- In view of the above, it is clear that one should strive to maximize the expected number of objects revisited—subject to the endurance reserve constraint. Thus, the payoff functional is:

$$J = E_{\tau_1,...,\tau_n}\Big(\sum_{i=1}^{n} u_i \Big). \tag{2}$$

Optimizing the sequential inspection process calls for the solution of a stochastic optimal control problem with a scalar state c, a binary decision/control variable u, nonlinear dynamics (1), the linear in the control performance index (2), and a scalar random variable τ whose statistics are known. The scalar random variable τ directly affects the dynamics (1), similar to the decision/control variable's action, however its realization is known at decision time.

The gist of the sequential inspection optimization problem then is: while a revisit of say, object i, always increases the average information about object i

and, consequently, the average information about the set of objects increases, the endurance reserve $2\tau_i$ that must be expanded while revisiting object i might be prohibitively high. Now, the endurance cost $2\tau_i$ is a random number. Hence, when the endurance cost $2\tau_i$ is high, it makes sense to hold back and not revisit object i and instead directly continue to object $i+1$, in the hope that the endurance cost of revisiting object $i+1$ might come in lower. If this is indeed the case, then object $i+1$ is revisited while at the same time the endurance reserve is parsimoniously used and is preserved, so that the opportunity to revisit additional objects in the future is still there and the payoff (2) is maximized.

The solution entails the calculation of a sequence of threshold functions $d_i(c_i)$, $i = 1, ..., n$. The optimal control law is $u_i^*(c_i, \tau_i) = 1$, provided that $\tau_i \leq d_i(c_i)$, and $u_i^*(c_i, \tau_i) = 0$ if $\tau_i > d_i(c_i)$. This, and more complex versions of the sequential inspection problem, are solved in [2]–[4] using Stochastic Dynamic Programming (SDP).

The SDP problem considered herein has some traits of popular TV game shows. To highlight the psychological angle of such a game, assume that the player is given a budget of $\$1$ in small change, say pennies. During the game the player encounters a succession of, say $n = 10$, beggars. Each beggar will specify a “donation" he would like to receive. The amount requested by each beggar is a random number of pennies, say, between 1 penny and 20 pennies, chosen according to, e.g., a uniform probability distribution. The player can hand the beggar the requested amount, or he can decline to do so. When the beggar is given the alm, he hands the player $\$1,000,000$ in cash. When the generous beggar walks away empty handed, the player receives nothing. The player desires to maximize his expected payoff.

In this simple TV game show, problem $k = 1, ..., 10$. At stage k, when the kth beggar is encountered, the player’s remaining capital is c_k; $c_0 = \$1$. The amount requested by the kth beggar is $2\tau_k$. When the decision variable $u_k = 1$, the player obliges and hands the beggar the requested alm, and when $u_k = 0$, the player chooses to pass. The dynamics of the game are given by Eq. (1) and the player’s payoff is given by Eq. (2).

In summary, the sequential inspection problems considered herein are stochastic resource allocation problems that can be formulated as SDPs with n stages, one state variable, and a scalar random variable that affects the nonlinear dynamics, with a known p.d.f., and whose realization is known at decision time.

Motivated by this insight, a novel LQG control paradigm which captures the essence of the sequential inspection optimization problem is introduced. An LQG game against Nature is formulated and its closed form solution is given.

3 LQG Game

The LQG multivariable analog of the sequential inspection problem is addressed—we refer to the LQG game against Nature. In the state $x \in R^n$ and the Player’s decision variable $u \in R^m$, however the dynamics are linear and the cost is quadratic.

The dynamics are also affected by Nature's input, which is the random variable v. We momentarily assume v is Gaussian distributed. Most importantly, the realization of the random variable v is known at decision time. Thus, the information structure of the sequential inspection problem is preserved.

While it is comfortable to fall back to, and revisit, the familiar LQG paradigm, the fact that the realization of the random variable is *known* at decision time renders an interesting and novel LQG control paradigm. Indeed, both the original sequential inspection problem and its LQG analog are solved using Stochastic Dynamic Programming (SDP); an N stage SDP is solved.

The linear dynamics are:

$$x_{k+1} = Ax_k + Bu_k + Cv_k \ , \ \ x_0 \equiv x_0 \ , \ \ k = 0, ..., N-1. \tag{3}$$

The state $x_k \in R^n$, the Player's control variable $u_k \in R^m$, and Nature's input is the random signal $v_k \in R^l$,

$$v_k \ \sim \ N(0, V) \ ,$$

where V is a real, symmetric, and positive definite $l \times l$ matrix. The Player is interested in minimizing the performance functional

$$J(u_0, ..., u_{N-1}) = E_{v_0,...,v_{N-1}} \ (\sum_{i=1}^{N} x_i^T Q x_i + u_{i-1}^T R u_{i-1} \ ). \tag{4}$$

At decision time k, the Player knows the state x_k and the random variable v_k. Thus, his optimal control law is of the form

$$u_k^* = u_k^*(x_k, v_k). \tag{5}$$

The optimal control law entails both state feedback, as in conventional LQG control, and also knowledge of the current realization of the random variable. The action of the latter is similar to that of an additional control variable, except that this second decision maker, Nature, uses a fixed random draw policy. The strategy (5) is akin to a stroboscopic strategy in differential games [5].

3.1 Stochastic Dynamic Programming

The value function $V_k(x_k)$ satisfies the functional equation of Stochastic Dynamic Programming (SDP),

$$\begin{aligned} V_k(x_k) =& E_{v_k} \ (\ Min_{u_k} \ [u_k^T R u_k + x_{k+1}^T Q x_{k+1} + V_{k+1}(x_{k+1})] \ ) \ , \\ & k = N-2, ..., 0 \end{aligned}$$

where x_{k+1} is given by the dynamics Eq. (3). Concerning the dynamic programming stage $N-1$: the terminal value function

$$V_{N-1}(x_{N-1}) = E_{v_{N-1}} \ (\ Min_{u_{N-1}} \ [u_{N-1}^T R u_{N-1} + x_N^T Q x_N] \ )$$

where

$$x_N = Ax_{N-1} + Bu_{N-1} + Cv_{N-1}.$$

Evaluate the terminal value function first. To this end, calculate

$$\begin{aligned}
u_{N-1}^T & Ru_{N-1} + x_N^T Q x_N \\
&= u_{N-1}^T Ru_{N-1} + (Ax_{N-1} + Bu_{N-1} + Cv_{N-1})^T \\
&\quad Q(Ax_{N-1} + Bu_{N-1} + Cv_{N-1}) \\
&= u_{N-1}^T (R + B^T QB) u_{N-1} + 2u_{N-1}^T B^T QAx_{N-1} + 2u_{N-1}^T B^T \\
&\quad QCv_{N-1} \\
&\quad + 2v_{N-1}^T C^T QAx_{N-1} + x_{N-1}^T A^T QAx_{N-1} + v_{N-1}^T C^T QCv_{N-1}.
\end{aligned}$$

Hence, the optimal control is linear:

$$\begin{aligned}
u_{N-1}^* & (x_{N-1}, v_{N-1}) \\
&= -(R + B^T QB)^{-1} B^T QAx_{N-1} - (R + B^T QB)^{-1} B^T QCv_{N-1}.
\end{aligned}$$

Next calculate

$$\begin{aligned}
u_{N-1}^{*^T} & Ru_{N-1}^* + x_N^{*^T} Q x_N^* \\
&= -(x_{N-1}^T A^T QB + v_{N-1}^T C^T QB)(R + B^T QB)^{-1} \\
&\quad (B^T QAx_{N-1} + B^T QCv_{N-1}) \\
&\quad + x_{N-1}^T A^T QAx_{N-1} + v_{N-1}^T C^T QCv_{N-1} + 2v_{N-1}^T C^T QAx_{N-1} \\
&= -x_{N-1}^T A^T QB(R + B^T QB)^{-1} B^T QAx_{N-1} \\
&\quad - 2v_{N-1}^T C^T QB(R + B^T QB)^{-1} B^T QAx_{N-1} \\
&\quad - v_{N-1}^T C^T QB(R + B^T QB)^{-1} B^T QCv_{N-1} + x_{N-1}^T A^T QAx_{N-1} \\
&\quad + v_{N-1}^T C^T QCv_{N-1} + 2v_{N-1}^T C^T QAx_{N-1}.
\end{aligned}$$

Thus,

$$\begin{aligned}
u_{N-1}^{*^T} & Ru_{N-1}^* + x_N^{*^T} Q x_N^* \\
&= x_{N-1}^T A^T [Q - QB(R + B^T QB)^{-1} B^T Q] Ax_{N-1} \\
&\quad + 2v_{N-1}^T C^T Q[I - B(R + B^T QB)^{-1} B^T Q] Ax_{N-1} \\
&\quad + v_{N-1}^T C^T [Q - QB(R + B^T QB)^{-1} B^T Q] Cv_{N-1}.
\end{aligned}$$

Hence,

$$\begin{aligned}
E_{v_{N-1}} & \left(u_{N-1}^{*^T} Ru_{N-1}^* + x_N^{*^T} Q x_N^* \right) \\
&= x_{N-1}^T A^T [Q - QB(R + B^T QB)^{-1} B^T Q] Ax_{N-1} \\
&\quad + Trace\, (C^T [Q - QB(R + B^T QB)^{-1} B^T Q] CV),
\end{aligned}$$

in other words, the terminal value function

$$\begin{aligned} V_{N-1}(x_{N-1}) \;=\; & x_{N-1}^T A^T [Q - QB(R + B^T QB)^{-1} B^T Q] A x_{N-1} \\ & + \; Trace\;(\; C^T [Q - QB(R + B^T QB)^{-1} B^T Q] CV \;). \end{aligned}$$

Ansatz. The value function

$$V_k(x_k) = x_k^T P_k x_k + p_k \;, \quad k = 0, ..., N-1,$$

where P_k are real symmetric positive semi-definite matrices and p_k are non-negative numbers.
At stage k, $k = N-2, ..., 0$, the SDP recursion is then

$$\begin{aligned} x_k^T P_k x_k + p_k =& E_{v_k}\;(\; Min_{u_k}\;[u_k^T R u_k + x_{k+1}^T Q x_{k+1} \\ & + x_{k+1}^T P_{k+1} x_{k+1} + p_{k+1}]\;), \end{aligned}$$

where x_{k+1} is given by the dynamics Eq. (3).
Now:

$$\begin{aligned} u_k^T R u_k +& x_{k+1}^T Q x_{k+1} + x_{k+1}^T P_{k+1} x_{k+1} + p_{k+1} \\ =& u_k^T R u_k + x_{k+1}^T (Q + P_{k+1}) x_{k+1} + p_{k+1}. \end{aligned}$$

Hence, similar to the $N-1$ stage, and setting $Q := Q + P_{k+1}$, we obtain the linear optimal control law

$$u_k^*(x_k, v_k) = -[R + B^T(Q + P_{k+1})B]^{-1} B^T (Q + P_{k+1})[A x_k + C v_k] \qquad (6)$$

and

$$\begin{aligned} u_k^{*^T} R u_k^* +& x_{k+1}^{*^T} Q x_{k+1}^* + x_{k+1}^{*^T} P_{k+1} x_{k+1}^* + p_{k+1} \\ =& x_k^T A^T [Q + P_{k+1} - (Q + P_{k+1}) B (R \\ & + B^T (Q + P_{k+1}) B)^{-1} B^T (Q + P_{k+1})] A x_k + 2 v_k^T C^T (Q \\ & + P_{k+1})[I - B(R + B^T (Q + P_{k+1}) B)^{-1} B^T (Q \\ & + P_{k+1})] A x_k + v_k^T C^T [Q + P_{k+1} - (Q + P_{k+1}) B (R \\ & + B^T (Q + P_{k+1}) B)^{-1} B^T (Q + P_{k+1})] C v_k + p_{k+1}. \end{aligned}$$

We calculate

$$\begin{aligned} E_{v_k}\;(\; u_k^{*^T} R u_k^* + x_{k+1}^{*^T} Q x_{k+1}^* + x_{k+1}^{*^T} P_{k+1} x_{k+1}^* + p_{k+1}\;) \\ = x_k^T A^T [Q + P_{k+1} - (Q + P_{k+1}) B (R \\ + B^T (Q + P_{k+1}) B)^{-1} B^T (Q + P_{k+1})] A x_k \\ + Trace(\; C^T [Q + P_{k+1} - (Q + P_{k+1}) B (R \\ + B^T (Q + P_{k+1}) B)^{-1} B^T (Q + P_{k+1})] CV\;) \\ + p_{k+1}. \end{aligned}$$

Thus:

$$
\begin{aligned}
& x_k^T P_k x_k + p_k \\
& = x_k^T \{A^T [Q + P_{k+1} - (Q + P_{k+1}) B (R + B^T (Q + P_{k+1}) B)^{-1} B^T \\
& \quad (Q + P_{k+1})] A\} x_k \\
& \quad + Trace(\ C^T [Q + P_{k+1} - (Q + P_{k+1}) B (R + B^T (Q + P_{k+1}) B)^{-1} B^T \\
& \quad (Q + P_{k+1})] CV\ ) \\
& \quad + p_{k+1},
\end{aligned}
$$

wherefrom the recursion is obtained

$$
\begin{aligned}
P_k &= A^T \{[Q + P_{k+1} - (Q + P_{k+1}) B [R + B^T (Q + P_{k+1}) B]^{-1} \\
& \quad B^T (Q + P_{k+1})]\} A \qquad (7) \\
p_k &= Trace(C^T [Q + P_{k+1} - (Q + P_{k+1}) B [R + B^T (Q + P_{k+1}) B]^{-1} \\
& \quad B^T (Q + P_{k+1})] CV) \\
& \quad + p_{k+1}\ , \quad k = N-2, ..., 0. \qquad (8)
\end{aligned}
$$

The boundary condition at $k = N - 1$ is

$$
\begin{aligned}
P_{N-1} &= A^T [Q - QB(R + B^T QB)^{-1} B^T Q] A \qquad (9) \\
p_{N-1} &= Trace\ (\ C^T [Q - QB(R + B^T QB)^{-1} B^T Q] CV\ ). \qquad (10)
\end{aligned}
$$

The recursions (7) and (8) are solved backward in time and one obtains the real, symmetric, positive semi-definite matrix P_0, and the non-negative scalar p_0. The optimal (minimal) cost is:

$$
J^* = x_0^T P_0 x_0 + p_0. \qquad (11)
$$

We note that the optimal control (6) is a function of just one variable, namely $Ax_k + Cv_k$.

4 Discussion

Note that:

- The covariance V of the Gaussian random variable does not affect the matrices P_k and the latter are propagated as in deterministic LQ control, where $V = 0$. The scalar p_k features in the stochastic case only and is determined by the Gaussian random variable's covariance matrix V <u>and</u> the matrix P_{k+1}. Indeed, $p_k \equiv 0\ \forall\ k = 0, ..., N-1\ iff$ the covariance matrix $V = 0$.
- The calculation of the matrices P_k is decoupled from the calculation of the scalars p_k. The calculation of p_k does however require the knowledge of the P_{k+1} matrix. Hence, one first solves the recursion (7) and obtains the P_k matrix sequence, following which the recursion (8) is solved.

- Using the boundary conditions
$$P_N = 0 \tag{12}$$
$$p_N = 0 \tag{13}$$
allows us to start the recursions (7) and (8) at $k = N - 1$.
- The recursions (7) and (8) require the inversion of an $m \times m$ matrix. Since $m \leq n$, this is a relatively small matrix and in the special case of a scalar input a matrix inversion is not required, for then the expression
$$[R + B^T(Q + P_{k+1})B]^{-1} = \frac{1}{r + b^T(Q + P_{k+1})b}$$
where $b \in R^n$ is the input vector and the scalar $r > 0$ is the control effort weight.

We shall require the following.

Matrix Inversion Lemma 1. Assuming that the relevant matrices are invertible and/or have compatible dimensions, as required, the following holds:

$$(A_1 - A_2 A_4^{-1} A_3)^{-1} = A_1^{-1} + A_1^{-1} A_2 (A_4 - A_3 A_1^{-1} A_2)^{-1} A_3 A_1^{-1}.$$

$Q.E.D.$

Assuming that the matrix Q is positive definite and applying the Matrix Inversion Lemma (MIL) changes the recursion Eqs. (7) and (8). It is then required to invert two $n \times n$ matrices in the recursion equations. Nevertheless, it is instructive to apply the MIL, for it then becomes clear that under the assumption of Q positive definite, the solution of the recursion equations, namely the matrices P_k, are positive definite, and, in addition, the scalars p_k are positive.

Setting $A_1 = Q^{-1},\ A_2 = B,\ A_3 = B^T,\ A_4 = -R$ in the recursion equations, we obtain

$$P_{N-1} = A^T(Q^{-1} + BR^{-1}B^T)^{-1}A$$
$$p_{N-1} = Trace(\ C^T A^T (Q^{-1} + BR^{-1}B^T)^{-1} ACV\ ).$$

Similarly,

$$Q + P_{k+1} - (Q + P_{k+1})B(R + B^T(Q + P_{k+1})B)^{-1}B^T(Q + P_{k+1})$$
$$= [(Q + P_{k+1})^{-1} + BR^{-1}B^T]^{-1}$$

and therefore the new recursion equations are:

$$P_k = A^T[(Q + P_{k+1})^{-1} + BR^{-1}B^T]^{-1}A \tag{14}$$
$$p_k = p_{k+1} + Trace(\ C^T[(Q + P_{k+1})^{-1} + BR^{-1}B^T]CV\ )\ ,$$
$$k = N - 1, ..., 0 \tag{15}$$

with the boundary conditions (12) and (13), respectively. The new recursion Eqs. (14) and (15) are written in a more compact form, and, it is now clear that when $Q > 0$, the matrices $P_k > 0$ and the scalar sequence $p_k > 0$, $\forall$ $k = 0, ..., N-1$. At the same time, the old recursion Eqs. (7) and (8) are more efficient from a computational point of view, for they require the inversion of an $m \times m$ matrix instead of two $n \times n$ matrices, and, moreover, $m < n$.

We note that the recursions (7) or (14) for the P_k matrices are exactly the same recursions as obtained in classical LQG control, and, for that matter, also in deterministic LQ optimal control. The recursions (8) or (15) for the scalar sequence p_k are not the same as in conventional LQG control, and, of course, the stroboscopic feedback strategy (6) is a unique feature of the LQG sequential inspection problem.

In summary, a game of chance is considered and we have shown that the following holds.

Theorem 2. Consider the LQG optimal control problem (3) and (4) where the realization of Nature's random input v is known at decision time. The Player's optimal strategy is the feedback/stroboscopic control law given by Eq. (6), where, similar to the classical LQG and LQR optimal control problems, the matrices P_k are given by the recursion (7) with $P_N = 0$. The optimal cost is given by Eq. (11), which requires the solution of the additional scalar recursion (8) for p_k, with $p_N = 0$. The latter is different from the scalar recursion encountered in classical LQG optimal control. In the special case where the state error weighting matrix Q is positive definite, the respective matrix and scalar recursions (14) and (15) are used.

$Q.E.D.$

A game of chance interpretation of the LQG optimal control problem might require the real symmetric matrix Q to be negative definite. The existence of a solution then hinges on the existence of a solution to the matrix Riccati difference Eq. (7); in other words, the invertibility of the matrix $[Q+P_{k+1}-(Q+P_{k+1})B[R+B^T(Q+P_{k+1})B]$ is no longer guaranteed and we require

$$det\ ([Q + P_{k+1} - (Q + P_{k+1})B[R + B^T(Q + P_{k+1})B]) \neq 0\ , \qquad k = 0, 1, ..., N-1.$$

Note that the Player needs to know V only for the purpose of calculating ahead of time his expected payoff. In other words, if we consider the Player to be a gambler where $u_k^T R u_k$ is the running effort expanded to play the game and $x^T Q x_k$ is the running reward, and where $Q < 0$, then calculating the expected "cost" will help the gambler decide whether he wants to play; once he decides to play, the player is committed to N rounds and he cannot quit.

5 Nature's Choice

Consider a dynamic game setting. The Player strives to minimize the quadratic performance functional (4), which can be interpreted as his loss function, and

Nature is now the maximizing player. Since Nature is oblivious of the state x_k, it is choosing a random control input: strictly speaking, Nature must choose a mixed strategy. Now, the first two moments of the random variables v_k are 0 and V—Nature, in its quest to maximize the Player's loss, is constrained to choose random variables whose first and second moments are specified.

Revisiting the development in Section 3 above, one realizes that the Gaussian assumption is not critical. The following holds.

- Nature is <u>not</u> restricted to using an input which is Gaussian.
- Nature is free to choose the p.d.f. of the random variable, provided its first two moments are as specified; the expectation of the random variable being zero, is inconsequential. Thus, Nature is free to choose its mixed strategy, subject to the constraints on the first and second moments of its random control input.
- The p.d.f. chosen by Nature does not affect the value function, provided the first and second moments of Nature's random variable are fixed.
- The Player's strategy is optimal vis a vis <u>all</u> of the random strategies that Nature could throw at him, a.k.a., all the p.d.f.s that "Nature" could possibly employ.
- The optimal feedback/stroboscopic strategy (6) of the Player is not affected by the first and second moments of the random variable.
- The specified first and second moments of the random variable affect the value function.

In summary, we have shown that in the dynamic game where Nature's choice of its random input v_k is constrained s.t. the first two moments of the random variable are specified, the following holds.

Theorem 3. Consider the zero sum LQ dynamic game (3) and (4) where the realization of Nature's random input v is known at decision time. The maximizing player, Nature, does not have access to state information and therefore opts for a mixed strategy, however the "Nature" player is constrained: the first two moments of the random variable v_k are specified to be 0 and V, where V is a real symmetric positive definite matrix. Then the random variable need not be Gaussian and the results (6)–(8) and (11)–(13) hold for all p.d.f.s chosen by Nature. Furthermore, the actual p.d.f. of the random variable chosen by Nature does not affect the value of the game. The value of the game is however affected by the matrix V.

$Q.E.D.$

Remark 1. The assumption that the expectation of Nature's input is zero is inconsequential and Theorem 3 can easily be extended to include the case of a nonzero first moment of the random variable. The important thing is that the expectation of v_k is specified.

The assumption that Nature's strategy is Gaussian is not critical. Theorems 2 and 3 hold in the more general case where Nature's strategy is any p.d.f. whose first moment is zero and the second moment is the real symmetric positive definite matrix V. In order to calculate the value function, the Player does not need to know Nature's p.d.f.—knowledge of Nature's constraints, namely, knowledge of the first and second moments of Nature's random variable v_k, suffices. Thus, Nature's strategy entails randomization, however the actual strategy employed by Nature, that is, the p.d.f. of the control variable v chosen by Nature, is not known to the Player when he designs his optimal feedback/stroboscopic strategy, as is indeed the case in Game theory.

6 Partial State Information

So far, it was tacitly assumed that full state information is available to the player. In conformity with the classical LQG optimal control paradigm, the problem formulation is now generalized, to include partial state information. The Gaussian assumption is required.
Specifically, the initial state information is

$$x_0 \ \sim \ N(\overline{x}_0, \Pi_0), \tag{16}$$

where $\overline{x}_0 \in R^n$ and Π_0 is a real, symmetric, and positive definite $n \times n$ matrix. In addition, the measurement equation is introduced

$$z_{k+1} = Hx_{k+1} + \zeta_{k+1}. \tag{17}$$

The measurements $z_k \in R^r$, $k = 1, ..., N-1$. The measurement error

$$\zeta_k \ \sim \ N(0, R) \ , \tag{18}$$

where R is a real, symmetric and positive definite $r \times r$ matrix. The information available to the player at decision time k is $\overline{x}_0$, and his measurements sequence $z_1, ..., z_k$; the covariances Π_0 and R are also known.

As before, the random variable v_k—Nature's randomly chosen control variable—is known at decision time k.

Similar to Kalman's solution [6] of the LQG optimal control problem where the process noise v_k is *not* known at decision time, the *Separation Principle* also applies in the LQG game against Nature where at decision time k the random input v_k is known. However, the player's optimal strategy in the game against Nature entails a modification of Kalman's solution of the LQG optimal control problem. Consider the Kalman filter state estimation algorithm: The Kalman gain and the calculation of the updated covariance of the state estimation error Π^+_{k+1} are the same as in classical LQG optimal control, however the new propagation equations for the prior state estimate and the covariance of the prior state's estimation error are as follows:

$$\hat{x}^-_{k+1} = A\hat{x}^+_k + Bu_k + Cv_k \ , \quad \hat{x}^+_0 = \overline{x}_0 \tag{19}$$

and

$$\Pi_{k+1}^{-} = A\Pi_k^{+}A^T \ , \quad \Pi_0^{+} = \Pi_0 \ , \ k = 0, ..., N-1 \ , \tag{20}$$

respectively.
In summary, the following holds.

Theorem 4. Consider the dynamic optimization problem (3), (4) where the realization of Nature's random input v is known at decision time, the initial state uncertainty is specified according to Eq. (16), the Player's measurement equation is (17), and the sensor noise statistics are given in Eq. (18). The Player's optimal strategy entails the following two steps, taken in tandem.

(1) At decision time k, $k = 1, ..., N-1$, and once the measurement z_k has been received, obtain the minimum variance state estimate $\hat{x}_k^{+}$. To this end, run a $modified$ Kalman filter based on the dynamics (3) and the measurement Eq. (17): The modification to the Kalman filter is specified in Eqs. (19) and (20); at decision time $k = 0$, $\hat{x}_0^{+} \equiv \overline{x}_0$.
(2) Use the feedback/stroboscopic control law (6) given in Theorem 2, where x_k in Eq. (6) is replaced by the minimum variance state estimate $\hat{x}_k^{+}$.

$Q.E.D.$

Nature does not have access to state information and therefore resorts to a mixed strategy. At the same time, and also in the case where the Player has partial state information, Nature does not have to employ a Gaussian p.d.f.; Nature can choose any p.d.f., as long as the first two moments of the random inputs are as specified. Indeed, Theorem 4 hinges on the fact that the “process noise" covariance matrix V does not feature in the modified Kalman filter equations used in our game against Nature. This, in turn, follows from the stroboscopic information structure, namely, the Player's knowledge at decision time of the realization of the random variable. This is the reason that the Kalman filter, modified according to Eqs. (19) and (20), applies—provided the initial state information and the measurement noise in Eqs. (16) and (18) are Gausssian. The Gaussian assumption on the initial state information and on the measurement noise is critical, but Nature's choice, that is, the “process noise" p.d.f., need not be Gaussian. This is different from the standard LQG optimal control problem where the realization of the random variable is not known at decision time, and, for the Kalman solution to be correct, it is crucial for also the process noise to be Gaussian.

7 Conclusion

Stochastic decision and control problems have dynamics which are driven by the decision variable and by a random input, often referred to as process noise. Sequential inspection problems fall in this class. Optimizing the sequential inspection process calls for the solution of a stochastic optimal control problem with a scalar state c, nonlinear dynamics (1), a binary decision/control variable u and an additional input—the random variable τ, whose statistics are known. The random variable τ directly affects the dynamics (1), similar to the decision/control variable's action, however its realization is $known$ at decision time. Indeed, in a broad class of stochastic control problems, the random input, or disturbance, is known at decision time. Motivated by this insight, a novel Linear Quadratic Gaussian control problem where the realization of the random disturbance is known at decision time is introduced. A game is played where Nature, oblivious of the state, randomly chooses the disturbance signal and the Player, having access to the random disturbance information, employs a feedback/stroboscopic strategy. Nature's mixed strategy choice need not necessarily be Gaussian, however the first two moments of the random variable are specified; the value of the game is in part determined by the pre-specified first two moments, but not by the p.d.f. chosen by Nature. The closed-form optimal solution is derived for the case where full state information is available to the Player, and also for the case where partial state information is available. In the case where full state information is available, the optimal controller (6) is linear in the state and in the disturbance. The gain for the state is identical to the LQR gain for the state. In the case where partial state information is available, the state information x_k in the optimal controller (6) is replaced by the minimum variance state estimate $\hat{x}_k^+$. The latter is obtained by running a $modified$ Kalman filter, where the prior state vector estimate and prior state's estimation error covariance matrix are propagated according to Eqs. (19) and (20), respectively.

Disclaimer

The views expressed in this article are those of the author and do not reflect the official policy of the United States Air Force, Department of Defense, or the U.S. Goverment.

References

[1] *Cooperative Operations in UrbaN TERrain (COUNTER)*, AFMC 04-197, Approved for public release and obtainable upon request from AFRL/VACA.

[2] Girard A., Pachter M., and Chandler P., *Decision Making Under Uncertainty and Human Operator Model for Small UAV Operations*, AIAA Guidance, Navigation and Control Conference, Keystone, CO, August 21–24, 2006.

[3] Pachter M., *Optimal Sequential Inspection - A Game Against Nature*, presented at the 12th International Symposium on Dynamic Games and Applications, July 3-6, 2006, Sophia Antipolis, France. 2006.

[4] Pachter M., Dharba S., and Chandler P., *Optimal MAV Operations in an Uncertain Environment*, to appear in the International Journal of Nonlinear and Robust Control, 2008.

[5] Hajek O., *Pursuit Games*, Academic Press, New York,(1976).

[6] Maybeck P., *Stochastic Models, Estimation and Control, Vol. 3,* Academic Press, New York (1982), pp. 17.

Is Deterrence Evolutionarily Stable?

David Ellison
Ecole Polytechnique
91128 Palaiseau Cedex France
david@ellison.net

Michel Rudnianski
CEREM &IIM (CNAM)
2 rue Conté, Paris 75003
michel.rudnianski@cnam.fr

Abstract

Deterrent behaviors are currently observed among animals. Similarly, deterrence has spread in human societies long before the nuclear era. That seems to speak for deterrence's efficiency, or to state things in the language of evolution, for its fitness. The literature contains many case studies about the effect of deterrence on evolution. This paper aims at extending these studies, providing a general framework for analyzing the fitness of deterrent behaviors through the combination of two approaches: Evolutionary Games and Games of Deterrence.

Key words. deterrence, playability, positive playability, playability dynamics, stability, asymptotic stability, evolutionary stability.

AMS Subject Classifications. 91A22, 91A43, 91A80.

1 Introduction

In this paper, we shall use a core element of Evolutionary Games theory, the Replicator Dynamic, which describes the quantitative evolution of several species inside a given population. In the standard Replicator Dynamic [4], interactions between species at a given instant are modelled as a two player quantitative matrix game (Nash game), in which each species is a strategy. The Replicator Dynamic then associates with the game a dynamic system representing the population's evolution, and analyzes the existence and properties of these dynamic system equilibria, especially *evolutionary stable equilibria*.

Nash games will be replaced here by qualitative games called *Games of Deterrence*. These games designed to specifically deal with threshold values (a central concept in the analysis of deterrence), consider only two kinds of outcomes: *acceptable* (noted 1) and *unacceptable* (noted 0) ones.

It has been shown [9] that each matrix game of deterrence can be associated with a bipartite graph, called *graph of deterrence*, and that all graphs of deterrence can

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_18,

be gathered in seven types, such that the game solution set is defined by the type. Games of deterrence have been initially developed by Rudnianski [9], in order to account for threshold issues at the heart of the French strategic doctrine of nuclear deterrence theorized as the Relation from the Weak to the Strong, and which aimed at deterring a potential enemy by acquiring the capacity of inflicting the latter casualties which would be over the threshold of acceptability for him. The game theoretic tools developed in this framework proved later to have potential applications in many different fields, some times quite far from nuclear strategy. For instance:
- International conflicts and negotiations [10,11,17,18,19]
- Prevention of congestion in digital communication networks and road traffic [3]
- Business Process Reengineering [1,15]
- Multi-Criteria Decision Making (with particular examples pertaining to the choice of a business strategy, and results which under a given set of conditions display consistency with the Cournot solution, and highlight the Condorcet Paradox [13]
- Valuation of inference chains in the framework of propositional logic, and argumentation [12,16], etc.

The present paper proposes a first exploration of the relation between deterrence and evolution.

Its first part will recall the basic properties of the Replicator Dynamic and matrix Games of Deterrence. In particular, a typology of games of deterrence will be proposed, that leans on graphs (called graphs of deterrence) associated with the games. A second part will focus on the particular case of symmetric games of deterrence that will be used in the Replicator Dynamic, and for which the number of game types reduces to three.

We shall thus have set up a firm ground to explore the relationship between the type of game and evolution. More specifically, the paper will give conditions for which the proportion of non playable strategies vanishes with time. These conditions will then be used to analyze games associated with particular graphs of deterrence.

2 Fundamental properties of the Replicator Dynamic

Given a *population* comprised of *n species,* let us associate with the interaction between two individuals belonging to different species a pair of outcomes describing the interaction's positive or negative impact on each of the individuals. The interaction's outcome will be measured by its consequence on the species' reproduction.

Frequency of interactions involving a species depends on the proportion of the species' members in the population.

Definition 1

Given:
- *the individual outcome* u_{iq} *of an individual of species i after interacting with an*

individual of species q
- the proportions $\theta_1, ..., \theta_n$ *of the various species in the total population, let:*
1. ***the average outcome of the species i*** $u_i = \sum_{q=1}^{n} \theta_q u_{iq}$
2. ***the average outcome of the population*** $u_T = \sum_{i=1}^{n} \theta_i u_i = \sum_{i=1}^{n} \sum_{q=1}^{n} \theta_i \theta_q u_{iq}$

u_T, u_i, and θ_i are functions of time. But for the sake of simplicity, we shall not mention the variable t explicitly, unless necessary.

A species' growth depends on its own outcome, while the evolution of its proportion in the population depends on its relative outcome $u_i - u_T$.
Thus, the fundamental equation of the Replicator Dynamic considers that for every species i:
$\theta'_i = \theta_i(u_i - u_T)$.

It can be noticed that the sum of the θ'_i equals 0, and that the definition here above is consistent with the fact that the sum of proportions equals 1.

Definition 2

1. ***The population's profile*** *is the vector* $\theta = (\theta_1, ..., \theta_n)$.

2. ***The outcome of a mixed strategy*** θ *after interacting with a mixed strategy* $\overline{\theta}$ *is* $u(\theta, \overline{\theta}) = \sum_{i=1}^{n} \sum_{q=1}^{n} \theta_i \overline{\theta}_q u_{iq}$.

3. θ *is* ***an evolutionary stable strategy (ESS)*** *if for any other profile* $\overline{\theta}$*:*
- either $u(\theta, \theta) > u(\overline{\theta}, \theta)$
- or $u(\theta, \theta) = u(\overline{\theta}, \theta)$ *and* $u(\theta, \overline{\theta}) > u(\overline{\theta}, \overline{\theta})$.

3 Games of deterrence

3.1 Playability and deterrence

In everyday life, a number of decisions are based, not on a rigorous optimization process, difficult to implement, but rather on a dichotomization of the outcome space between acceptable outcomes (noted 1) and unacceptable ones (noted 0). Games of deterrence have been specifically developed to represent such situations, and analyze the strategies' playability.

For the sake of simplicity, in the following, we shall only consider two-player matrix games, and recall the basic definitions and properties presented in [9].

The definitions hereafter extend straightforwardly to N-player games.

Thus, let us consider a matrix game with the two players E and R having respective strategic sets S_E (card $S_E = n$) and S_R (card $S_R = p$)

Definition 3

Let $(u_{er}, v_{er}) \in \{0,1\} \times \{0,1\}$ *be the outcome pair associated with the strategic pair* $(e,r) \in S_E \times S_R$.
If $u_{er} = 1, u_{er}$ *is said to be an* ***acceptable outcome*** *for player E.*
If on the opposite $u_{er} = 0, u_{er}$ *is said to be an* ***unacceptable outcome*** *for player E.*
Similar definitions apply by analogy to player R with matrix V.

Definition 4

A strategy e of E is ***safe*** *iff* $\forall r \in S_R, u_{er} = 1$.

Definition 5

1. Let us associate with each strategy $s \in S_E \times S_R$ *an index J(s), and with each player P an index* j_p *such that given* $e \in S_E$*:*
- If e is safe then $J(e) = 1$.
- If e is not safe, $J(e) = \prod_{r \in S_R} [1 - (1 - u_{er})J(r)](1 - j_E)(1 - j_R)$,
with $j_E = \prod_{e \in S_E} (1 - J(e))$ *;* $j_R = \prod_{r \in S_R} (1 - J(r))$.
J(s) and j_P *will be called* ***index of positive playability*** *of strategy s, and* ***index of playability by default*** *of player P, respectively.*

2. If $J(e) = 1$*, strategy* $e \in S_E$ *is said to be* ***positively playable****.*

3. If E has no positively playable strategy, (i.e., if $j_E = 1$*), all strategies* $e \in S_E$ *are said to be* ***playable by default****.*
Similar definitions apply by analogy to strategies r of S_R *.*

4. A strategy $s \in S_E \cup S_R$ *is* ***playable*** *iff it is either positively playable or playable by default.*

Definition 6

1. The system ***S*** *of all J(e),* $e \in S_E$*, and J(r),* $r \in S_R$*, is called the* ***playability system*** *of the game.*

2. A ***solution*** *of* ***S*** *is a consistent set* $\{J(e_1), J(e_2), ..., J(e_n), J(r_1), J(r_2), ..., J(r_p)\}$.
In general, there is no uniqueness of the solution.

Definition 7

1. A structure $(S_E, S_R, U, V, \boldsymbol{S})$ *is called* ***matrix game of deterrence.***

2. A strategic pair $(e, r) \in S_E \times S_R$ *is said to be an* ***equilibrium*** *of the game of deterrence if both strategies are playable for some solution of the playability system.*

Definition 8

Given a strategic pair $(e, r) \in S_E \times S_R$*, e is termed* ***deterrent strategy*** *vis-à-vis r iff:*
(1) e is playable
(2) $v_{re} = 0$
(3) $\exists r_i \in S_R : J(r_i) = 1.$

It has been shown [9] that a strategy $r \in S_R$ is playable iff there is no strategy $e \in S_E$ deterrent vis-à-vis r. Thus, the study of deterrence properties amounts to analyzing the strategies' playability properties.

Example 1

	r_1	r_2	r_3
e_1	(0,0)	(1,1)	(1,1)
e_2	(1,1)	(1,0)	(0,0)

The playability system writes:
$J(e_1) = [1 - J(r_1)].[1 - j_E].[1 - j_R]$
$J(e_2) = [1 - J(r_3)].[1 - j_E].[1 - j_R]$
$j_E = [1 - J(e_1)].[1 - J(e_2)]$
$J(r_1) = [1 - J(e_1)].[1 - j_E].[1 - j_R]$
$J(r_2) = J(r_3) = [1 - J(e_2)].[1 - j_E].[1 - j_R]$
$j_R = [1 - J(r_1)].[1 - J(r_2)].[1 - J(r_3)]$

One can easily establish that the playability system has three solutions:
- $\{J(e_1) = 1; J(e_2) = 0; J(r_1) = 0; J(r_2) = 1; J(r_3) = 1\}$
- $\{J(e_1) = 0; J(e_2) = 1; J(r_1) = 1; J(r_2) = 0; J(r_3) = 0\}$
- all strategies are playable by default.

In the first case, e_1 is deterrent vis-à-vis r_1, and r_3 is deterrent vis-à-vis e_2.
In the second case, e_2 is deterrent vis-à-vis r_2 and r_3, while r_1 is deterrent vis-à-vis e_1.
In the third case, there is no relation of deterrence.

3.2 Graphs of deterrence

Definition 9

*Given a game of deterrence (S_E, S_R, U, V,**S**):*
*1. A **graph of deterrence** is a bipartite graph G on $S_E \times S_R$ such that, given $(e, r) \in S_E \times S_R$, there is an arc of origin e (resp. (r) and extremity r (resp. e), iff $v_{er} = 0$ (resp. $u_{er} = 0$).*

*2. An **E-path** (resp. **R-path**) is a path the root of which is a safe strategy of S_E (resp. of S_R).*

*3. A **C-graph** is a graph that includes neither and E-path, nor an R-path.*

It has been shown [9] that:
1. If G is an E-path (resp. R-path), the only positively playable strategy for E (resp. R) is the root, which is a safe strategy, while all strategies of R (resp. of E) are playable by default.

2. If G is a C-graph, a solution of **S** satisfies the following properties:
i) for any strategy s_0,
$$J(s_0) = (1 - j_E)(1 - j_R)(1 - \prod_{s \in N(s_0)} (1 - J(s)) \prod_{s' \in N'(s_0)} (1 - J(s')),$$
where $N(s_0)$ (resp. $N'(s_0)$) is the set of the first strategies met when following G backward from s_0, and belonging to the same strategic set as s_0 (resp. to the other);
ii) on a path, the positive playability of a vertex is determined by the parity of its distance to the origin of the path;
iii) each player has at least one non positively playable strategy.

3. Through appropriate cuts, it is always possible to break down G into connected parts, each one being an E-path, R-path, or C-graph.

Hence, depending on the presence of these components in the graph, one can distinguish between 7 types of games : E, R, C, E-R, E-C, R-C, and E-R-C. This typology leads in turn to the following.

<u>Classification Theorem</u> [9]
(1) Given a game of deterrence, its playability system's solution set is not empty.
(2) The game type defines the properties of the solution set.

Thus, in particular:
a) if the game is of type E or R, on each path all strategies of odd rank except the root are not playable, while all strategies of even rank are playable by default ;
b) if the game is of type E-R or E-R-C, on each path strategies of odd rank are positively playable while strategies of even rank are not playable.

It stems straightforwardly that in games of type E, R, or E-R, the playability system displays a unique solution.
Also, games of type E-R-C have at least two solutions: in the component of type C of the graph, all E's strategies are positively playable and all R's strategies are not playable, and vice-versa

It follows from (1) that every game of deterrence has an equilibrium, but the above shows that this equilibrium may not be unique.

In the above example, the graph of deterrence associated with the game matrix is:

$$e_1 \rightleftharpoons r_1$$
$$r_3 \rightleftharpoons e_2 \rightarrow r_2$$

4 Symmetric games of deterrence

Definition 10

A symmetric matrix game is a game in which both players have the same strategic set S, and such that whatever (i,j) $\in S^2$, if the outcome pair associated with the strategic pair (i,j) is (a,b), then the outcome pair associated with the strategic pair (j,i) is (b,a).

As in Sec. 2, we shall focus on a single population comprised of n species, and consider only two-player games, in which the (abstract) players have the same strategic set which is precisely the set of species. Furthermore, given a pair of species (i,j) the outcome for species i of an interaction with species j should not depend on which player chose strategy i, it follows that the games considered will be symmetric.

It should be noticed that in case of multiple solutions of the playability system, a symmetric game may display asymmetric solutions: for a given solution, the same strategy may be playable for one player, and not playable for the other.

Proposition 1

In a symmetric game of deterrence $j_R = j_E$.

Proof. If player E's strategies are playable by default, the latter has no safe strategy. By reason of symmetry, the same goes for player R. In other words, for each strategy j of R, there is a strategy i of E such that for R the outcome associated with the strategic pair (i,j) is 0. As by assumption, i is playable, j cannot be positively playable. Consequently, strategies of player R are also playable by default.

Proposition 2

In a symmetric game, the solution set is comprised of:

(1) either symmetric solutions, that is solutions in which a given strategy i has the same playability properties for both players
(2) or pairs of asymmetric solutions, such that if a strategy i has a given playability property in one solution of the pair for one player, it has the same playability property for the other player in the other solution of the pair

Proof. The conclusion stems straightforwardly from the symmetric features of the game.

It follows from definition 10 that a symmetric game can only be of type C, E-R or E-R-C.

5 Evolution and playability: general properties

Unless otherwise specified, in this section, all games considered are symmetric games of deterrence associated with the Replicator Dynamic.

Definition 11

A strategy i is said to be isotropic if $\forall j \neq i, u_{ji} = a$ *(a = 0 or 1).*

Proposition 3

If a game has an isotropic strategy i, then the respective evolutions of the remaining strategies do not depend on the proportion of i in the population.

Proof. Let $(j,k) \in (S - \{i\})^2$,

$$(\frac{\theta_j}{\theta_k})' = \frac{\theta_j'\theta_k - \theta_j\theta_k'}{\theta_k^2} = \frac{\theta_j(u_j - u_T)\theta_k - \theta_k(u_k - u_T)\theta_j}{\theta_k^2}$$
$$= \frac{\theta_j}{\theta_k}(u_j - u_k) = \frac{\theta_j}{\theta_k}\sum_{m=1}^{n}\theta_m(u_{jm} - u_{km})$$

It then stems from the isotropy of i that $(\frac{\theta_j}{\theta_k})' = \frac{\theta_j}{\theta_k}\sum_{m \neq i}\theta_m(u_{jm} - u_{km})$

In other words, $(\frac{\theta_j}{\theta_k})'$ does not depend on θ_i.

Proposition 4

Consider a game with a symmetric solution including non playable by default strategies.
Let :
- $\{1,...,k-1\}$ be the set of positively playable strategies (for both players)
- $\{k,...,n\}$ be the set of non playable strategies (for both players)
- $\Theta_k(t) = \sum_{j \geq k}\theta_j(t)$
If at initial time, for every species $i < k$, the proportion $\theta_i(0)$ of i in the population is strictly superior to the sum of proportions $\Theta_k(0)$, then:

(1) $\forall i < k, \lim_{t \to +\infty} \theta_i(t) = l_i \neq 0$

(2) $\forall j \geq k, \lim_{t \to +\infty} \theta_j(t) = 0$

Proof.

1) Let:

$\theta_m(t) = \min_{i<k} \theta_i(t)$ and $\epsilon(0) = \theta_m(0) - \Theta_k(0)$

By assumption, $\epsilon(0) > 0$.

$\forall t \in [0, \infty[, \forall (i, i') \in \{1, ..., k-1\}^2, u_{ii'} = 1$

It follows that for $i < k, u_i(t) = \sum_{q=1}^{n} \theta_q(t) u_{iq} \geq \sum_{q=1}^{k-1} \theta_q(t) = 1 - \Theta_k(t)$.

Hence, $u_i(t) \geq 1 - \Theta_k(t)$ (1)

For $j \geq k, \exists i' < k$ such that $u_{ji'} = 0$.

Now $u_j(t) = \sum_{q=1}^{n} \theta_q(t) u_{jq} = \sum_{q \neq i'} \theta_q(t) u_{jq} \leq \sum_{q \neq i'} \theta_q(t) = 1 - \theta_{i'}(t)$.

Hence, $u_j(t) \leq 1 - \theta_m(t)$. (2)

(a) $\forall q \in \{1, ..., n\}, \theta_q$ is C^∞.

Hence, θ_m is piecewise C^∞, and has a left and right derivative at the points where it is discontinuous.

Let $(\theta'_m)^+(t)$ be the right derivative of $\theta_m(t)$.

There exists $i < k$ such that $(\theta'_m)^+(t) = \theta'_i(t) = \theta_i(t)(u_i(t) - u_T(t))$

It then stems from definition of $\theta_m(t)$ and inequality (1) that

$\theta'_i(t) \geq \theta_m(t)(1 - \Theta_k(t) - u_T(t))$.

Consequently, $(\theta'_m)^+(t) \geq \theta_m(t)(1 - \Theta_k(t) - u_T(t))$. (3)

Moreover, $\Theta'_k(t) = \sum_{j \geq k} \theta'_j(t) = \sum_{j \geq k} \theta_j(t)(u_j(t) - u_T(t))$.

Hence, it stems from (2) that:

$\Theta'_k(t) \leq \sum_{j \geq k} \theta_j(t)(1 - \theta_m(t) - u_T(t)) = \Theta_k(t)(1 - \theta_m(t) - u_T(t))$ (4)

(3) and (4) then imply :

$$(\theta'_m)^+(t) - \Theta'_k(t) \geq \theta_m(t)(1 - \Theta_k(t) - u_T(t)) - \Theta_k(t)(1 - \theta_m(t) - u_T(t))$$
$$\geq (\theta_m(t) - \Theta_k(t))(1 - u_T(t))$$

Consequently, if $(\theta_m - \Theta_k)$ is positive on a time interval I, then $(\theta_m - \Theta_k)$ is an increasing function over I.

It then stems from the assumption according to which $(\theta_m - \Theta_k)(0) = \epsilon(0) > 0$, that $\forall t > 0, \theta_m(t) - \Theta_k(t) \geq \epsilon(0)$ (5)

(b) $\Theta_k' = \sum\limits_{j \geq k} \theta_j u_j - \Theta_k u_T = \sum\limits_{j \geq k} \theta_j u_j - \Theta_k \sum\limits_{q=1}^{n} \theta_q u_q$
$\Theta_k' = (1 - \Theta_k) \sum\limits_{j \geq k} \theta_j u_j - \Theta_k \sum\limits_{i<k} \theta_i u_i$

It then stems from (1) and (2) that:
$\Theta_k' \leq \Theta_k(1 - \Theta_k)(1 - \theta_m) - \Theta_k(1 - \Theta_k)^2 = \Theta_k(1 - \Theta_k)(\Theta_k - \theta_m)$
Hence, (5) implies that $\Theta_k' \leq -\epsilon(0)\Theta_k(1 - \Theta_k)$.
On the whole, Θ_k decreases exponentially towards 0,
which implies that:
- $\forall j \geq k, \lim\limits_{t \to \infty} \theta_j = 0$
- Θ_k can be integrated over $[0, \infty[$

(2) By definition, $\forall (i, i') \in \{1, ..., k-1\}^2, u_{ii'} = 1$
Furthermore, as $\sum\limits_{i=1}^{k-1} \theta_i(t) = 1 - \Theta_k(t)$,
$u_T = \sum\limits_{q=1}^{n} \sum\limits_{q'=1}^{n} \theta_q \theta_{q'} u_{qq'} \geq \sum\limits_{i=1}^{k-1} \sum\limits_{i'=1}^{k-1} \theta_i \theta_{i'} u_{ii'} = (1 - \Theta_k)^2$
Whence $(1 - \Theta_k)^2 \leq u_T \leq 1$.

Now, given $i < k$, it stems from (1) that $(1 - \Theta_k) \leq u_i \leq 1$.
Hence, $|u_i - u_T| \leq 1 - (1 - \Theta_k)^2$.
$|\frac{\theta_i'}{\theta_i}| \leq 2\Theta_k - \Theta_k^2 \leq 2\Theta_k$
Now it stems from b), that Θ_k can be integrated.
Hence, θ_i'/θ_i can be integrated over $[0, \infty[$, and $\lim\limits_{t \to \infty} ln(\theta_i)$ is finite.

It follows that $\lim\limits_{t \to \infty} \theta_i \neq 0$.

Interpretation
Let:
- $\theta = (\theta_1, \theta_2, ..., \theta_n)$
- A = $\{\theta(0)$ / $\theta_j(0) = 0 \Leftrightarrow$ j is not playable$\}$
- B = $\{\theta(0)$ / i is positively playable $\Rightarrow \theta_i(0) > \Theta_k(0)\}$

A is a set of equilibria, because the outcomes of all species associated with a positively playable strategy are equal.
B is a neighbourhood of A.
According to Proposition 4, B is included in the reunion of the attraction basins of A's equilibria.
Therefore, this reunion is a neighbourhood of A.
This property is similar to the definition of asymptotic stability. If we extend this definition to include the asymptotic stability of sets of equilibria, then we can say that A is asymptotically stable.

Example 2

Let us consider a population in which individuals may adopt three possible types of behaviours:
- aggressive (a)
- defensive (d)
- passive (p)

Let us furthermore assume the following:
- Whenever two individuals of the same type interact, the outcome for each one is 1, which means between other things that an aggressive individual will not try to attack another aggressive individual (maybe because of the fear of the outcome)
- A defensive type when meeting an aggressive individual will respond to the aggression and inflict damages to the aggressor, with the consequent that the outcome pair will be (0,0)
- When meeting a defensive or a passive type, the defensive type does not attack and the outcome pair is (1,1)
- A passive type never responds, and gets a 0 when attacked, and 1 otherwise.

Under the above assumptions the game can be represented by the following matrix:

	a	d	p
a	(1,1)	(0,0)	(1,0)
d	(0,0)	(1,1)	(1,1)
p	(0,1)	(1,1)	(1,1)

One can easily establish that the game of deterrence has three solutions:
1) all strategies are playable by default
2) a is positively playable, while d and p are not playable
3) d and p are playable, while a is not

Proposition 4 applies to the last two solutions:
It tells us that if at initial time, more than half of the population is aggressive, then the whole population becomes aggressive, whereas if the aggressive population is the smallest of the three, it then disappears.

It stems straightforwardly from the game structure that the profile $(1, 0, 0)$ (corresponding to the whole population playing a), is an evolutionary stable strategy.
However, the profiles $(0, x, 1 - x)$ with $0 < x < 1$ (corresponding to the whole population playing d or (p) are not evolutionarily stable strategies.

Proposition 5

If the Replicator Dynamic has an equilibrium $(\theta_i^E)_{1\leq i\leq n}$ such that all θ_i^E do not equal 0, and if the playability system has a symmetrical solution including only one playable strategy, then $(\theta_i^E)_{1\leq i\leq n}$ is an unstable equilibrium.

Proof. Suppose that 1 is the only playable strategy. Then 1 is positively playable and deterrent vis-à-vis all other strategies, which implies in turn that 1 is isotropic. It stems from the definition of an equilibrium that at the equilibrium point E, there is u^E such that $\forall i, u_i^E = u_T^E = u^E$.
One can move away from E by changing:
- θ_1^E into $\theta_1(0) = \theta_1^E + \epsilon(0)$
- θ_k^E into $\theta_k(0) = (1 - \frac{\epsilon(0)}{1-\theta_1^E})\theta_k^E, \forall k \geq 2$.

It then stems from Proposition 3 that the relative proportion of the $\theta_k, k \geq 2$, remains unchanged.
The vector $(\theta_i)_{1\leq i\leq n}$ then evolves along a straight line of $\mathbf{R}^n$, and the parameter $\epsilon = \epsilon(t)$ suffices to describe the evolution process.
$\forall k \geq 2, \forall t, \theta_k(t) = (1 - \frac{\epsilon(t)}{1-\theta_1^E})\theta_k^E$

$$
\begin{aligned}
u_1(t) = \sum_{k=1}^{n} \theta_k(t)u_{1k} &= u^E + \epsilon(t) - \frac{\epsilon(t)}{1-\theta_1^E} \sum_{k\geq 2} \theta_k^E u_{1k} \\
&= u^E + \epsilon(t) + \frac{\epsilon(t)}{1-\theta_1^E}\theta_1^E - \frac{\epsilon(t)}{1-\theta_1^E} u^E \\
&= (1 - \frac{\epsilon(t)}{1-\theta_1^E})u^E + \frac{\epsilon(t)}{1-\theta_1^E}
\end{aligned}
$$

$\forall i \geq 2$,

$$u_i(t) = \sum_{k=1}^{n} \theta_k(t)u_{ik} = u^E - \frac{\epsilon(t)}{1-\theta_1^E} \sum_{k\geq 2} \theta_k^E u_{ik} = u^E - \frac{\epsilon(t)}{1-\theta_1^E} u^E$$

$$u_T(t) = \sum_{k=1}^{n} \theta_k(t)u_k(t) = (1 - \frac{\epsilon(t)}{1-\theta_1^E})u^E + \frac{\epsilon(t)}{1-\theta_1^E}\theta_1^E + \frac{\epsilon^2(t)}{1-\theta_1^E}$$

$$\theta_1'(t) = (\epsilon(t) - \frac{\epsilon^2(t)}{1-\theta_1^E})\theta_1(t)$$

$$\epsilon'(t) = (\epsilon(t) - \frac{\epsilon^2(t)}{1-\theta_1^E})(\theta_1^E + \epsilon(t))$$

$\epsilon(t)$ satisfies an autonomous equation describing a system with three equilibria: $0, 1 - \theta_1^E$ and $-\theta_1^E$, where 0 is unstable.
Hence, the equilibrium $(\theta_i^E)_{1\leq i\leq n}$ is unstable.

6 Evolution and game type

Proposition 6

In a game of type E-R or E-R-C, $\forall i \in \{1, ..., n\}, \theta_i$ has a limit when t tends toward the infinite.

Proof. Let i be a safe strategy and j any strategy.
θ_i is increasing and $\theta_i \leq 1$.

Hence, θ_i has a limit a when t tends toward the infinite.
$(\frac{\theta_j}{\theta_i})' = \frac{\theta_j(u_j - u_T)\theta_i - \theta_j\theta_i(1-u_T)}{\theta_i^2} = \frac{\theta_j}{\theta_i}(u_j - 1) < 0.$
Hence, θ_j/θ_i is decreasing and strictly positive.
Hence, θ_j/θ_i has a limit b when t tends toward the infinite,
whence, $\lim\limits_{t\to\infty} \theta_j = ab.$

Proposition 7

In a game of type E-R or E-R-C, the proportions of strategies belonging to an E-path or an R-path have a:
- zero limit when they are not playable
- non zero limit when they are positively playable.

Proof. The idea of the proof is to follow the E-paths and R-paths with a recurrence based on the safe strategies, and show that the proportions of the strategies of even rank decrease exponentially towards 0. In order to show that all the strategies belonging to the E-paths and R-path have been reached, we will prove that the remaining strategies belong to a C-graph by considering the solutions of the playability system and using the classification theorem.

Let us consider a game of type E-R or E-R-C.
Let S be the strategic set common to both players, and θ a solution of the Replicator Dynamic.
It stems from proposition 6 that $\theta(t)$ tends towards an equilibrium θ^E when t tends towards the infinite.

(a) At the equilibrium point, outcomes of all strategies, the proportion of which in the population is non zero, are equal.
Moreover, the proportions of safe strategies being increasing, their limit is not 0.
Therefore, at equilibrium, outcomes of all strategies, the proportion of which in the population is not zero, equal 1.
Hence, $\lim\limits_{t\to\infty} u_T = 1.$

(b) Let i and j be two strategies such that $u_{ij} = 0$ and $\theta_j^E \neq 0$
$\theta_i'/\theta_i = u_i - u_T \leq 1 - \theta_j - u_T.$
Now, it stems from (a) that $\lim\limits_{t\to\infty}(1 - \theta_j - u_T) = -\theta_j^E.$

Hence, θ_i decreases exponentially, $\theta_i^E = 0$ and θ_i can be integrated over $[0, \infty[$.

(c) Let i be a strategy such that $\forall j \in S, u_{ij} = 0 \Rightarrow \theta_j$ can be integrated
Let $D_i = \{j \in S/u_{ij} = 0\}$ and $\theta_{D_i} = \sum\limits_{j\in D_i} \theta_j$
$\theta_i'/\theta_i = u_i - u_T = 1 - u_T - \theta_{D_i}$
Hence, $(ln(\theta_i))' \geq -\theta_{D_i}.$

Now $-\theta_{D_i}$ can be integrated over $[0, \infty[$.
Therefore, $\lim_{t\to\infty} ln(\theta_i) > -\infty$.
Hence, $\theta_i^E > 0$.

(d) Consider the sequences of sets (X_k) and (Y_k) such that:
X_1 = {i / i is safe}
$Y_1 = \{j/\exists i \in X_1 : u_{ji} = 0\}$
And more generally, for any $k \geq 2$,
$X_k = \{i/u_{ij} = 0 \Rightarrow j \in Y_{k-1}\}$
$Y_k = \{j/\exists i \in X_k : u_{ji} = 0\}$

By construction, the sequences (X_k) and (Y_k) are strictly increasing for the relation of inclusion, until k reaches a value m from which these sequences remain constant. Let $X = \bigcup_{k=1}^{\infty} X_k = X_m$ and $Y = \bigcup_{k=1}^{\infty} Y_k = Y_m$.

$\forall i \in X_1$, i is positively playable, hence $\forall j \in Y_1$, j is not playable.
If $\forall i \in X_k$, i is positively playable, then $\forall j \in Y_k$, j is not playable.
If $\forall j \in Y_k$, j is not playable, then $\forall i \in X_{k+1}$, i is positively playable.
Hence, by recurrence, $\forall i \in X$, i is positively playable, and $\forall j \in Y$, j is not playable.

Let $Z = S - (X \cup Y)$.
If Z is empty, the playability system has a unique solution for which no strategy is playable by default. According to the classification theorem, the game is then of type E-R.

Assume now that Z is not empty, and let $z \in Z$.
$z \notin Y$, hence, by definition of $Y, \forall i \in X, u_{zi} = 1$
$z \notin X$, hence by definition of $X, \exists z' \in S - Y : u_{zz'} = 0$
Hence, $\exists z' \in Z : u_{zz'} = 0$.
Thus, $\forall i \in X, \forall k \in X \cup Z, u_{ik} = u_{ki} = 1$
and $\forall j \in Y, \exists i \in X : u_{ji} = 0$.
and $\forall j \in Y \cup Z, \exists i \in X \cup Z : u_{ji} = 0$

Hence, the playability system has at least two solutions:

- i is positively playable for player E $\Leftrightarrow i \in X$
  and j is positively playable for player R $\Leftrightarrow i \in X \cup Z$
- i is positively playable for player E $\Leftrightarrow i \in X \cup Z$
  and j is positively playable for player R $\Leftrightarrow i \in X$

Hence, the game is of type E-R-C, the sub-graph associated with $X \cup Y$ being the E-R component of the game, and the sub-graph associated with Z being a C-graph.

By using (b) and c), one shows by recurrence that:
- $\forall i \in X, \theta_i^E > 0$
- $\forall j \in Y, \theta_j^E = 0$ and θ_j is integrable.

Consequences:
- In a symmetric game, the proportion associated with a strategy that is positively playable (resp. non playable) in all solutions of the playability system, has a non-zero (resp. zero) limit when t tends toward the infinite.
- the proportion of the population associated with a strategy being deterred by a safe strategy, tends toward zero when t tends toward the infinite, and is integrable over $[0, \infty[$.

Proposition 8

If the ordered vertices 1,2,...,2p of an E-path or an R-path are such that:
$\theta_2(0) \le \theta_4(0) \le ... \le \theta_{2k-2}(0) \le \theta_{2p}(0) \le \theta_{2p-1}(0) \le \theta_{2p-3}(0) \le ... \le \theta_3(0) \le \theta_1(0)$,
then $\forall t, \theta_2(t) \le \theta_4(t) \le ... \le \theta_{2p-2}(t) \le \theta_{2p}(t) \le \theta_{2p-1}(t) \le \theta_{2p-3}(t) \le ... \le \theta_3(t) \le \theta_1(t)$.

Proof.
$u_1 = 1$, and $\forall i \in [2, 2p], u_i = 1 - \theta_{i-1}$

Let's assume that the result is true at time t:
Then, $\forall k < p, u_{2k}(t) = 1 - \theta_{2k-1}(t) \le u_{2k+2}(t) = 1 - \theta_{2k+1}(t)$.
Hence, $\theta'_{2k}(t) = \theta_{2k}(t)(u_{2k}(t) - u_T(t)) \le \theta_{2k+2}(t)(u_{2k+2}(t) - u_T(t)) = \theta'_{2n+2}(t)$.

Similarly, $u_{2k-1}(t) = 1 - \theta_{2k-2}(t) \ge 1 - \theta_{2k}(t) = u_{2k+1}(t)$.
Hence, $\theta'_{2k-1}(t) \ge \theta'_{2k+1}(t)$
$u_{2p-1}(t) = 1 - \theta_{2p-2}(t) \ge u_{2p}(t) = 1 - \theta_{2p}(t) \ge 1 - \theta_{2p-1}(t)$
Hence, $\theta'_{2p-1}(t) \ge \theta'_{2p}(t)$.
So the proposition is true whatever t.
It applies in particular if at t = 0 all proportions equal 1/(2p)

Example 3
Let us consider a game of type E-R, the matrix of which is:

	i	j	k
i	(1,1)	(1,0)	(1,1)
j	(0,1)	(1,1)	(1,0)
k	(1,1)	(0,1)	(1,1)

The game has only one solution for which i and k are positively playable while j is not playable. It then stems from Proposition 7 that $\lim_{t \to \infty} \theta_j = 0$.

So when $t \to \infty$, the only profiles we have to consider are those of type $(x, 0, 1-x)$.

Here as in example 2, we see that the set of equilibria exhibit stability properties which may categorize them not exactly as evolutionary stable equilibria, but as evolutionary stable equilibrium sets (ESES):
$\forall \theta^1, \theta^2, \theta^3, \theta^4 \in$ ESES, $u(\theta^1, \theta^2) = u(\theta^3, \theta^4)$
and $\forall \overline{\theta} \notin$ ESES, $\exists \theta \in$ ESES :
- either $u(\theta, \theta) > u(\overline{\theta}, \theta)$
- or $u(\theta, \theta) = u(\overline{\theta}, \theta)$ and $u(\theta, \overline{\theta}) > u(\overline{\theta}, \overline{\theta})$.

Now one should not derive from the above propositions and examples that deterrence is a necessary condition for reaching an evolutionary stable equilibrium, as the following example shows:

Example 4
Let us consider the following game borrowed from Gintis [5]:

	i	j	k
i	(1,1)	(1,1)	(0,0)
j	(1,1)	(0,0)	(1,0)
k	(0,0)	(0,1)	(0,0)

One can establish quite easily that i is an ESS. The associated game of deterrence which is of type C, has only one solution, for which all strategies are playable by default. Hence, positive playability is not a necessary condition:

7 Evolution and fuzzy playability

The indices of positive playability that we have been considering until now are binary ones. But the playability system of a game may have solutions in which these indices (as well as the indices of playability by default) may take non binary values. We shall then speak of fuzzy playability. This is the case, for instance, if one considers a game of type E or of type R comprised of a single path.

It has thus been shown [12] that:
$J(2k-1) = \dfrac{v(1+v^{2k-3})}{1+v}$, and $J(2k) = \dfrac{v(1-v^{(2k-2)})}{1+v}$,
where $v = 1 - j_R$.

One can see that the strategies of E have a positive playability index which decreases with k, while on the opposite strategies of R have a positive playability index which increases with k.

So we observe a phenomenon similar to the result of Proposition 8, except, that we consider positive playability indices on the one hand, and proportions of species in the total population on the other. The reason behind this similarity lies in the fact that the Replicator Dynamic is actually a long run process of strategy selection. Those strategies which replicate the most are those which are the best fitted to their environment, while on the opposite the strategies, the proportion of which in the total population decreases or even vanishes with time, are those which are ill fitted or even not fitted at all to their environment. Somehow this selection process tells us how a given strategy is playable in the long run, and therefore there is nothing surprising in the fact that strategies' fitness and playability follow similar trajectories. Of course the graphs of deterrence are not the same, since in the Replicator case, it is comprised of an E-path and a R-path, which precludes having fuzzy playability indices, while in the fuzzy game case, there is only one E-path. But these are just formal differences due to the fact that in the Replicator case, only symmetric games are considered. If we focus on the issue of strategy selection, these differences become irrelevant.

8 Conclusion

The developments above have pointed out a correlation between the existence and properties of equilibria in the Replicator Dynamic on the one hand, and strategies' playability in the associated game of deterrence, on the other. In particular, for a strategy to be positively playable and—even more—deterrent vis-à-vis other strategies, seems to be an important factor for the strategy's survival, and the non playable strategies tend to disappear. At a refined level, the breakdown of games of deterrence plays an important role in establishing some correspondence between properties of equilibrium existence in the Replicator Dynamics and the game type. Strong results have been obtained for E-R and E-R-C type games.

A comparison has been undertaken between the results found in some games of type E-R, and the conclusions obtained in developing the theory of fuzzy games of deterrence. What is left at this level is to see whether one could go beyond rough similarities, and determine a quantitative relationship between species' proportions and strategies' fuzzy playability. The nature of the correspondence in the case of C-type games, seems to be more uncertain.

More generally the developments here above are a tentative work which aims at paving the way for a more general theory linking deterrence and evolutionary stability.

References

[1] H. Bestougeff, M. Rudnianski, *Games of Deterrence and Satisficing Models Applied to Business Process Modeling*, in Proceedings AAAI Spring Symposium, Stanford 1998.

[2] R. Cressmann, *The Stability Concept of Evolutionary Game Theory*, Springer, Berlin, 1992.

[3] M. Rudnianski, P. Fabiani, *On Games of Deterrence*, in Qualitative Reasoning and Decision Technologies, N. Piera Carreté, M.G. Singh, Eds., pp 55-64, CIMNE Barcelona, 1973.

[4] D. Fudenberg, D.K. Levine, *The Theory of Learning in Games*, MIT Press, Cambridge, 1998.

[5] H. Gintis, *Game Theory Evolving*, Princeton University Press, Princeton, 2000.

[6] J. Hoffbauer and K. Sigmund, *Evolutionary Games and Population Dynamics*, Cambridge University Press, Cambridge, 1998.

[7] M. A. Nowak, *Evolutionary Dynamics*, The Belknap Press of Harvard University Press, Cambridge 2006.

[8] Randolph M. Nesse, *Evolution and the Capacity for Commitment*, Ronald Sage Foundation, N.Y., 2001.

[9] M. Rudnianski, *Deterrence Typology and Nuclear Stability, A Game Theoretical Approach*, in Defense Decision Making, R. Avenhaus, H. Karkar, M. Rudnianski, Eds, Springer Verlag, pp 137-168, Heidelberg, 1991.

[10] M. Rudnianski, *Multipolar Deterrence in a Dynamic Environment*, in IEEE Systems, Man and Cybernetics 1995, vol 5, pp 4279 – 4285.

[11] M. Rudnianski, *Nuclear Proliferation and Deterrence Stability*, in Models for the Security Policy in the Post-Cold War Era, R Huber, R. Avenhaus, Eds., pp-153-182, Nomos Verlag, Baden-Baden, 1996.

[12] M. Rudnianski, *Deterrence, Fuzzyness and Causality*, in Proceedings of ISAS 96, pp 473 -480, IIIS, Orlando 1996.

[13] M. Rudnianski, *Games of Deterrence for Multiple Criteria Decision Making*, in Proceedings of ISAS 98, vol 3, N. C. Callaos Ed., pp 633-639, Orlando, 1998.

[14] M. Rudnianski, H. Bestougeff, *Modeling Tasks and Teams through Game Theoretical Agents*,in NASA First Goddard Workshop on Formal Approaches to Agent Based Systems, Lectures Notes in Computer Sciences N° 1871, pp 235-249, Springer Verlag, Berlin 2001.

[15] M. Rudnianski, H. Bestougeff, *Modeling Multi-agent Interactions Through State Transitions in Games of Deterrence* , Proceedings AAAI Spring Symposium, PP 159-162, AAAI Press, Stanford, 2001 .

[16] M. Rudnianski, H. Bestougeff, *Multi-Agent Systems Reliability, Fuzzyness and Deterrence*, in MG Hinchey & al, NASA FAABS 2004, LNAI 3228, pp 41-56, Springer Verlag, Berlin, 2005 .

[17] M. Rudnianski, H. Bestougeff, Bridging Games and Diplomacy, presented at the 2005 American Association for the Advancement of Science Conference, to be published in Formal Models of International Negotiation, R; Avenhaus, G-O Faure, V. Kremeniuk, B. Zartmann, Eds., PIN, Vienna, 2006.

[18] M. Rudnianski, A. d'Assignies, From MAD to MAD, in Strategic Stability in the Post-Cold War World and the Future of Nuclear Disarmament, M.L. Best, J. Hughes-Wilson, A. Piontkowsky, Eds., NATO ASI series, Disarmament Technologies, vol 3, pp 229-259, Kluwer Academic Publishers, 1995.

[19] M. Rudnianski, N. Ojeh, NATO and European Enlargment. A Methodological Perspective, in Defense Analysis for the 21st Century : Issues, Approaches, Models, R. Huber, W. Hoffmann, Eds., Nomos Verlag, Baden-Baden, 1999.

Evolutionarily Robust Strategies: Two Nontrivial Examples and a Theorem

A. J. Shaiju
Laboratoire I3S
UNSA-CNRS
06903 Sophia Antipolis Cedex, France
s.ainikkal@adfa.edu.au

Pierre Bernhard
Laboratory I3S
University of Nice Sophia Antipolis and CNRS
Sophia Antipolis, France
Pierre.Bernhard@polytech.unice.fr

Abstract

We revisit the relationship between evolutionarily stable strategy (ESS) and related topics such as evolutionarily robust strategy (ERS) on the one hand, and stability of the replicator dynamics on the other hand, when the phenotypic set is continuous. The state of the population considered is a measure over the phenotypic set. Thus, topological considerations come into play that make the situation much more difficult than for a finite phenotypic set. As a consequence, the issue of the asymptotic stability of an ERS is not settled at this time. We give one partial new result in that direction.

It has also been noticed in the literature that there is a dearth of concrete examples of mixed ESS or ERS in the literature. Actually none seems to be known if the "kernel" of the game is continuous. We provide two such examples, one a convex combination of two Dirac measures and one family with the Lebesgue measure as an ERS.

Key words. Evolutionary games, ESS, Replicator dynamics, Darwinian dynamics, stability.

AMS Subject Classifications. Primary 91A22; Secondary 91A10, 92D25, 91B02.

1 Introduction

The concept of evolutionarily stable strategy (ESS) as the core concept of Evolutionary Game Theory was introduced by Maynard-Smith and Price in [6]. The very intent of the definition is to imply that in a population in an ESS state, a small sub-population in a different state will die out for lack of fitness in Darwinian dynamics.

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*, Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_19,

A mathematical translation of that intuition is given by the fact that in the case of a finite phenotypic set, an ESS is an asymptotically stable state of the associated replicator dynamics. But analogous results for the case of a continuous phenotypic set are not proved, in general, and this remains an open conjecture.

As a matter of fact, the state of a population is a probability measure over the phenotypic set. Thus, if this set is finite, the state space is isomorphic to $\mathbb{R}^n$, while if the phenotypic set is infinite, say continuous, one ends up with a true measure dynamics. Hence, there are as many concepts of asymptotic stability as there are topologies on the set of measures. We seek results in the weak topology, which seems to be the most significant one as far as the modelization is concerned.

The result is known for an evolutionarily robust strategy (ERS, a strengthening of ESS, strictly stronger in the infinite case) if it is a Dirac measure or if the underlying game is "doubly symmetric". We prove the result for the case of a global ERS strategy.

Note that, in the case of finite phenotype set, both ESS and ERS concepts coincide due to the equivalence of strong and weak topologies. In this paper, we study problems with the infinite phenotype set. Such problems are important in Biology and Economics (See, for instance, [1] and [9]).

A further difficulty in investigating this topic is that there does not seem to be in the literature an example of an ERS which is not a Dirac measure. (This point itself has been raised in the literature.) We provide two very different (families of) such examples, and the analysis that substantiates the claim that they are indeed ERS.

The rest of this paper is organized as follows. Sec. 2 is devoted to present the definitions and some properties of ESS, ERS, and related topics. In Sec. 3, we give a brief description of strong and weak topologies on Δ, the set of all strategies (probability measures on the phenotypic set) which will be useful in subsequent discussions. Sec. 4 begins with the description of replicator dynamics and its stability. At the end of this section, namely in Sec. 4.6, we prove our first result—the weak asymptotic stability of global ERS. Then, in Sec. 5, we present two new (families of) examples of ERS that are not Dirac measures.

2 ESS and related concepts

2.1 Definition of ESS

We investigate an evolutionary game in a classical linear framework. The *pure strategy set* of the underlying evolutionary game is a compact subset K of the Euclidean space $\mathbb{R}^d$. We denote a pure strategy using letters $x, y, z, \ldots$. A *mixed strategy* is therefore a probability measure on the Borel σ-algebra $\mathcal{B} = \mathcal{B}(K)$ of K. The mixed strategy simplex (that is, the set of all probability measures on K) is denoted by Δ. We denote a particular mixed strategy using capital letters $P, Q, R, \ldots$. We are given a *fitness map*, a continuous function $u : K \times K \to \mathbb{R}$. Here, $u(x, y)$ represents the fitness gained by the animal adopting x when compete against another animal adopting y. It will be naturally extended to $K \times \Delta$ and to $\Delta \times \Delta$ as:

$$u(x,R) = \int_K u(x,y)R(\mathrm{d}y)\,, \qquad u(Q,R) = \iint_{K\times K} u(x,y)Q(\mathrm{d}x)R(\mathrm{d}y)\,.$$

Remark 2.1. We shall think of a mixed strategy as a distribution of behaviours among a polymorphic population of animals, each using a fixed pure strategy. It could as well be the probability distribution of the strategies used by every animals of a monomorphic population where all individuals are random players.

We are now ready for the definition of ESS.

Definition 2.1. $P \in \Delta$ is called an evolutionarily stable strategy (ESS), if for any mutant strategy $R \neq P$, there is a $\varepsilon(R) \in (0,1]$ such that:

$$u(P, \varepsilon R + (1-\varepsilon)P) > u(R, \varepsilon R + (1-\varepsilon)P) \qquad \text{for all } 0 < \varepsilon < \varepsilon(R). \quad (1)$$

Note that (1) can also be stated as:

$$\varepsilon[u(P,R) - u(R,R)] + (1-\varepsilon)[u(P,P) - u(R,P)] > 0, \qquad \text{for all } 0 < \varepsilon < \varepsilon(R). \quad (2)$$

The Definition 2.1 is essentially by Taylor & Jonker [11], and, in view of (2), is equivalent to the original definition by Maynard Smith [7]:

$$\forall Q \in \Delta, \quad u(Q,P) \le u(P,P)\,, \quad (3)$$
$$\forall R \neq P, \quad [u(R,P) = u(P,P)] \Rightarrow [u(P,R) > u(R,R)]\,. \quad (4)$$

It is convenient to introduce the *best response* map $\mathbb{R}(P)$ defined as $\mathbb{R}(P) = \{R \mid u(R,P) = \max_Q u(Q,P)\}$. Then, $P \in \Delta$ is an ESS iff

$$P \in \mathbb{R}(P) \text{ and } \forall R \in \mathbb{R}(P) \setminus \{P\},\ u(R,R) < u(P,R). \quad (5)$$

It shall be convenient to introduce the notation, for any P and Q in Δ,

$$\begin{aligned} \sigma(x,P) &:= u(x,P) - u(P,P)\,, \\ \sigma(Q,P) &:= \textstyle\int_K \sigma(x,P)Q(\mathrm{d}x) = u(Q,P) - u(P,P)\,, \end{aligned} \quad (6)$$

so that condition (3) above writes $\sigma(Q,P) \le 0$, $\forall Q \in \Delta$.

Any P satisfying that condition, i.e., $P \in \mathbb{R}(P)$, is said to be a *symmetric Nash equilibrium*, and the set of symmetric Nash strategies is denoted by Δ^{NE}.

Remark 2.2. We note that $P \in \Delta^{NE}$ iff (P,P) is a Nash equilibrium of the two-person game with pure strategy set K,and payoff functions $u_1(x,y) = u(x,y)$, $u_2(x,y) = u(y,x)$.

P is a *strict symmetric Nash equilibrium* if $\mathbb{B}\mathbb{R}(P) = \{P\}$. The set of such strategies is denoted by Δ^{SNE}. The set of ESS strategies is denoted by Δ^{ESS}. From (5), it now follows that

$$\Delta^{SNE} \subset \Delta^{ESS} \subset \Delta^{NE}. \tag{7}$$

We can also show, as in the finite case, that there always exists a symmetric Nash equilibrium; that is, $\Delta^{NE} \neq \emptyset$. But Δ^{ESS} can be empty, as the next example illustrates.

2.1.01 Example $K = [-1, 1]$, $u(x, y) = x^2 y$. For any $x \in K$ and $R \in \Delta$,

$$u(x, R) = x^2[R],$$

where $[R] := \int_K y\, R(dy)$ is the average value of R. Now

$$\begin{aligned}
\Delta^{NE} &= \Big\{R \in \Delta \,:\, R \in \mathbb{B}\mathbb{R}(R)\Big\} \\
&= \Big\{R \in \Delta \,:\, u(x, R) \leq u(R, R) \quad \text{for all } -1 \leq x \leq 1\Big\} \\
&= \Big\{R \in \Delta \,:\, x^2[R] \leq [R] \int_{-1}^{1} z^2\, R(dz), \quad \text{for all } -1 \leq x \leq 1\Big\} \\
&= \Big\{R \in \Delta \,:\, x^2[R] \leq [R] \int_{-1}^{1} z^2\, R(dz), \quad \text{for all } -1 \leq x \leq 1\Big\}.
\end{aligned}$$

This suggests that we can write Δ^{NE} as the union of two disjoint sets Δ_1^{NE} and Δ_2^{NE}; where

$$\Delta_1^{NE} = \Big\{R \in \Delta \,:\, [R] = 0\Big\},$$

and

$$\begin{aligned}
\Delta_2^{NE} &= \Big\{R \in \Delta \,:\, [R] > 0 \text{ and } \int_{-1}^{1} z^2\, R(dz) = 1\Big\} \\
&= \Big\{R \in \Delta \,:\, R = \alpha\delta_{-1} + (1-\alpha)\delta_1, \quad 0 \leq \alpha < \frac{1}{2}\Big\}.
\end{aligned}$$

Here, δ_x denote the Dirac measure at x.
To prove $\Delta^{ESS} = \emptyset$, in view of (7), it suffices to show that no strategy $P \in \Delta^{NE}$ is an ESS. We have only two cases to consider here. In both cases (that is when $P \in \Delta_1^{NE}$ or $P \in \Delta_2^{NE}$), $\delta_1 \in \mathbb{B}\mathbb{R}(P)$. But $u(\delta_1, \delta_1) = 1 \geq \int_K x^2\, P(dx) = u(P, \delta_1)$. Therefore, $P \notin \Delta^{ESS}$. Hence, no symmetric Nash equilibrium $P \in \Delta^{NE}$ is an ESS, and $\Delta^{ESS} = \emptyset$. □

Nevertheless, there are many games with $\Delta^{ESS} \neq \emptyset$. We give here two almost trivial examples.

(1) $K = [a, b]$, $u(x, y) = f(x) + g(y)$, where f is a continuous function on $[a, b]$ with a strict maximum at x_0, and g is any continuous function.

Now, for any $x \in K$,

$$u(x, \delta_{x_0}) - u(\delta_{x_0}, \delta_{x_0}) = f(x) - f(x_0).$$

Therefore, δ_{x_0} is a strict Nash equilibrium, and hence an ESS. □

(2) $K = [0, 1]$, $u(x, y) = -xy$. Then $\mathbb{BR}(\delta_0) = \Delta$ and for any $R \neq \delta_0$, and

$$u(\delta_0, R) - u(R, R) = [R]^2 > 0.$$

Therefore, δ_0 is an ESS, but not a strict Nash equilibrium. □

2.2 ESS and uninvadability

Definition 2.2. In Definition 2.1, let $\varepsilon_P(R) := \max\{\varepsilon(R)\}$. It is called the *invasion barrier* of P against R

We define a strategy to be *uninvadable* if it has a uniform invasion barrier, i.e., the following.

Definition 2.3. (Vickers & Cannings [12]) A strategy P is *uninvadable* if $\inf_{R \neq P} \varepsilon_P(R) > 0$.

Notation: To simplify the presentation, we use the notation $h_{R,P}(\varepsilon) := u(R, \varepsilon R + (1 - \varepsilon)P) - u(P, \varepsilon R + (1 - \varepsilon)P)$.

An uninvadable strategy is clearly an ESS. However, unlike in the finite case, the converse is not true, as the following example shows.

2.2.02 Example K any compact interval containing 0, $0 < a < b$, $u(x, y) = bxy - ax^4$. Since $u(x, 0) = -ax^4$, δ_0 is a strict Nash equilibrium, and hence an ESS. Now, for any nonzero x in K,

$$\begin{aligned} h_{\delta_x,\delta_0}(\varepsilon) &= \varepsilon u(x, x) + (1 - \varepsilon)u(x, 0) \\ &= \varepsilon(bx^2 - ax^4) + (1 - \varepsilon)(-ax^4) \\ &= (\varepsilon b - ax^2)x^2. \end{aligned}$$

This implies that $\varepsilon_{\delta_0}(\delta_x) = \min(1, \frac{ax^2}{b})$. Clearly, this invasion barrier tends to zero with x, and so δ_0 is not uninvadable. □

2.3 Strong uninvadability and evolutionary robustness

For a fixed P, let $R_\varepsilon := \varepsilon R + (1-\varepsilon)P$. We first observe that $P \in \Delta$ is uninvadable iff there exists $\varepsilon_0 > 0$ such that:

$$\forall R \neq P,\ \forall \varepsilon \in (0, \varepsilon_0), \quad h_{R,P}(\varepsilon) = u(R, R_\varepsilon) - u(P, R_\varepsilon) < 0\,.$$

Notice that $R - P = (1/\varepsilon)(R_\varepsilon - P)$, so that

$$h_{R,P}(\varepsilon) = u(R-P, R_\varepsilon) = \frac{1}{\varepsilon}[u(R_\varepsilon, R_\varepsilon) - u(P, R_\varepsilon)].$$

Therefore, P is uninvadable iff there exists $\varepsilon_0 > 0$ such that:

$$\forall R \neq P,\ \forall \varepsilon \in (0, \varepsilon_0), \quad \sigma(P, R_\varepsilon) := u(P, R_\varepsilon) - u(R_\varepsilon, R_\varepsilon) > 0\,. \tag{8}$$

This condition says that for Q ($Q \neq P$) close to P in some sense,

$$\sigma(P, Q) > 0.$$

In order to make the concept of 'nearness' precise, we need to equip $\Delta = \Delta(K)$ with a topology (say τ). With respect to this topology τ, we can define evolutionary stability as follows.

Definition 2.4. A strategy $P \in \Delta$ is called

- *locally superior* (w.r.t. τ) if there exists a τ-neighborhood G of P such that

$$\sigma(P, Q) > 0 \quad \text{for all } Q \in G \setminus \{P\}\,,$$

- *strongly uninvadable* if it is locally superior w.r.t. the strong (variational) topology (Bomze [2]),
- *evolutionary robust* if it is locally superior w.r.t. the weak topology. (Oechssler and Riedel [9]),
- *globally evolutionarily robust* if $\sigma(P, R) > 0$ for all $R \in \Delta \setminus \{P\}$.

We state the following easy theorem:

Theorem 2.1. *Let $P \in \Delta$. Then*
P evolutionary robust $\Longrightarrow$ P strongly uninvadable $\Longrightarrow$ P is an ESS.

Proof. This is, respectively, because of the fact that a weak neighborhood is a strong neighborhood, and $\varepsilon R + (1-\varepsilon)P \to P$ strongly as $\varepsilon \to 0$. [1] ∎

But the reverse implications are not true, in general. Nevertheless, these reverse implications do hold true in the finite case; that is, when K is a finite set. In this case, strong and weak topologies coincide. Furthermore, in this case, a neighborhood of P consists only of Q of the form $Q = \varepsilon R + (1-\varepsilon)P$. This is not true in the infinite case.

[1] This will be clearer when we define explicitly the strong and weak topologie, in the next section.

3 Strong and weak topologies on Δ

For the sake of completeness, we provide more details regarding strong and weak topologies on Δ. We view Δ as a subset of the linear space $\mathcal{M}$ of all finite signed measures on K.

If $\mathcal{M}$ is equipped with the *variational norm*

$$\|\mu\|_{var} = \sup_{|f|\leq 1,\, f \text{ measurable}} \left| \int_K f(x)\,\mu(dx) \right|,$$

then $\mathcal{M}$ is a Banach space. The topology generated by this norm on $\mathcal{M}$, is referred to as the *strong topology*.

We denote $\int_K f(x)\mu(dx)$ by $\langle f, \mu \rangle$. $\mu_n \to \mu$ in the strong topology iff $\langle f, \mu_n \rangle \to \langle f, \mu \rangle$ uniformly for all continuous f with $|f| \leq 1$.

For $Q, R \in \Delta$, it can be shown (see Lemma 1, p.360, Shiryaev [10]) that

$$\|Q - R\|_{var} = 2 \sup_{B \in \mathcal{B}(K)} |Q(B) - R(B)|.$$

For this reason, $R_n \to R$ in the strong topology iff $R_n(B) \to R(B)$ uniformly for $B \in \mathcal{B}(K)$. For instance, if $x \neq y$, $\|\delta_x - \delta_y\| = 2$. This implies that Δ, equipped with this topology, is not compact.

On the other hand, the Banach space $(\mathcal{M}, \|\cdot\|_{var})$ is the dual space of $C^0(K)$.[2] The weak topology on probabilities is its $weak^*$ topology. Therefore, equipped with that weak topology, it is compact (and metrizable).

If R_n, R are probability measures, then $R_n \to R$ in the weak topology iff $\langle f, R_n \rangle \to \langle f, R \rangle$ for every measurable $f \in C(K)$ with $|f| \leq 1$. It can be shown that $R_n \to R$ in the weak topology iff $R_n(B) \to R(B)$ for all $B \in \mathcal{B}(K)$ with $R(\partial B) = 0$.[3]

There are various (equivalent, of course) metrics which generate the weak topology on Δ. One is the *Prohorov metric*:

$$\rho(Q, R) = \\ \inf\{\varepsilon > 0 \mid Q(C) \leq R(C^\varepsilon) + \varepsilon,\ R(C) \leq Q(C^\varepsilon) + \varepsilon, \text{for all closed } C \subset K\},$$

where $C^\varepsilon = \{x \in K \mid \inf_{y \in K} |y - x| < \varepsilon\}$.

Another metric generating the weak topology is

$$d(Q, R) = \sup\{|\langle f, Q - R \rangle| \mid f \text{ Lipschitz continuous}, \|f\|_\infty + L(f) \leq 1\},$$

where $\|f\|_\infty = \sup_{x \in K} |f(x)|$ and $L(f) = \sup_{x \neq y} \frac{|f(x) - (f(y)|}{|x - y|}$.

---

[2] As usual, $C^0(K)$ denotes the space of continuous real valued functions on K, equipped with uniform norm.

[3] ∂B is the boundary of B.

Yet another one is defined using the set $M(Q,R)$ of measures over $K\times K$ whose marginals are Q and R, respectively, as

$$d^2(Q,R) = \inf_{\mu\in M(Q,R)} \int_{K\times K} \|x-y\|^2\,\mu(dx,dy)\,.$$

For more details on strong and weak convergence of probability measures, we refer to Shiryaev [10].

4 Replicator dynamics

4.1 Definition and existence

Let $Q(t)\in\Delta$ be the population state at time t. For a given state Q, the difference between the average fitness of a subpopulation B (a Borel subset of K) and the population average fitness is

$$\frac{1}{Q(B)}\int_B u(y,Q)\,Q(dy) - u(Q,Q) = \frac{1}{Q(B)}\int_B \sigma(y,Q)\,Q(dy).$$

If we assume that the rate of replication of the subpopulation B is the difference between the average fitness of the subpopulation B and the population average fitness, then we obtain the replicator dynamics in the form:

$$\dot{Q}(t)(B) = \int_B \sigma(y,Q(t))\,Q(t)(dy) \quad B\in\mathcal{B}(K).$$

In order to make the replicator equation look simpler, we introduce the notation:

$$F(Q)(B) := \int_B \sigma(y,Q)\;Q(dy) \quad B\in\mathcal{B}(K).$$

That is, $F(Q)$ is the measure which is absolutely continuous w.r.t. Q (denoted $F(Q)\ll Q$) with the Radon-Nikodym derivative $\frac{dF(Q)}{dQ}=\sigma(\cdot,Q)$.

The replicator dynamics can now be written as

$$\dot{Q}(t) = F(Q(t)). \tag{9}$$

Now, for any measurable f with $|f|\le 1$, and for $Q,R\in\Delta$,

$$\begin{aligned}
&|\langle f, F(Q)-F(R)\rangle| = |\langle f\sigma(\cdot,Q),Q\rangle - \langle f\sigma(\cdot,R),R\rangle|\\
&\quad\le |\langle f\sigma(\cdot,Q),Q-R\rangle| + |\langle f[\sigma(\cdot,Q)-\sigma(\cdot,R)],R\rangle|\\
&\quad\le \|\sigma\|_\infty\|Q-R\|_{var} + \|u\|_\infty\|Q-R\|_{var} + |\langle f[u(Q,Q)-u(R,R)],R\rangle|\\
&\quad\le 3\|u\|_\infty\|Q-R\|_{var} + |u(Q,Q)-u(R,R)|\\
&\quad= 3\|u\|_\infty\|Q-R\|_{var} + |u(Q,Q-R)| + |u(Q-R,R)|\\
&\quad\le 5\|u\|_\infty\|Q-R\|_{var}.
\end{aligned}$$

This yields

$$\|F(Q) - F(R)\|_{var} \leq 5\|u\|_\infty \|Q - R\|_{var}.$$

In a similar fashion, one can show that the map F is Lipschitz Continuous (w.r.t. the strong topology) on any bounded ball of the space $\mathcal{M}$ of all finite signed measures (Lipschitz constant depends on the radius of the ball). Therefore, F is locally Lipschitz continuous on the Banach space $(\mathcal{M}, \|\cdot\|_{var})$, and hence the replicator dynamics admits a unique solution, for each initial distribution $Q(0)$.

Furthermore,

$$\frac{d}{dt}\int_K Q(t)(dy) = \int_K \dot{Q}(t)(dy) = \int_K F(Q(t))(dy) = \sigma(Q(t), Q(t)) = 0.$$

As a result, the simplex Δ of all probability measures is invariant under the replicator dynamics. Thus, for any initial condition $Q \in \Delta$, the replicator dynamics admits a unique solution $Q(\cdot)$ defined (and stays in Δ) for all times t with $Q(0) = Q$.

4.2 Stationary states of the replicator dynamics

Let Δ^0 denote the set of all stationary points of the replicator dynamics in Δ. Now $Q \in \Delta^0$ iff $F(Q)(B) = 0$ for all B iff $\sigma(\cdot, Q) = 0$ a.e.(Q). That is, $\sigma(y, Q) = 0$ except on a Borel set $N \subset K$ with $Q(N) = 0$. This implies that

$$\Delta^0 = \{Q \in \Delta \,:\, \sigma(\cdot, Q) = 0 \text{ a.e.}(Q)\}.$$

For every $x \in K$,

$$\frac{\mathrm{d}F(\delta_x)}{\mathrm{d}\delta_x} = \sigma(\cdot, x) = 0 \qquad \text{a.e. } (\delta_x),$$

and so, as in the finite case, Δ^0 contains all pure strategies. It now follows that:

Theorem 4.1.

(1) $\{\delta_x \,:\, x \in K\} \cup \Delta^{NE} \subset \Delta^0$,
(2) $\Delta^{NE} \cap \mathrm{int}(\Delta) = \Delta^0 \cap \mathrm{int}(\Delta)$.

A state Q is Lyapunov stable, if any trajectory starting "nearby" Q stays "nearby" Q. Hence, stability crucially depends on the topology we use on Δ.

We now prove that any Lyapunov stable state (w.r.t. weak topology) is a Nash equilibrium.

Theorem 4.2. *If $Q \in \Delta$ is Lyapunov stable under* (9) *w.r.t the weak topology, then $Q \in \Delta^{NE}$.*

Proof. If $Q \notin \Delta^{NE}$, then there exists a pure strategy $x \in K$ such that

$$\sigma(x, Q) = u(x, Q) - u(Q, Q) > 0.$$

Since in the weak topology, $r \mapsto \sigma(x, R)$ is continuous, this implies that there exists a weak neighborhood G of Q and $\delta > 0$ such that

$$\sigma(x, R) \geq \delta \quad \text{for all } R \in G. \tag{10}$$

For $\varepsilon > 0$ small enough, the strategies $Q_\varepsilon(0) := \varepsilon\delta_x + (1-\varepsilon)Q \in G$. Now,

$$Q_\varepsilon(t)(x) = Q_\varepsilon(0)(x)\mathrm{e}^{\int_0^t \sigma(x, Q(s))\, ds} = \varepsilon\mathrm{e}^{\int_0^t \sigma(x, Q_\varepsilon(s))\, ds}$$

is the trajectory emanating from $Q_\varepsilon(0)$. This, in view of (10), contradicts the Lyapunov stability of the state Q. ∎

4.3 Limit states of replicator dynamics

Theorem 4.3. *If there is an interior trajectory $Q(t)$ with $\lim_{t\to\infty} Q(t) = P \in \mathrm{int}(\Delta)$ (in the weak topology), then $P \in \Delta^{NE}$.*

Proof. If $Q \notin \Delta^{NE}$ then, as in the proof of Theorem 4.2, there exist a pure strategy $x \in K$, a weak neighborhood G of P and $\delta > 0$ such that

$$\sigma(x, R) \geq \delta \quad \text{for all } R \in G.$$

This implies that, by continuity, that there exists an open set I containing x, with $P(I) \neq 0$ (because P is an interior point), such that

$$F(R)(I) = \int_I \sigma(y, R)\, R(dy) \geq \frac{\delta}{2} P(I) \quad \text{for all } R \in G.$$

Hence, since for some t_0 and $t \geq t_0$ $Q(t) \in G$,

$$Q(t)(I) = Q(0)(I) + \int_0^t F(Q(s))\, \mathrm{d}s \geq Q(t_0)(I) + \frac{\delta}{2}(t - t_0)P(I)$$

and $Q(t)(I)$ diverges. This contradiction proves the theorem. ∎

4.4 Asymptotically stable states of replicator dynamics and locally superior strategies

One would like to prove, as in the finite case, that an evolutionary robust strategy (that is, locally superior w.r.t. the weak topology) P is Lyapunov stable and attracting in the weak topology (henceforth called *weakly asymptotically stable*). The main difficulty here is that the analogue of the Lyapunov function used in the finite case, is not continuous in the weak topology. More precisely, the function[4]

$$V_P(Q) := \begin{cases} \int_K \ln(\frac{\mathrm{d}P}{\mathrm{d}Q})\mathrm{d}P & \text{if } P \ll Q \\ +\infty & \text{otherwise} \end{cases} \tag{11}$$

[4] $V_P(Q) = \int_K log(\frac{dP}{dQ}\, dP)$ is the Kullback-Leibler distance between P and Q.

is not well behaved w.r.t. weak topology. Nevertheless, one can establish that this map is well behaved along the trajectories of the replicator dynamics (9). We use the next lemma to establish this.

Lemma 4.1. *Let $Q(\cdot)$ solve the replicator Eq.* (9) *with initial state $Q(0)$. Then, for every $t > 0$, $Q(t) \approx Q(0)$* [5] *and*

$$\frac{\mathrm{d}Q(t)}{\mathrm{d}Q(0)} = \mathrm{e}^{\int_0^t \sigma(\cdot, Q(s))\,\mathrm{d}s}.$$

Proof. See Bomze [3]. ∎

One also has the following result:

Theorem 4.4.

(1) $V_P(Q) \geq 0$ with equality iff $Q = P$.

(2) If $V_P(Q(0)) < \infty$, then, for all $0 < t < \infty$, $V_P(Q(t)) < \infty$ and

$$\frac{d}{dt} V_P(Q(t)) = -\sigma(P, Q(t)).$$

Proof. Clearly, $V_P(P) = 0$. Let S be the support of P. For $Q \neq P$,

$$\begin{aligned} V_P(Q) &= -\textstyle\int_S \ln(\frac{\mathrm{d}Q}{\mathrm{d}P})\,\mathrm{d}P \\ &> -\textstyle\int_S (\frac{\mathrm{d}Q}{\mathrm{d}P} - 1) \\ &= 1 - Q(S) \\ &\geq 0. \end{aligned}$$

To prove (ii), let $V_P(Q(0)) < \infty$. For $0 < t < \infty$, by the above lemma,

$$\begin{aligned} \ln(\tfrac{\mathrm{d}P}{\mathrm{d}Q(t)}) &= \ln \tfrac{\mathrm{d}P}{\mathrm{d}Q(0)} + \ln \tfrac{\mathrm{d}Q(0)}{\mathrm{d}Q(t)} \\ &= \ln \tfrac{\mathrm{d}P}{\mathrm{d}Q(0)} - \textstyle\int_0^t \sigma(\cdot, Q(s))\,\mathrm{d}s. \end{aligned}$$

This equality of functions holds for $y \in K$ a.e (Q(0)). This implies that

$$V_P(Q(t)) = V_P(Q(0)) - \int_0^t \sigma(P, Q(s))\,\mathrm{d}s.$$

This gives the required result. ∎

Because of the lack of continuity of this Lyapunov function, the stability result for the continuous case has not been settled completely. Nevertheless, we can prove the stability in some special cases, as the next two theorems show.

---

[5] $Q(t) \approx Q(0)$ means that $Q(t) \ll Q(0)$ and $Q(0) \ll Q(t)$.

Theorem 4.5 (Oechssler & Riedel [9])**.** *If $u(x,y)=u(y,x)$, $\forall x,y\in K$ (that is, if the game is doubly symmetric), then an evolutionarily robust strategy is weakly asymptotically stable.*

Proof. Let $P\in\Delta$ be an evolutionarily robust strategy, and G a weak neighborhood of P such that

$$u(P,Q)-u(Q,Q)>0 \quad \forall Q\in G\setminus\{P\}.$$

Consider the function

$$W(Q)=u(P,P)-u(Q,Q)=u(P,P)-u(Q,P)+u(P,Q)-u(Q,Q).$$

Therefore, $W(Q)\geq 0$ for all $Q\in G$, with equality iff $Q=P$. In addition, W is weakly continuous. Since the weak topology is compact, it is enough to prove that, for every trajectory $Q(\cdot)$ in G, the derivative of $t\mapsto W(Q(t))$ is negative.

$$\begin{aligned} -\frac{\mathrm{d}}{\mathrm{d}t}W(Q(t)) &= \frac{\mathrm{d}}{\mathrm{d}t}u(Q(t),Q(t))\\ &= 2u(F(Q(t)),Q(t))\\ &= 2\int_K u(y,Q(t))\sigma(y,Q(t))\,Q(t)(\mathrm{d}y)\\ &= 2\int_K \sigma^2(y,Q(t))\,Q(t)(dy)\end{aligned}$$

This always positive, as long as the trajectory stays inside G and $Q(0)\neq P$. ∎

Theorem 4.6. *Let $P=\delta_x$ be an ERS. Then P is a stable state (in the weak topology) of the replicator dynamics. Furthermore, $Q(t)\to P$ (in the weak topology) as $t\to\infty$, whenever $Q(0)$ is weakly near P and $Q(0)(\{x\})>0$.*

Proof. We have

$$\dot{Q}(t)(x)=\sigma(x,Q(t))Q(t)(x),$$

and hence

$$Q(t)(x)=Q(0)(x)\mathrm{e}^{\int_0^t \sigma(x,Q(s))\,ds}.$$

Now let $P=\delta_x << Q(0)$, since $Q(0)(x)>0$. If $Q(0)$ is close to δ_x, then $\sigma(x,Q(0))>0$ and hence $t\mapsto Q(t)(x)$ is initially (strictly) increasing. As a result, $Q(t)$ becomes (weakly) closer to δ_x than $Q(0)$, and therefore $\sigma(x,Q(t))>0$ for all t.

Now, since $Q(t)(x)\leq 1$, $\forall t$, we must have

$$\sigma(x,Q(t))\to 0 \quad \text{as } t\uparrow\infty.$$

Since Δ is compact in the weak topology, any trajectory $t\mapsto Q(t)$ must have limit points. Any such limit point R (this has to be near δ_x, by above arguments) satisfies $\sigma(x,R)=0$. This can happen only if $R=\delta_x$. This implies that if we start with a $Q(0)$ weakly close to δ_x, then $Q(t)$ converges weakly to δ_x as $t\to\infty$. ∎

4.5 Stability of global evolutionarily robust strategies

We establish now a new result concerning global ERS.

Theorem 4.7. *Let $P \in \Delta$, $\Omega_P = \{R \in \Delta : \sigma(P,R) = 0\}$, and $\Sigma_P = \{R \in \Delta : \sigma(P,R) \geq 0\}$. Let $Q(\cdot)$ be a trajectory of the replicator dynamics* (9) *with $V_P(Q(0)) < \infty$. If, in addition, $Q(t) \in \Sigma_P$ for all t large enough, then, Ω_P contains all the weak ω-limit points of the trajectory $Q(t)$.*

Proof. Let $t_0 \geq 0$ be such that

$$Q(t) \in \Sigma_P \quad \forall t \geq t_0.$$

From the proof of Theorem 4.4, it follows that, for $t \geq t_0$

$$V_P(Q(t)) = V_P(Q(t_0)) - \int_{t_0}^{t} \sigma(P, Q(s))\,\mathrm{d}s.$$

In particular,

$$\sup_{t \geq t_0} \left[\int_{t_0}^{t} \sigma(P, Q(s))\,\mathrm{d}s \right] \leq V_P(Q(t_0)) < \infty.$$

If, after time t_0, the trajectory $Q(\cdot)$ lies in Σ_P, then the map

$$[t_0, \infty) \ni t \mapsto \int_{t_0}^{t} \sigma(P, Q(s))\,\mathrm{d}s$$

is nondecreasing, and hence has to converge to a finite number, as $t \to \infty$. Since

$$s \mapsto \sigma(P, Q(s))$$

is Lipschitz continuous, it follows that

$$\lim_{s \to \infty} \sigma(P, Q(s)) = 0.$$

Now

$$R \mapsto \sigma(P, R)$$

is weakly continuous, and hence, for any weak limit point Q of the trajectory $Q(\cdot)$,

$$\sigma(P, Q) = 0.$$

∎

The next result is an immediate consequence of Theorem 4.7, since then $\Omega_P = \{P\}$ and $\Sigma_P = \Delta$:

Theorem 4.8. *Let $P \in \Delta$ be a global evolutionarily robust strategy. Let $Q(\cdot)$ be a trajectory of the replicator dynamics* (9) *with $V_P(Q(0)) < \infty$. Then, as $t \to \infty$, the trajectory $Q(t)$ converges weakly to P.*

5 Examples

In any area of mathematical research, the importance of ample analytical examples cannot be overemphasized. To the best of our knowledge, there is no nontrivial (that is, non-Dirac) examples of evolutionary robust strategies, in the literature, at least with a continuous fitness function u. We provide here, two such examples: one is a convex combination of Diracs, and the other one has density w.r.t. Lebesgue measure.

5.1 Example 1

We take $K = [0, 1]$. Let $\lambda \in (0, \infty)$, and $u(x, y) = \max\{x - y, \lambda(y - x)\}$. For convenience, let $\alpha = \frac{\lambda}{1+\lambda}$, $\beta = \frac{1}{1+\lambda}$.[6] ‘ Let δ_0 and δ_1 denote the Dirac measures concentrated at the points $\{0\}$ and $\{1\}$, respectively.

We claim the following.

Theorem 5.1. *The strategy $P = \alpha\delta_0 + \beta\delta_1$ is globally superior.*

Proof. The rest of this subsection is devoted to the proof of that theorem.

For any $0 \leq x \leq 1$,

$$u(x, 0) = x, \quad \text{and} \quad u(x, 1) = \lambda(1 - x).$$

Therefore, for $P = \alpha\delta_0 + \beta\delta_1$ and $0 \leq x \leq 1$,

$$\begin{aligned} u(x, P) &= \alpha u(x, 0) + \beta u(x, 1) \\ &= \alpha x + \beta\lambda(1 - x) \\ &= \alpha. \end{aligned}$$

This implies that $P \in \Delta^{NE}$ and $BR(P) = \Delta$; that is, all strategies gain the same fitness α against P. Our aim is to show that

$$\sigma(P, R) = u(P, R) - u(R, R) > 0 \qquad \forall R \in \Delta \setminus \{P\}. \tag{12}$$

Since for each $0 \leq y \leq 1$, the map $x \mapsto u(x, y)$ is convex, it follows that, for every $R \in \Delta$, the map $x \mapsto u(x, R)$ is also convex. As a result, the function

$$[0, 1] \ni x \mapsto u(x, R)$$

attains its maximum at either $x = 0$ or $x = 1$.

We denote $R(\{x\})$ by $R(x)$. We also denote the “average value” of R as

$$[R] := \int_{[0,1]} x \, R(\mathrm{d}x).$$

---

[6] Notice that it would be equivalent to start with two positive numbers α and $\beta = 1 - \alpha$ and to set $u(x, y) = \max\{\alpha(y - x), \beta(x - y)\}$.

For $0 \leq x \leq 1$,

$$u(0,x) = \lambda x, \quad \text{and} \quad u(1,x) = 1 - x.$$

This gives

$$u(0,R) = \lambda[R], \quad \text{and} \quad u(1,R) = 1 - [R].$$

For this reason, $u(0,R) = u(1,R)$ (resp. $u(0,R) < u(1,R)$, $u(0,R) > u(1,R)$) iff $[R] = \beta$ (resp. $[R] < \beta$, $[R] > \beta$).

It is also true that

$$[R] \leq 1 - R(0), \quad \forall R \in \Delta,$$

with equality iff $R(0) + R(1) = 1$.

We prove (12) by considering several cases.

Case 1: $R(0) > \alpha$.
In this case, $\alpha < 1 - [R]$, and so $u(0,R) < u(1,R)$. Because of this, $u(\cdot,R)$ attains maximum at $x = 1$. Now

$$\begin{aligned} u(R,R) &= R(0)u(0,R) + \int_{(0,1]} u(x,R)\, R(\mathrm{d}x) \\ &\leq R(0)u(0,R) + (1 - R(0))u(1,R) \\ &< \alpha u(0,R) + \beta u(1,R) \\ &= u(P,R). \end{aligned}$$

Case 2: $R(0) = \alpha$.
For $R \neq P$, we must have $R((0,1)) > 0$, and hence $R(1) < \beta$. Furthermore, $[R] < \beta$. Therefore, $u(0,R) < u(1,R)$. The convexity of $u(\cdot,R)$ now yields, for $0 < x < 1$,

$$\begin{aligned} u(x,R) &\leq (1-x)u(0,R) + xu(1,R) \\ &< u(1,R). \end{aligned}$$

This implies that

$$\begin{aligned} u(R,R) &= R(0)u(0,R) + \int_{(0,1]} u(x,R)\, R(\mathrm{d}x) \\ &< R(0)u(0,R) + (1 - R(0))u(1,R) \\ &= \alpha u(0,R) + \beta u(1,R) \\ &= u(P,R). \end{aligned}$$

Case 3: $R(0) < \alpha$ and $R(1) > \beta$.
In this case, $[R] > \beta$. Hence, $u(0,R) > u(1,R)$, and so $u(\cdot,R)$ attains maximum

at $x = 0$. Now

$$\begin{aligned} u(R,R) &= R(1)u(1,R) + \textstyle\int_{[0,1)} u(x,R)\, R(\mathrm{d}x) \\ &\leq R(1)u(1,R) + (1 - R(1))u(0,R) \\ &= u(1,R) + (1 - R(1))[u(0,R) - u(1,R)] \\ &< u(1,R) + \alpha[u(0,R) - u(1,R)] \\ &= \alpha u(0,R) + \beta u(1,R) \\ &= u(P,R). \end{aligned}$$

Case 4: $R(0) < \alpha$ and $R(1) = \beta$.
In this case, $R((0,1)) > 0$ and $[R] > \beta$. This yields $u(0,R) > u(1,R)$ and $x = 0$ is the maximizer of $u(\cdot,R)$. Now, as in Case 2, $u(x,R) < u(0,R)$ for $0 < x < 1$, and

$$\begin{aligned} u(R,R) &= R(1)u(1,R) + \textstyle\int_{[0,1)} u(x,R)\, R(\mathrm{d}x) \\ &< R(1)u(1,R) + (1 - R(1))u(0,R) \\ &= \alpha u(0,R) + \beta u(1,R) \\ &= u(P,R). \end{aligned}$$

Case 5: $R(0) < \alpha$ and $R(1) < \beta$.
Define the numbers $x_0 = x_0(R)$, $x_1 = x_1(R)$ as

$$\begin{aligned} x_0 &:= \inf\{x \in [0,1] \,:\, R([0,x]) \geq \alpha\}, \\ x_1 &:= \sup\{x \in [0,1] \,:\, R([x,1]) \geq \beta\}. \end{aligned}$$

Clearly, $x_0 > 0$ and $x_1 < 1$. Furthermore, $R([0,x_0]) \geq \alpha$ and $R([x_1,1]) \geq \beta$. However, we have more. More precisely, $x_0 \leq x_1$. To see this, let $x_0 > x_1$. Now for $x_0 > x > x_1$,

$$R([0,x]) < \alpha \quad \text{and } R([0,x)) > \alpha.$$

This is a contradiction. Thus, $0 < x_0 \leq x_1 < 1$.

We first consider the case $x_0 < x_1$. In this case, we must have

$$R([0,x_1]) = \alpha \quad \text{and} \quad R([x_1,1]) = \beta,$$

and hence, $R((x_0,x_1)) = 0$.

We now define

$$a := \int_{[0,x_0]} x\, R(\mathrm{d}x) \quad \text{and} \quad b := \int_{[x_1,1]} (1-x)\, R(\mathrm{d}x).$$

Note that a and b cannot vanish at the same time ($a = b = 0$ would imply that $1 = R(0) + R(1) < \alpha + \beta = 1$, a contradiction). Moreover, for $x \in (x_0, x_1)$, we have

$$\begin{aligned} u(x,R) &= \int_{[0,x_0]}(x-y)\,R(\mathrm{d}y) + \int_{[x_1,1]}\lambda(y-x)\,R(\mathrm{d}y) \\ &= \int_{[0,x_0]}(x-y)\,R(\mathrm{d}y) + \lambda\int_{[x_1,1]}[(1-x)-(1-y)]\,R(\mathrm{d}y) \\ &= xR([0,x_0]) + \lambda(1-x)R([x_1,1]) - a - \lambda b \\ &= \alpha x + \lambda\beta(1-x) - a - \lambda b \\ &= \alpha - a - \lambda b. \end{aligned}$$

In particular,

$$u(x_0,R) = u(x_1,R) = \alpha - a - \lambda b.$$

Similar calculations also yield

$$u(0,R) = \alpha + \lambda a - \lambda b, \quad u(1,R) = \alpha - a + b.$$

We now use the convexity of $u(\cdot,R)$ to get

$$u(x,R) \leq \begin{cases} u(0,R) - \frac{x}{x_0}[u(0,R) - u(x_0,R)] & ;\ 0 \leq x \leq x_0 \\ u(1,R) - \frac{1-x}{1-x_1}[u(1,R) - u(x_1,R)] & ;\ x_1 \leq x \leq 1. \end{cases}$$

Integrating $u(\cdot,R)$ w.r.t. the probability measure R, using the above calculations, yield

$$\begin{aligned} u(R,R) &\leq \alpha u(0,R) + \beta u(1,R) - \tfrac{a}{x_0}[u(0,R) - u(x_0,R)] - \\ &\qquad \tfrac{b}{1-x_1}[u(1,R) - u(x_1,R)] \\ &= u(P,R) - \tfrac{a}{x_0}(1+\lambda)a - \tfrac{b}{1-x_1}(1+\lambda)b \\ &= u(P,R) - (1+\lambda)[\tfrac{a^2}{x_0} + \tfrac{b^2}{1-x_1}] \\ &< u(P,R). \end{aligned}$$

Finally, when $x_0 = x_1$, we let $r_0 := R(x_0)$, and let $\gamma \in [0,1]$ such that

$$R([0,x_0)) = \alpha - \gamma r_0, \quad \text{and} \quad R((x_1,1]) = \beta - (1-\gamma)r_0.$$

Define now

$$a = \int_{[0,x_0)} x\,R(dx) + x_0\gamma r_0, \quad \text{and} \quad b = \int_{(x_0,1]} (1-x)\,R(dx) + (1-x_0)(1-\gamma)r_0.$$

Now we can proceed as above to complete the proof. □

5.2 Example 2

Let again $K = [0,1]$ and P, the Lebesgue measure on $[0,1]$. For $n = 1, 2, \ldots$, let

$$f_n(x) := \cos(n\pi x).$$

Let $a_n > 0$ for all n, and $\sum_{n=1}^{\infty} a_n < \infty$, g be any continuous function with $\langle g, P \rangle = 0$. Define

$$u(x,y) := g(y) - \sum_{n=1}^{\infty} a_n f_n(x) f_n(y).$$

We claim the following.

Theorem 5.2. *The Lebesgue measure P is a globally superior strategy.*

Proof. The rest of this subsection is devoted to the proof of that theorem.

Clearly, $\langle f_n, P \rangle = 0$ for all n, and so $u(P,R) = \langle g, R \rangle$ for every $R \in \Delta$. Now

$$u(R,R) = \langle g, R \rangle - \sum_{n=1}^{\infty} a_n |\langle f_n, R \rangle|^2.$$

This gives

$$u(P,R) - u(R,R) = \sum_{n=1}^{\infty} a_n |\langle f_n, R \rangle|^2.$$

This is obviously nonnegative, and vanishes iff

$$\langle f_n, R \rangle = 0 = \langle f_n, P \rangle \quad \forall n = 1, 2, \ldots.$$

This can happen iff

$$\langle f, R \rangle = \langle f, P \rangle$$

for all continuous functions f on $[0,1]$ (because any continuous function can be approximated pointwise by a finite linear combination of f_n's). Therefore, $R = P$. □

Acknowledgements

A.J. Shaiju gratefully acknowledges post-doctoral fellowship from the Ministry of Research, France. He is currently at school of ITEE, UNSW@ADFA, Canberra, ACT 2600, Australia.

References

[1] I. M. Bomze and B. M. Potscher *Game Theoretical Foundations of Evolutionary Stability*, Springer, Berlin, 1989.

[2] I. M. Bomze "Dynamical aspects of evolutionary stability", *Monatsh. Math.*, **110** (1990), 189-206.

[3] I. M. Bomze "Cross entropy, minimization in uninvadable states of complex populations", *J. Math. Biol.*, **30** (1991), 73–87.

[4] J. Hofbauer, P. Schuster and K. Sigmund "A note on evolutionary stable strategies and game dynamics", *J. Theor. Biol.*, **81** (1979), 609–612.

[5] J. Hofbauer and K. Sigmund *Evolutionary Games and Population Dynamics*, Cambridge University Press, Cambridge, 1998.

[6] J. Maynard Smith and G. R. Price "The logic of animal conflict", *Nature*, **246** (1973), pp. 15–18.

[7] J. Maynard Smith "The theory of games and the evolution of animal conflicts", *J. Theor. Biol.*, 1974

[8] J. Maynard Smith *Evolution and the Theory of Games*, Cambridge University Press, Cambridge, 1982.

[9] J. Oechssler and F. Riedel "On the dynamic foundation of evolutionary stability in continuous models", *J. Economic Theory*, **107** (2002), 223–252.

[10] A. N. Shiryaev *Probability*, 2nd ed, Springer, New York, 1995.

[11] P. D. Taylor and L. B. Jonker "Evolutionarily stable strategies and game dynamics", *Math. Biosci.*, **40** (1978), 145–156.

[12] G. T. Vickers and C. Cannings "On the definition of an evolutionarily stable strategy." *J. Theor. Biol.*, **129** (1987), pp. 349–353.

[13] J. W. Weibull *Evolutionary Game Theory*, MIT Press, Cambridge, 1995.

[14] E. Zeeman "Dynamics of the evolution of animal conflicts", *J. Theor. Biol.*, **89** (1981), pp. 249–270.

Nash Equilibrium in a Game Version of the Elfving Problem

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

Anna Krasnosielska
Faculty of Mathematics and Information Science
Warsaw University of Technology
00-661 Warsaw, Poland
akrasno@mini.pw.edu.pl

Abstract

Multi-person stopping games with a finite and infinite horizon, players' priorities, and observed rewards at jump times of a Poisson process are considered. The existence of a Nash equilibrium is proved and its explicit form is obtained for special classes of reward sequences. Our game is a generalization of the Elfving stopping time problem to the case of many players and modification of multi-person stopping games with priorities.

Key words. Stopping game, Nash equilibrium, Elfving problem, Poisson process.

AMS Subject Classifications. 60G40, 91A60.

1 Introduction

Suppose that there are m companies that are searching for a project to realize. The companies are ordered according to the quality of their products; the ordering is referred to as priority so that 1 refers to the best company and m to the worst. Each company wishes to accept one project, which gives the highest possible reward. On the market one observes a sequence of projects with values $X_n = Y_n r(\tau_n)$, $n = 1, 2, 3, ...$, which appear at times τ_n of jumps of a Poisson process and r is a discount function. The value X_n is interpreted as a company's reward.

We will formulate the problem as an m-person stopping game with priorities in which random offers are presented at jump times of a homogeneous Poisson process. An optimal selection strategy is, by definition, a Nash equilibrium point of the game.

Another application of the game concerns business lines. Let us assume that there are m players and each of them has one business line which transfers money.

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_20,

The business lines are ordered according to their quality. As previously mentioned, we can associate a priority with each line. On the market there are customers who want to transfer money. The customers appear sequentially at times τ_n. The n-th customer who transfers the value Y_n at τ_n yields the reward $Y_n r(\tau_n)$ to the player. Each player wishes to transfer as much money as possible but is allowed to make only one transfer of money.

The game is a generalization, to the case of several players, of the optimal stopping time problem formulated and solved first by Elfving [3], later considered also by Siegmund [14]. Stadje [15] considered the optimal multi-stopping time problem in the Elfving setting, in which the final reward is the sum of selected discounted random variables. Enns and Ferenstein [4] investigated the two-person nonzero-sum stopping game with priorities such that the rewards are observed at jump times of a Poisson process. The player who selects, in a sequential way with priority constraints, a random variable which is larger than the one selected by his opponent wins 1, otherwise 0. So the aim of each player is to maximize the probability of selecting the larger random variable. Priority means that if both players decide to select the observation just presented, the player with the higher priority gets it. Saario and Sakaguchi [10] considered multi-stopping best choice games with offers observed also at times of successive jumps of a Poisson process and with priority constraints. The players' final rewards were maxima of their selected offers. Porosiński and Szajowski [9] considered a two-person best choice full information problem with imperfect observations in a similar continuous time setting. Kurano, Nakagami and Yasuda [8] also considered a game variant of a stopping problem for jump processes. General models of multi-person discrete-time games with priorities were analyzed in Ferenstein [6], [7], Enns and Ferenstein [5], and Sakaguchi [13]. Dixon [2] considered three kinds of games with Poisson streams of offers consisting of independent and identically distributed random variables having the uniform distribution on $[0,1]$. The games had different reward structures. A sequential game with player's random priority played over two Poisson streams was investigated by Sakaguchi [11]. He obtained simultaneous differential equations for the equilibrium values of the game. Another kind of two-person game with a Poisson stream of offers is presented in Sakaguchi [12]. In that game the players have equal weights, that is, if both players want to accept the same reward, a lottery is used to ensure that each player can get it with probability $\frac{1}{2}$.

In this paper we formulate the game model, and obtain results on existence of a Nash equilibria and Bellmann equations.

2 One-person game—the Elfving problem

In this section we recall the Elfving stopping time problem. The basic assumptions and notations follow those of Chow et al. [1]. Let $(\Omega, \mathcal{F}, P)$ denote a basic probability space on which all random objects are considered. Let $Y_1, Y_2, ...$ be independent, nonnegative random variables with common piecewise continuous distribution function F and mean $\mu \in (0, \infty)$. Let $0 < \tau_1 < \tau_2 < ...$ be jump times of a homogeneous Poisson process $N(t), t \geq 0$, with intensity $\lambda = 1$ and $\tau_0 = 0$.

The random variables $Y_1, Y_2, ...$ are interpreted as the rewards observed successively by one player at times $\tau_1, \tau_2, ...$ We assume that the sequences $\{Y_n\}$, $\{\tau_n\}$ are independent. Furthermore, we are given a discount function $r : [0, \infty) \to [0, 1]$ which is non-increasing, continuous from the right, and $r(0) = 1$, $r(U) = 0$, where U is finite or infinite, and

$$\int_0^\infty r(u)du < \infty.$$

$X_n = Y_n r(\tau_n)$ is the reward for the player who accepts Y_n at time τ_n on the basis of observations $Y_1, ..., Y_n, \tau_1, ..., \tau_n$. Hence, $\mathcal{F}_n := \sigma(Y_1, ..., Y_n, \tau_1, ..., \tau_n)$ is the σ-field of the observed events till time τ_n. Let Λ denote the set of discrete stopping (selection) rules with respect to $\{\mathcal{F}_n\}$; that means $\sigma \in \Lambda$ iff $\{\sigma = n\} \in \mathcal{F}_n$. Let us recall that an optimal stopping rule $\hat{\sigma} \in \Lambda$ maximizes the player's mean reward in Λ, that is,

$$\sup_{\sigma \in \Lambda} E[X_\sigma] = E[X_{\hat{\sigma}}].$$

Putting $Z_n = (Y_n, \tau_n)$, we see that we are in a homogeneous Markov case.

Let $V_{1,1}(u)$ be the optimal mean reward for stopping the sequence $\{Y_n r(u + \tau_n)\}$, $u \in [0, U)$. Let us set

$$y_{1,1}(u) = \frac{V_{1,1}(u)}{r(u)} \qquad \text{if} \qquad 0 \leq u < U. \tag{1}$$

If $u \geq U$, then $y_{1,1}(u) = 0$. Let us define the stopping rules

$$\sigma_1 = \inf\{n \geq 1 : Y_n \geq y_{1,1}(\tau_n)\},$$

$$\sigma_1(u) = \inf\{n \geq 1 : Y_n \geq y_{1,1}(u + \tau_n)\}.$$

Note that $\sigma_1 = \sigma_1(0) = \hat{\sigma}$ is the optimal stopping rule for the reward sequence $\{X_n\}$. Set

$$G(y) = 1 - F(y), \tag{2}$$

$$H(y) = \int_y^\infty y' dF(y'), \tag{3}$$

$$f_{1,u}(v) = P(\tau_{\sigma_1(u)} > v).$$

Elfving obtained the formula (see Chow [1], pp. 114-115)

$$f_{1,u}(v) = \exp\Big[- \int_u^{u+v} G(y_{1,1}(v'))dv' \Big]. \tag{4}$$

Hence, one derives the optimal mean reward:

$$y_{1,1}(u)r(u) = V_{1,1}(u) = \int_u^U r(v)H(y_{1,1}(v)) \exp\left[- \int_u^v G(y_{1,1}(v'))dv' \right] dv. \tag{5}$$

Differentiating both sides of (5) with respect to u we obtain

$$\frac{d}{du}[r(u)y_{1,1}(u)] = -r(u)[H(y_{1,1}(u)) - y_{1,1}(u)G(y_{1,1}(u))]. \tag{6}$$

The theorem below answers the question how to find the optimal mean reward and optimal stopping rule.

Theorem 2.1. *(i) A piecewise continuous function $\tilde{y}(\cdot)$ satisfies* (5) *if and only if $\tilde{y}(\cdot)$ satisfies* (6) *and $r(u)\tilde{y}(u) \to 0$ as $u \to U$.*
(ii) If $\tilde{y}(\cdot)$ satisfies (5) *on $[0, U)$ and $\tilde{y}(\cdot) = 0$ on $[U, \infty)$, then $\tilde{y}(\cdot) = y_{1,1}(\cdot)$ in* (1)*, that is, it determines $V_{1,1}(\cdot)$ and*

$$\sigma_1 = \inf\{n \geq 1 : Y_n \geq \tilde{y}(\tau_n)\}.$$

Proof. See Chow [1], pp. 115-117. □

Remark 2.1. $V_{1,1}(0)$ is the optimal mean reward and $\tau_{\sigma_1(0)}$ is the time of optimal selection for the player.

Remark 2.2. Note that

$$V_{1,1}(u) = E(R^1_{\sigma_1(u)}),$$

where $R^1_{\sigma_1(u)}$ is the reward corresponding to the stopping rule $\sigma_1(u)$:

$$R^1_{\sigma_1(u)} = Y_{\sigma_1(u)} r(u + \tau_{\sigma_1(u)}).$$

3 Model of the *m*-person game

In this section we present a generalization of the Elfving stopping time problem to the case of a multi-person, multi-stopping game. Suppose that there are $m > 1$ ordered players who sequentially observe rewards Y_n at times $\tau_n,\ n = 1, 2, ...$ Players' indices $1, 2, ..., m$ correspond to their ordering (ranking) called priority so that 1 refers to the player with the highest priority and m to the lowest one. Each player is allowed to obtain one reward at time of its appearance on the basis of the past and current observations and taking into account the players' priorities. More precisely, Player i, say, who has just decided to make a selection at τ_n gets the reward $Y_n r(\tau_n)$ if and only if he has not obtained any reward before and there is no player with higher priority (lower order) who has also decided to take the current reward. As soon as the player gets the reward he quits the game. The remaining players select rewards in the same manner, their priorities remain as previously.

Now we will define the game in a normal form. As before, Λ is the set of stopping rules with respect to $\{\mathcal{F}_n\}$. Let S_i denote Player i's strategy set, $i = 1, ..., m$. We define $s_i = (s_1^i, s_2^i, ..., s_m^i) \in S_i$ if $s_j^i \in \Lambda$, $j = 1, ..., m$, and $s_1^i \leq s_2^i \leq ... \leq s_m^i$ a.s. To define the players' rewards under the game strategy $s = (s_1, s_2, ..., s_m) \in S_1 \times S_2 \times ... \times S_m$ we will need more notation. Let $\sigma_1 = \min\{s_1^1, ..., s_1^m\}$ determine the first selection at time τ_{σ_1}. The reward X_{σ_1} is obtained by Player L_1, where $L_1 = \min\{j : s_1^j = \sigma_1, j \in \{1, ..., m\}\}$, because of the assumed priorities. The second selection is made at τ_{σ_2}, where $\sigma_2 = \min\{s_2^j : j \neq L_1\}$, by Player L_2, $L_2 = \min\{j : s_k^j = \sigma_2, j \neq L_1\}$. Similarly, we define the k-th selection, $2 < k \leq m$. Let $\sigma_k = \min\{s_k^j : j \neq L_1, L_2, ..., L_{k-1}\}$, $L_k = \min\{j : s_k^j = \sigma_k, j \neq L_1, L_2, ..., L_{k-1}\}$. At time τ_{σ_k} Player L_k obtains the reward X_{σ_k}. Under the game strategy $s = (s_1, ..., s_m)$, the reward for Player j, $j = 1, ..., m$, is

$$R^j(s) = \sum_{k=1}^{m} X_{\sigma_k} \mathbb{I}(L_k = j),$$

and the mean reward is $V^j(s) = E(R^j(s))$, where $\mathbb{I}(A)$ is the indicator function of the event A. Let us recall that the game strategy $\hat{s} = (\hat{s}_1, \hat{s}_2, ...\hat{s}_m) \in \prod_{j=1}^m S_j$ is a Nash equilibrium strategy if for any strategy $s = (s_1, s_2, ..., s_m) \in \prod_{j=1}^m S_j$ we have

$$V^j(\hat{s}) \geq V^j(\hat{s}^{-i}, s_i), \quad \text{for} \quad j = 1, 2, ..., m.$$

Using the results of Stadje [15] on existence of optimal k stopping rules in the Poisson arrival setting and the results of Ferenstein [6], [7] on existence of a Nash equilibrium for stopping games with priorities in the discrete time case we conclude that there exists a Nash equilibrium strategy of the game $(S_1 \times ... \times S_m, V_1, ..., V_m)$. In the next section we describe the optimal behavior of the players corresponding to a Nash equilibrium.

From now on we will denote by $\sigma_{m,k}(u)$, $\tau_{\sigma_{m,k}(u)}$, and $V_{m,k}(u)$, respectively, the optimal stopping rule, the time of the optimal stopping selection and the optimal mean reward of the player with priority k in the case where there are m players in the game which starts at time u, $u \geq 0$. For $m = k = 1$ we have $\sigma_{1,1}(u) = \sigma_1(u)$.

3.1 Optimal mean reward for player with the lowest priority

Let $V_{m,m}(u)$ be the optimal mean reward for the player with the lowest priority, where $m = 2, 3, ...$ Let us note that

$$V_{1,1}(u) \geq V_{2,2}(u) \geq ... \geq V_{m,m}(u). \tag{7}$$

Let us define the rule of the first stopping by one of the players in the m-person game:

$$\sigma_m(u) = \inf\{n \geq 1 : Y_n r(u + \tau_n) \geq V_{m,m}(u + \tau_n)\}, \qquad u \geq 0,$$

which may be rewritten as

$$\sigma_m(u) = \inf\Big\{n \geq 1 : Y_n \geq y_{m,m}(u+\tau_n)\Big\},$$

where we have set

$$y_{m,m}(u) = \frac{V_{m,m}(u)}{r(u)} \qquad \text{if} \quad 0 \leq u < U. \tag{8}$$

If $u \geq U$, then $y_{m,m}(u) = 0$. Let us note that

$$\sigma_1(u) \geq \sigma_2(u) \geq ... \geq \sigma_m(u). \tag{9}$$

The inequalities (7) can be rewritten as follows:

$$y_{1,1}(u) \geq y_{2,2}(u) \geq ... \geq y_{m,m}(u), \tag{10}$$

where

$$y_{i,i}(u) = \frac{V_{i,i}(u)}{r(u)} \qquad \text{for} \quad i = 1, 2, ..., m.$$

Let $\sigma_{m,m}(u)$ be the optimal stopping rule for Player m. If $\sigma_m(u)$ is not optimal for Players $1, ..., m-1$ (it is not profitable for Players $1, ..., m-1$ to stop at $\tau_{\sigma_m(u)}$), then $\sigma_{m,m}(u) = \sigma_m(u)$. On the other hand, if it is profitable for one of Players $1, ..., m-1$ to stop, then Player m cannot stop and he will have the lowest priority in the $m-1$-person game starting at $\tau_{\sigma_m(u)}$ (his optimal stopping rule is greater than $\sigma_m(u)$). Hence, his optimal selection is made according to the stopping rule $\sigma^{\sigma_m}_{m-1,m-1}(u)$. Therefore, his optimal stopping rule is as follows:

$$\sigma_{m,m}(u) = \begin{cases} \sigma_m\,(u) & \text{if} \quad \sigma_{m-1}(u) > \sigma_m(u) \\ \sigma^{\sigma_m}_{m-1,m-1}(u) & \text{if} \quad \sigma_{m-1}(u) = \sigma_m(u). \end{cases} \tag{11}$$

The stopping rule $\sigma^{\sigma_m}_{m-1,m-1}(u)$ will be found by recursion. The following notation will be convenient:

$$\sigma^t_i(u) = \inf\{n > t : Y_n r(u+\tau_n) \geq V_{i,i}(u+\tau_n)\}. \tag{12}$$

$\sigma^t_i(u)$ is the optimal rule of the first selection after time t, $t \geq 0$, in the i-person game. Let

$$A^m_m(u) = \sigma_m(u),$$

$$A^m_k(u) = \inf\{n > A^m_{k+1}(u) : Y_n r(u+\tau_n) \geq V_{k,k}(u+\tau_n)\},$$

for $k = m-1, ..., 1$. We will omit u in $A^m_k(u)$ to simplify the notation. Note that

$$A^m_{m-1} = \sigma^{\sigma_m}_{m-1}(u).$$

Let us denote additionally

$$\sigma_{2,2}^{A_3^m}(u) = \begin{cases} \sigma_2^{A_3^m}(u) & \text{if} \quad \sigma_1^{A_3^m}(u) > \sigma_2^{A_3^m}(u) \\ \sigma_{1,1}^{\sigma_2^{A_3^m}}(u) & \text{if} \quad \sigma_1^{A_3^m}(u) = \sigma_2^{A_3^m}(u), \end{cases}$$

which can be rewritten as follows:

$$\sigma_{2,2}^{A_3^m}(u) = \begin{cases} A_2^m & \text{if} \quad \sigma_1^{A_3^m}(u) > \sigma_2^{A_3^m}(u) \\ \sigma_{1,1}^{A_2^m}(u) & \text{if} \quad \sigma_1^{A_3^m}(u) = \sigma_2^{A_3^m}(u). \end{cases}$$

Let us define by recursion for $k = 2, 3, ..., m-1$ the stopping rules

$$\sigma_{k,k}^{A_{k+1}^m}(u) = \begin{cases} \sigma_k^{A_{k+1}^m}(u) & \text{if} \quad \sigma_{k-1}^{A_{k+1}^m}(u) > \sigma_k^{A_{k+1}^m}(u) \\ \sigma_{k-1,k-1}^{A_k^m}(u) & \text{if} \quad \sigma_{k-1}^{A_{k+1}^m}(u) = \sigma_k^{A_{k+1}^m}(u), \end{cases}$$

which may be rewritten as

$$\sigma_{k,k}^{A_{k+1}^m}(u) = \begin{cases} A_k^m & \text{if} \quad \sigma_{k-1}^{A_{k+1}^m}(u) > \sigma_k^{A_{k+1}^m}(u) \\ \sigma_{k-1,k-1}^{A_k^m}(u) & \text{if} \quad \sigma_{k-1}^{A_{k+1}^m}(u) = \sigma_k^{A_{k+1}^m}(u). \end{cases}$$

The above procedure allows us to determine $\sigma_{m-1,m-1}^{\sigma_m}(u)$. Let us note that

$$\sigma_{m-1}(u) = \sigma_m(u) \quad \text{iff} \quad Y_{\sigma_m(u)} r(u + \tau_{\sigma_m(u)}) \geq V_{m-1,m-1}(u + \tau_{\sigma_m(u)}),$$
$$\sigma_{m-1}(u) > \sigma_m(u) \quad \text{iff} \quad Y_{\sigma_m(u)} r(u + \tau_{\sigma_m(u)}) < V_{m-1,m-1}(u + \tau_{\sigma_m(u)}).$$

We set

$$f_{m,u}(v) = P(\tau_{\sigma_m(u)} > v).$$

Considerations similar to those in Elfving [3] yields

$$f_{m,u}(v) = \exp\Big[- \int_u^{u+v} G(y_{m,m}(v'))dv' \Big]. \tag{13}$$

Now we will derive the equation for the optimal mean reward of Player m,

$$V_{m,m}(u) = E(R^m_{\sigma_{m,m}(u)}),$$

where $R^m_{\sigma_{m,m}(u)}$ is the reward of Player m under the optimal stopping rule $\sigma_{m,m}(u)$, that is,

$$R^m_{\sigma_{m,m}(u)} = \begin{cases} Y_{\sigma_m(u)} r(u + \tau_{\sigma_m(u)}) & \text{if} \quad \sigma_{m-1}(u) > \sigma_m(u) \\ V_{m-1,m-1}(u + \tau_{\sigma_m(u)}) & \text{if} \quad \sigma_{m-1}(u) = \sigma_m(u). \end{cases}$$

We have

$$V_{m,m}(u) = E\Big[Y_{\sigma_m(u)} r(u+\tau_{\sigma_m(u)})\mathbb{I}(Y_{\sigma_m(u)} < y_{m-1,m-1}(u+\tau_{\sigma_m(u)}))$$

$$+V_{m-1m-,1}(u+\tau_{\sigma_m(u)})\mathbb{I}(Y_{\sigma_m(u)} \geq y_{m-1,m-1}(u+\tau_{\sigma_m(u)}))\Big]$$

$$= E\Bigg[E\Big[Y_{\sigma_m(u)} r(u+\tau_{\sigma_m(u)})\mathbb{I}(Y_{\sigma_m(u)} < y_{m-1,m-1}(u+\tau_{\sigma_m(u)}))$$

$$+V_{m-1,m-1}(u+\tau_{\sigma_m(u)})\mathbb{I}(Y_{\sigma_m(u)} \geq y_{m-1,m-1}(u+\tau_{\sigma_m(u)}))|\tau_{\sigma_m(u)}, \sigma_m(u)\Big]\Bigg].$$

Let us note that the conditional distribution of $Y_{\sigma_m(u)}$ given $\tau_{\sigma_m(u)}, \sigma_m(u)$ is the same as the conditional distribution of Y given $\{Y \geq y_{m,m}(u+\tau_{\sigma_m(u)})\}$, where $Y, \tau_{\sigma_m(u)}$ are independent and Y has the distribution function F. Hence, using (10) we have

$$V_{m,m}(u) = \int_0^{U-u} \frac{r(u+v)\int_{y_{m,m}(u+v)}^{y_{m-1,m-1}(u+v)} ydF(y)}{G(y_{m,m}(u+v))}(-f'_{m,u}(v))dv$$

$$+\int_0^{U-u} \frac{V_{m-1,m-1}(u+v)\int_{y_{m-1,m-1}(u+v)\vee y_{m,m}(u+v)}^{\infty} dF(y)}{G(y_{m,m}(u+v))}(-f'_{m,u}(v))dv.$$

Set

$$E_{m,m}(u,\tilde{v}) = \exp\Big[-\int_u^{\tilde{v}} G(y_{m,m}(v'))dv'\Big].$$

Using the formula (13) we have

$$V_{m,m}(u) = \int_0^{U-u} r(u+v)\int_{y_{m,m}(u+v)}^{y_{m-1,m-1}(u+v)} ydF(y)E_{m,m}(u,u+v)dv$$

$$+\int_0^{U-u} V_{m-1,m-1}(u+v)\int_{y_{m-1,m-1}(u+v)}^{\infty} dF(y)E_{m,m}(u,u+v)dv.$$

Now a change of variable gives

$$V_{m,m}(u) = \int_u^U r(\tilde{v})[H(y_{m,m}(\tilde{v})) - H(y_{m-1,m-1}(\tilde{v}))]E_{m,m}(u,\tilde{v})d\tilde{v}$$

$$+\int_u^U V_{m-1,m-1}(\tilde{v})[1-F(y_{m-1,m-1}(\tilde{v}))]E_{m,m}(u,\tilde{v})d\tilde{v}. \tag{14}$$

Differentiating (14) with respect to u and using (8) we obtain the following equation:

$$\frac{d}{du}V_{m,m}(u) = \frac{d}{du}[r(u)y_{m,m}(u)]$$

$$= r(u)\Big(G(y_{m,m}(u))y_{m,m}(u) - H(y_{m,m}(u)) + H(y_{m-1,m-1}(u))\Big) \\ -V_{m-1,m-1}(u)[1 - F(y_{m-1,m-1}(u))], \qquad (15)$$

where $0 \le u < U$.

Remark 3.1. Equation (15) is a first-order non-linear differential equation which can be solved numerically with the condition

$$r(u)y_{m,m}(u) \to 0 \quad \text{as} \quad u \to U.$$

Remark 3.2. $\sigma_{m,m}(0)$ is the optimal stopping rule for Player m, $\tau_{\sigma_{m,m}(0)}$ is the time of his optimal selection and $V_{m,m}(0)$ is his optimal mean reward, where $m = 2, 3, ...$

As in the Elfving problem, we get the solution of equation (14) from (15) because of the following theorem.

Theorem 3.1. *(i) A piecewise continuous function $\tilde{y}(\cdot)$ satisfies* (14) *if and only if $\tilde{y}(\cdot)$ satisfies* (15) *and $r(u)\tilde{y}(u) \to 0$ as $u \to U$.*
(ii) If $\tilde{y}(\cdot)$ satisfies (14) *on $[0, U)$ and $\tilde{y}(\cdot) = 0$ on $[U, \infty)$, then $\tilde{y}(\cdot) = y_{m,m}(\cdot)$ in* (8)*, that is, it determines $V_{m,m}(\cdot)$ and*

$$\sigma_m(0) = \inf\{n \ge 1 : Y_n \ge \tilde{y}(\tau_n)\}.$$

3.2 Optimal mean rewards of remaining players

We will show that $\sigma_{m,k}(u) = \sigma_{k,k}(u)$ for each m, where k is the player's priority (k-th rank, where 1 refers to the highest priority), $k = 1, 2, ..., m - 1$. As a consequence, we will have $V_{m,k}(u) = V_{k,k}(u)$. The mean reward $V_{m,m}(u)$ is found separately.

We will show by induction with respect to m, $m = 2, 3, ...$, that $\sigma_{m,k}(u) = \sigma_{k,k}(u)$ and $V_{m,k}(u) = V_{k,k}(u)$ for $k = 1, ..., m - 1$.

The statements of the theorems below are intuitively clear. Player k's decision is not influenced by decisions of players with lower priorities. In the case of discrete time and finite horizon and general reward structures these theorems were proved in Enns and Ferenstein [5]. Sakaguchi [13] considered a game version of the no-information best-choice problem, with players' priorities and discrete observation time. The theorems below generalize the results of Enns and Ferenstein [5] and Sakaguchi [13] to the case of Poisson-arriving offers and a random number of offers, which is a consequence of introducing of a discount function.

The proofs of the following theorems appear in the Appendix.

Theorem 3.2. *The optimal stopping rule for Player 1 in the m-person game starting at $u \ge 0$, $m = 2, 3, ...$, is the optimal stopping rule of the one-player game, that is,*

$$\forall u \ge 0 \quad \sigma_{m,1}(u) = \sigma_1(u).$$

Theorem 3.3. *The optimal stopping rule for Player k, $k = 2, 3, ..., m - 1$, in the m-person game starting at $u \geq 0$, $m = 3, 4, ...$, is the optimal stopping rule of the player with the lowest priority in the k-person game:*

$$\forall u \geq 0 \quad \sigma_{m,k}(u) = \sigma_{k,k}(u).$$

As a result, the optimal mean reward of Player k, $k = 1, 2, ..., m - 1$, in the m-person game is

$$V_{m,k}(u) = E(R^k_{\sigma_{m,k}(u)}) = E(R^k_{\sigma_{k,k}(u)}) = V_{k,k}(u),$$

where $R^k_{\sigma_{m,k}(u)}$ is the reward of Player k,

$$R^k_{\sigma_{m,k}(u)} = R^k_{\sigma_{k,k}(u)},$$

where

$$R^k_{\sigma_{k,k}(u)} = \begin{cases} Y_{\sigma_k(u)} r(u + \tau_{\sigma_k(u)}) & \text{if} \quad \sigma_1(u) \geq ... \geq \sigma_{k-1,k-1}(u) > \sigma_{k,k}(u) \\ V_{k-1,k-1}(u + \tau_{\sigma_k(u)}) & \text{if} \quad \sigma_1(u) \geq ... \geq \sigma_{k-1,k-1}(u) = \sigma_{k,k}(u). \end{cases}$$

Remark 3.3. From Theorems 3.2 and 3.3 we can get by recursion the optimal mean rewards for Player k, $k = 1, ..., m - 1$, in the m-person game.

Remark 3.4. $\sigma_{m,k}(0)$ is the optimal stopping rule for Player k, $\tau_{\sigma_{m,k}(0)}$ is the time of his optimal selection and $V_{m,k}(0)$ is his optimal mean reward, where $k = 1, 2, ..., m$.

4 Conclusion and example

4.1 Non-homogeneous Poisson process

Instead of the game with a homogeneous Poisson process we can consider the model with a non-homogeneous Poisson process with intensity function $p(u)$. (See Chow [1], pp. 113-114.) It is sufficient to apply the transformation of the time scale

$$\tilde{u} = \int_0^u p(u')du',$$

which is a $1 - 1$ map of the positive u-axis into the finite or infinite interval

$$0 \leq \tilde{u} \leq \tilde{U} = \int_0^\infty p(u')du'.$$

4.2 Example

We will consider the three-person game. Let $\{Y_i\}$ be the sequence of iid r.v's with the exponential distribution with mean 1. Assume that $r(u) = 1$ for $u \in [0, U)$ and 0 otherwise. Then, $F(y) = 1 - \exp(-y)$, $G(y) = \exp(-y)$, $H(y) = (y + 1)\exp(-y)$. The differential equation (6) becomes

$$\frac{dy_{1,1}(u)}{du} = -\exp(-y_{1,1}(u)). \tag{16}$$

Solving Eq. (16) with the boundary condition $y_{1,1}(U) = 0$ we get $V_{1,1}(u) = y_{1,1}(u) = \ln(1+U-u)$. Hence, the optimal mean reward is $V_{1,1}(0) = \ln(1+U)$. Equation (15) with $V_{1,1}(u)$ as above becomes

$$\frac{d}{du}y_{2,2}(u) = -\exp(-y_{2,2}(u)) + \frac{1}{1+U-u}.$$

Hence, using the boundary condition $y_{2,2}(U) = 0$ we get

$$y_{2,2}(u) = \ln\left(\frac{2+2U+U^2-2u-2Uu+u^2}{2(1+U-u)}\right)$$

and the optimal reward

$$V_{2,2}(0) = y_{2,2}(0) = \ln\left(\frac{2+2U+U^2}{2(1+U)}\right).$$

Using Theorem 3.2 for $m = 2$ and $u = 0$ we get $V_{2,1}(0) = V_{1,1}(0) = y_{1,1}(0) = \ln(1+U)$. In a similar way, we obtain $V_{3,3}(u)$. Note that $y_{3,3}(u)$ satisfies the equation

$$\frac{d}{du}y_{3,3}(u) = -\exp(-y_{3,3}(u)) + \frac{2(1+U-u)}{2+2U+U^2-2u-2Uu+u^2}$$

with the boundary condition $y_{3,3}(U) = 0$. Hence,

$$y_{3,3}(u) = \ln\left(\frac{-6u-6Uu-3U^2u+3u^2+3Uu^2-u^3+6+6U+3U^2+U^3}{3(2+2U+U^2-2u-2Uu+u^2)}\right).$$

Therefore, the optimal mean reward for Player 3 is

$$V_{3,3}(0) = y_{3,3}(0) = \ln\left(\frac{6+6U+3U^2+U^3}{3(2+2U+U^2)}\right).$$

Using Theorems 3.2 and 3.3 for $m = 3$ and $u = 0$, we get the optimal mean rewards for Players 1 and 2, respectively:

$$V_{3,1}(0) = V_{1,1}(0) = \ln(1+U),$$

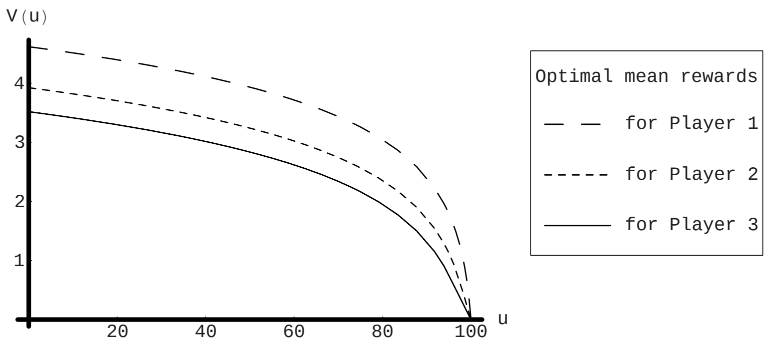

Figure 1: The optimal mean rewards for the players in the three-person game.

$$V_{3,2}(0) = V_{2,2}(0) = \ln\left(\frac{2+2U+U^2}{2(1+U)}\right).$$

The graphs of the optimal mean rewards $V_{3,i}(u)$ are displayed in Fig. 1.

5 Appendix

The proofs below use induction.

5.1 Proof of Theorem 3.2

Proof. 1. Let $m = 2$. It is easy to see that the player with the highest priority will stop according to the rule $\sigma_2(u)$ (at time $\tau_{\sigma_2(u)}$) if $\sigma_1(u) = \sigma_2(u)$ and according to the rule $\sigma_1^{\sigma_2}(u)$ (at time $\tau_{\sigma_1^{\sigma_2}(u)}$) if $\sigma_1(u) > \sigma_2(u)$. Hence, his optimal stopping rule is

$$\sigma_{2,1}(u) = \begin{cases} \sigma_2(u) & \text{if} \quad \sigma_1(u) = \sigma_2(u) \\ \sigma_1^{\sigma_2}(u) & \text{if} \quad \sigma_1(u) > \sigma_2(u). \end{cases} \tag{17}$$

Therefore, for each $u \geq 0$, we have $\sigma_{2,1}(u) = \sigma_1(u)$.
2. Let $m > 2$ and assume that

$$\forall u \geq 0 \quad \sigma_{m-1,1}(u) = \sigma_{m-2,1}(u) = ... = \sigma_{1,1}(u) = \sigma_1(u).$$

We will show $\sigma_{m,1}(u) = \sigma_1(u)$. By the definition (12) we have

$$\sigma_1^{\sigma_m}(u) = \sigma_1(u) \quad \text{if} \quad \sigma_1(u) > \sigma_m(u).$$

From the induction assumption we have

$$\sigma_{m-1,1}^{\sigma_m}(u) = \sigma_1^{\sigma_m}(u). \tag{18}$$

Hence,

$$\sigma_{m-1,1}^{\sigma_m}(u) = \sigma_1^{\sigma_m}(u) = \sigma_1(u) \quad \text{if} \quad \sigma_1(u) > \sigma_m(u), \tag{19}$$

where $\sigma_{m-1,1}^{\sigma_m}(u)$ is the optimal stopping rule greater than $\sigma_m(u)$ of the player with the highest priority in the $m-1$-person game. Let us note that two cases are possible:

(i) $\sigma_1(u) = \sigma_m(u)$,
(ii) $\sigma_1(u) > \sigma_m(u)$.

In case (i), $\sigma_m(u)$, the first selection rule, coincides with the optimal selection rule in the one-player game; so it is clear that Player 1 makes the first selection. In case (ii), it is not profitable for Player 1 to stop, so after the first selection according to $\sigma_m(u)$ (after time $\tau_{\sigma_m(u)}$) we have the $m-1$-person game and Player 1 keeps the highest priority. Hence,

$$\sigma_{m,1}(u) = \begin{cases} \sigma_m(u) & \text{if} \quad \sigma_1(u) = \sigma_m(u) \\ \sigma_{m-1,1}^{\sigma_m}(u) & \text{if} \quad \sigma_1(u) > \sigma_m(u). \end{cases}$$

Hence, using (19) we have $\sigma_{m,1}(u) = \sigma_1(u)$. □

5.2 Proof of Theorem 3.3

Proof. 1. Proof for $m = 3$. We will find the optimal strategy $\sigma_{3,2}(u)$ of Player 2. Three cases are possible:

(i) $\sigma_1(u) > \sigma_2(u) = \sigma_3(u)$,
(ii) $\sigma_1(u) = \sigma_2(u) = \sigma_3(u)$,
(iii) $\sigma_1(u) \geq \sigma_2(u) > \sigma_3(u)$.

In case (i), it is not profitable for Player 1 to stop according to the rule $\sigma_3(u)$ while for Players 2 and 3 it is profitable to stop, so Player 2 will stop according to this rule. Hence, $\sigma_{3,2}(u) = \sigma_3(u)$ in this situation. In case (ii), it is profitable for each player to stop according to the rule $\sigma_3(u)$, so when Player 1 quits the game, Player 2 will be the most privileged player in the two-person game. Hence, $\sigma_{3,2}(u) = \sigma_{2,1}^{\sigma_3}(u)$. In case (iii), it is profitable only for Player 3 to stop. Hence, Player 2 will have the lowest priority in the two-person game after the selection made by Player 3. Hence, $\sigma_{3,2}(u) = \sigma_{2,2}^{\sigma_3}(u)$. As a consequence, we have

$$\sigma_{3,2}(u) = \begin{cases} \sigma_3(u) & \text{if} \quad \sigma_1(u) > \sigma_2(u) = \sigma_3(u) \\ \sigma_{2,1}^{\sigma_3}(u) & \text{if} \quad \sigma_1(u) = \sigma_2(u) = \sigma_3(u) \\ \sigma_{2,2}^{\sigma_3}(u) & \text{if} \quad \sigma_1(u) \geq \sigma_2(u) > \sigma_3(u). \end{cases}$$

Using (18) we have

$$\sigma_{3,2}(u) = \begin{cases} \sigma_2(u) & \text{if} \quad \sigma_1(u) > \sigma_2(u) = \sigma_3(u) \\ \sigma_1^{\sigma_2}(u) & \text{if} \quad \sigma_1(u) = \sigma_2(u) = \sigma_3(u) \\ \sigma_{2,2}^{\sigma_3}(u) & \text{if} \quad \sigma_1(u) \geq \sigma_2(u) > \sigma_3(u). \end{cases}$$

Applying the formula (11) for $m = 2$ we have

$$\sigma_{3,2}(u) = \begin{cases} \sigma_{2,2}(u) & \text{if} \quad \sigma_1(u) \geq \sigma_2(u) = \sigma_3(u) \\ \sigma_{2,2}^{\sigma_3}(u) & \text{if} \quad \sigma_1(u) \geq \sigma_2(u) > \sigma_3(u). \end{cases}$$

Hence, $\sigma_{3,2}(u) = \sigma_{2,2}(u)$.
2. Let $m > 3$ and assume that

$$\forall u \geq 0 \quad \sigma_{m-1,k}(u) = \sigma_{m-2,k}(u) = ... = \sigma_{k,k}(u) \quad \text{for} \quad k = 2, 3, ..., m-2.$$

3. We will show that the optimal strategy $\sigma_{m,k}(u)$ of Player k is $\sigma_{m,k}(u) = \sigma_{k,k}(u)$ for $k = 2, 3, ..., m-1$. Three cases are possible:

(i) $\sigma_1(u) \geq ... \geq \sigma_{k-1}(u) > \sigma_k(u) = ... = \sigma_m(u)$,
(ii) $\sigma_1(u) \geq ... \geq \sigma_{k-1}(u) = \sigma_k(u) = ... = \sigma_m(u)$,
(iii) $\sigma_1(u) \geq ... \geq \sigma_k \geq ... \geq \sigma_{i-1}(u) > \sigma_i(u) = ... = \sigma_m(u)$ for some $i = k+1, ..., m$.

In case (i), it is not profitable for Players 1,..., $k-1$ to stop according to the rule $\sigma_m(u)$. For Players k,..., m it is profitable to stop, so Player k will stop according to this rule. Hence, $\sigma_{m,k}(u) = \sigma_m(u)$ in this situation. In case (ii), it is profitable for each Player $k-1, ..., m$ to stop according to $\sigma_m(u)$, so when one of Players $1, ..., k-1$ quits the game, Player k will have priority $k-1$ in the $m-1$-person game. Hence, $\sigma_{m,k}(u) = \sigma_{m-1,k-1}^{\sigma_m}(u)$. In case (iii), it is profitable for some Player i to stop. Hence, Player k will have priority k in the $m-1$-person game after the selection made by Player i. Hence, $\sigma_{m,k}(u) = \sigma_{m-1,k}^{\sigma_m}(u)$ in this situation. As a consequence, we have

$$\sigma_{m,k}(u) = \begin{cases} \sigma_m(u) & \text{if } \sigma_1(u) \geq ... \geq \sigma_{k-1}(u) > \sigma_k(u) = ... = \sigma_m(u) \\ \sigma_{m-1,k-1}^{\sigma_m}(u) & \text{if } \sigma_1(u) \geq ... \geq \sigma_{k-1}(u) = \sigma_k(u) = ... = \sigma_m(u) \\ \sigma_{m-1,k}^{\sigma_m}(u) & \text{if } \sigma_1(u) \geq ... \geq \sigma_{i-1}(u) > \sigma_i(u) = ... = \sigma_m(u) \\ & \text{for some } i = k+1, ..., m. \end{cases}$$

In case (ii), we have $\sigma_{m-1,k-1}^{\sigma_m}(u) = \sigma_{m-1,k-1}^{\sigma_k}(u)$. From the induction assumption $\forall u \geq 0$ $\sigma_{m-1,k}(u) = \sigma_{k,k}(u)$ we have $\sigma_{m-1,k}^{\sigma_m}(u) = \sigma_{k,k}^{\sigma_m}(u)$. Additionally, we have

$$\sigma_{k,k}^{\sigma_m}(u) = \sigma_{k,k}(u) \quad \text{if} \quad \sigma_k(u) > \sigma_m(u).$$

Finally, we obtain

$$\forall u \geq 0 \quad \sigma_{m-1,k}^{\sigma_m}(u) = \sigma_{k,k}^{\sigma_m}(u) = \sigma_{k,k}(u) \quad \text{if} \quad \sigma_k(u) > \sigma_m(u).$$

Hence,

$$\sigma_{m,k}(u) = \begin{cases} \sigma_k(u) & \text{if } \sigma_1(u) \geq ... \geq \sigma_{k-1}(u) > \sigma_k(u) = ... = \sigma_m(u) \\ \sigma_{m-1,k-1}^{\sigma_k}(u) & \text{if } \sigma_1(u) \geq ... \geq \sigma_{k-1}(u) = \sigma_k(u) = ... = \sigma_m(u) \\ \sigma_{k,k}(u) & \text{if } \sigma_1(u) \geq ... \geq \sigma_{i-1}(u) > \sigma_i(u) = ... = \sigma_m(u) \\ & \text{for some } i = k+1, ..., m. \end{cases}$$

From the induction assumption $\forall u \; \sigma_{m-1,k-1}(u) = \sigma_{k-1,k-1}(u)$ we get

$$\forall u \quad \sigma^{\sigma_k}_{m-1,k-1}(u) = \sigma^{\sigma_k}_{k-1,k-1}(u).$$

Hence,

$$\sigma_{m,k}(u) = \begin{cases} \sigma_k(u) & \text{if} \quad \sigma_1(u) \geq ... \geq \sigma_{k-1}(u) > \sigma_k(u) = ... = \sigma_m(u) \\ \sigma^{\sigma_k}_{k-1,k-1}(u) & \text{if} \quad \sigma_1(u) \geq ... \geq \sigma_{k-1}(u) = \sigma_k(u) = ... = \sigma_m(u) \\ \sigma_{k,k}(u) & \text{if} \quad \sigma_1(u) \geq ... \geq \sigma_{i-1}(u) > \sigma_i(u) = ... = \sigma_m(u) \\ & \text{for some} \quad i = k+1, ..., m. \end{cases}$$

Using the definition (11) for $m = k$ we obtain

$$\sigma_{m,k}(u) = \begin{cases} \sigma_{k,k}(u) & \text{if} \quad \sigma_1(u) \geq ... \geq \sigma_{k-1}(u) \geq \sigma_k(u) = ... = \sigma_m(u) \\ \sigma_{k,k}(u) & \text{if} \quad \sigma_1(u) \geq ... \geq \sigma_{i-1}(u) > \sigma_i(u) = ... = \sigma_m(u) \\ & \text{for some} \quad i = k+1, ..., m. \end{cases}$$

Therefore, $\sigma_{m,k}(u) = \sigma_{k,k}(u)$ for $k = 2, 3, ..., m-1$. □

Acknowledgements

We are grateful to the referees for helpful comments and bibliographical refinements.

References

[1] Chow Y. S., Robbins H., Siegmund D., *Great Expectations: The Theory of Optimal Stopping*, Houghton Mifflin Company, Boston, 1971.

[2] Dixon M. T., *Equilibrium points for three games based on the Poisson process*, J. Appl. Prob. **30** (1993), No. 3, 627-638.

[3] Elfving G., *A persistency problem connected with a point process*, J. Appl. Prob. **4** (1967), 77-89.

[4] Enns E.G., Ferenstein E.Z., *A competitive best-choice problem with Poisson arrivals*, J. Appl. Prob. **27** (1990), 333-342.

[5] Enns E.G., Ferenstein E.Z., *On a multi-person time-sequential game with priorities*, Sequent. Anal. **6** (1987), No. 3, 239-256.

[6] Ferenstein E.Z., *A variation of the Dynkin's stopping game*, Math. Japonica **38** (1993), 371-379.

[7] Ferenstein E.Z., *On infinite-horizon multi stage stopping games with priorities*, in A. Wieczorek, M. Malawski and A. Wiszniewska-Matyszkiel (eds.), *Game Theory and Mathematical Economics*, Banach Center Publications **71**, Warsaw, (2006), 93-102.

[8] Kurano M., Nakagami J., Yasuda M., *A Game Variant of Stopping Problem on Jump Processes with a Monotone Rule*, in J. A. Filar, V. Gaitsgory, K. Mizumaki (eds.) *Advances in Dynamic Games and Applications*, Birkhäuser, Boston, (2000), 257-266.

[9] Porosiński Z., Szajowski K., *On continuous-time two person full-information best choice problem with imperfect observation*, Sankhyā: Ind. J. Statist., **58** (1996), Series A, 186-193.

[10] Saario V., Sakaguchi M., *Multistop best choice games related to the Poisson process*, Math. Japonica **37** (1992), 41-51.

[11] Sakaguchi M., *Best-choice problems with random priority on a two-Poisson stream*, Math. Japonica **36** (1991), No. 4, 731-745.

[12] Sakaguchi M., *Better-than-opponent: A stopping game for Poisson-arriving offers*, Sci. Math. Japonica **56** (2002), No. 3, 457-473.

[13] Sakaguchi M., *Multiperson multilateral secretary problems*, Math. Japonica **34** (1989), No. 3, 459-473.

[14] Siegmund D., *Some problems in the theory of optimal stopping*, Ann. Math. Statist. **38** (1967), 1627-1640.

[15] Stadje W., *An Optimal k-Stopping Problem for the Poisson process*, in P. Bauer, F. Konecny and W. Wetz (eds.) *Proceedings of the 6th Pannonian Symposium on Mathematical Statistics* **B**, Reidel, Dordrecht, (1986), 231-244.

Cooperative Strategies in Stopping Games

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

Diana Cierpiał
Department of Mathematics and Computer Science
Wrocław University of Technology
Wrocław, POLAND

Abstract

This article considers a cooperative approach to 2-player stopping games. It is assumed that the players cooperate in such a way as to maximise the sum of their expected payoffs. Hence, the value of the game to a coalition of both players is given by the optimal expected reward in an auxiliary problem, which involves the double stopping of a Markov chain by an individual. The value of the game is defined using a recursive procedure based on the form of the subgames played at each moment. The concept of Shapley value is used to define the value of a subgame. It is shown that, in games where one player always has priority, the Shapley value of the cooperative game is simply the unique Nash value of the game. In games involving random priority, players must coordinate their actions and use side payments to achieve the Shapley value. The side payments and coordination may be described in a preplay agreement made between the players. Aspects of the dynamic rationality of such agreements are considered using the concept of subgame consistency. An example based on the job search problem is given.

Key words. Stopping game, cooperative game, Shapley value, subgame consistency.

AMS Subject Classifications. Primary 91A25; Secondary 91A12, 91A50.

1 Introduction

The theory of stopping games was initiated by Dynkin [5], who considered a zero-sum game in which two players observe a sequence of N objects whose values $X_1, X_2, \ldots, X_N$ are i.i.d. random variables. The n-th object is observed at moment n ($n = 1, 2, \ldots, N$). It is assumed that N is finite. Each player may obtain at most one object and an object can be accepted only at the moment of its appearance. Dynkin assumed that one player could stop at odd moments and the other at even moments. The class of pure stopping times is broad enough to define equilibria

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_21,

in such games. The theory of such games has been extended to models of non-zero-sum games in which players simultaneously observe the sequence (see, e.g., Ohtsubo [11], Ferenstein [6]). When such models are considered, it is necessary to resolve the problem of which player obtains an object when more than one wishes to accept it. This led to discussion of the concept of the priority of a player (see Fushimi [8] and Szajowski [18]).

In games with random priority there may be a multitude of equilibria and many of these equilibria may lead to a specific criterion being satisfied. Szajowski [18] presents examples of such equilibria in a 2-player stopping game where the criterion for choosing an equilibrium is to equalise the payoffs of the players. The set of pure stopping times is not rich enough to describe all the Nash equilibria in stopping games. For this reason, Yasuda [19] introduced the notion of randomised stopping times.

When communication is possible, but side payments are not, the concept of a correlated equilibrium is a suitable concept to define a solution to a game. Correlated equilibria in matrix games were introduced by Aumann [2]. Ramsey and Szajowski [15], [16] introduced the concept of correlated stopping times and correlated equilibria in 2-player stopping games. Players can use communication and signals (for example, the result of a coin toss) to correlate their actions. Aumann [2] showed that the set of Nash equilibria of a game is a subset of the set of correlated equilibria. Hence, if the problem of selecting a Nash equilibrium exists, then the problem of selecting a correlated equilibrium also exists. However, the concept of correlated equilibria assumes that communication between the players is possible. Thus, the problem of equilibrium selection can be solved by the players adopting some criterion for choosing an equilibrium.

When side payments between players are possible, we should consider the game as a cooperative game. A preplay agreement defining the actions to be taken and the side payments to be made in each possible state will be called a cooperative strategy. It is assumed that when the players form a coalition they maximise the sum of their expected payoffs. It follows that this value is given by the solution of a stopping problem in which the coalition may choose up to two objects from a sequence. This will be called the auxiliary problem. The agreement is used to split the payoffs obtained in a way that is seen to be fair by both players. The concept of Shapley value will be used to define a "fair split". The preplay agreement is used to assign each of the two objects taken to one the players. The payoff of an individual is the value of the object obtained plus the side payment obtained from the other player (or minus this payment if a player must pay the other player).

The concept of a cooperative game and the value of a cooperative game was introduced by Shapley [17]. His initial concept of the value of such a game has been extended by, e.g., assuming the players have different bargaining strengths and can choose their level of commitment to a coalition (see, e.g., Amer and Carreras [1]). Other concepts of solution have also been suggested, for example the Banzhaf value [3] and the solidarity value introduced by Nowak and Radzik [10]. For applications of cooperative game theory to dynamic games see, Filar and Petrosjan [7], Petrosjan et al. [13] (their approach is slightly different, since coalition forming is considered as a dynamic process, in this article coalition forming is assumed to constitute a preplay stage of a dynamic game), and Yeung and Petrosjan [20].

This paper is intended as a step in developing concepts of cooperative solutions to stopping games. However, many of the concepts used can be applied to dynamic games in general. One obvious need for future development is the application of such ideas to multi-player games. The layout of the article is as follows. Section 2 outlines the model, describes some possible formulations of the game as a cooperative game, and derives the Shapley value. Section 3 is devoted to cooperative strategies. Such strategies are defined by a preplay agreement between the players which describes the actions to be taken in each possible state and the side payments made between the players. A cooperative strategy is derived which satisfies the following two conditions:
1) When the players follow this cooperative strategy, the vector of their expected payoffs is equal to the Shapley value.
2) A cooperative solution should be subgame consistent (see Yeung and Petrosjan [21]). That is to say that: (a) given both players are still searching and n objects are still to appear then any renegotiation would lead to the same cooperative strategy being followed; and (b) there can be no possible subgame in which the expected reward of either player at the cooperative solution is less than his/her expected reward at the mixed equilibrium solution.
When a game is not specifically defined as a cooperative game, it is necessary to define the characteristic function. In the case of matrix games, the payoff to a coalition $\mathcal{S}$ is normally defined to be the minimax payoff of coalition $\mathcal{S}$ in the two-player game in which the players are $\mathcal{S}$ and $\mathcal{S}^c$ (where ^{c} denotes the complement of a set). However, different definitions may be appropriate for different games (see, e.g., Petrosjan and Zaccour [14]). Three approaches are suggested here. Section 4 presents an example based on the job search problem.

2 Formulation of a 2-player stopping game as a cooperative game and its Shapley value

We consider the following 2-person stopping game. The players simultaneously observe a sequence of objects of values $X_1, X_2, \ldots, X_N$ defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with state space $(\mathbb{E}, \mathcal{B})$. Define $\mathcal{F}_i = \sigma(X_1, X_2, \ldots, X_i)$ to be the σ-field generated by the random variables $X_1, X_2, \ldots, X_i$. It is assumed that the distribution of X_i is continuous. The i-th object appears at moment i. A player can obtain at most one object and an object can be obtained only at the moment of its appearance.

If only one player wishes to accept an object (denoted as the action s), then that player obtains the object and stops observing the sequence. The remaining player (who chooses action c) is free to continue observing the sequence. Suppose both players are still searching and both wish to accept the object appearing at moment i. A random lottery is used to assign this object to one of the players and the other player continues observing the sequence. In this case, it is assumed that Player 1 obtains the object with probability α_i, $0 < \alpha_i < 1$. We will refer to α_i as the priority of Player 1 at moment i. Let $(\beta_1, \beta_2, \ldots, \ldots, \beta_N)$ be a vector of independent random variables from the uniform distribution on $[0, 1]$, which is independent of $(X_1, X_2, \ldots, X_N)$. This vector defines the result of any necessary random lottery, i.e., when both players wish to accept an object, Player 1 obtains it iff $\beta_i < \alpha_i$.

Since utility is transferable, it may be assumed that X_i is the utility of object i to the players and that X_i is a finite, non-negative random variable. The payoff of a player when he does not accept any object is defined to be 0.

We now define the auxiliary problem. In such a problem a lone player (who may be understood as a coalition formed by both players) is allowed to sequentially choose at most two objects from the sequence without recall. Suppose the i-th object is the first to be accepted. The problem then reduces to a restricted problem in which one object may be accepted from a sequence of objects of values $X_{i+1}, X_{i+2}, \ldots X_n$. Since the payoff is given by the sum of the objects taken, the player should simply maximise the expected value of the second object taken. It follows that the value of the first object taken only influences the optimal strategy through its effect on F_j, for $j \geq i$. In particular, if the values of the objects are independent, then the optimal strategy in this reduced problem does not depend on the value of first object taken. Define $u_{j,n}$ to be the optimal expected payoff when at most j objects can be accepted from n remaining objects. The optimality equations are given by

$$u_{1,n+1} = E[\max\{X_{N-n}, u_{1,n}\}]; \quad u_{2,n+1} = E[\max\{X_{N-n} + u_{1,n}, u_{2,n}\}], \quad (1)$$

where expectations are taken with respect to $\mathcal{F}_{N-n-1}$. It should be noted that this value function is $\mathcal{F}_{N-n-1}$-measurable. However, for notational ease this dependency is not reflected in the notation used. This is true of all the value functions considered. If the player has just one choice remaining, when n objects are still to appear, an object of value x should be accepted iff $x \geq u_{1,n}$. If the player has two choices remaining, such an object should be accepted iff $x \geq u_{2,n} - u_{1,n}$.

Since $u_{1,0} = u_{2,0} = 0$, these payoff functions can be calculated by recursion. The expected value of the cooperative game to a coalition formed by the 2 players with n moments remaining, is $u_{2,n}$. It is intuitively clear that $u_{2,n} \leq 2u_{1,n}$ (this can be proven by induction).

Define the Shapley value of the game when n objects are yet to appear and neither player has accepted an object to be $(w_{1,n}, w_{2,n})$. Let $\{v_n(\mathcal{S}; x)\}_{\mathcal{S}=\{1\},\{2\},\{1,2\}}$ be the characteristic function of the cooperative game derived from the subgame played when an object of value x appears and n objects are yet to appear. Since a coalition of both players maximises the sum of the payoffs to the players, we have

$$v_n(\{1, 2\}; x) = \max\{w_{1,n} + w_{2,n}, x + u_{1,n}\}. \quad (2)$$

In order to define $v_n(\{1\};x)$ and $v_n(\{2\};x)$, it is assumed that if the two players do not cooperate in the final n moments, then they play the mixed equilibrium in the non-cooperative n-stage stopping problem. Let $(z_{1,n}, z_{2,n})$ be the value of the game where n objects are yet to appear and neither player has accepted an object at this mixed equilibrium. This value can be calculated recursively. The value of the $n+1$-stage stopping game at this mixed equilibrium is the expected value of the following subgame played when an object of value x appears and n objects are still to appear.

$$\begin{array}{c} \\ s \\ c \end{array} \begin{array}{c} \begin{array}{cc} \text{s} \hspace{14em} & \text{c} \end{array} \\ \left(\begin{array}{cc} (\alpha_{N-n}x + (1-\alpha_{N-n})u_{1,n}, \alpha_{N-n}u_{1,n} + (1-\alpha_{N-n})x) & (x, u_{1,n}) \\ (u_{1,n}, x) & (z_{1,n}, z_{2,n}) \end{array} \right). \end{array}$$

This will be referred to as game $\mathcal{G}_n(x)$ (the matrix will also be referred to as $\mathcal{G}_n(x)$). It should be noted that when at least one of the players accepts an object, the sum of the expected payoffs is always $x + u_{1,n}$.

The mixed Nash equilibrium for this subgame is as follows (see Neumann et al. [9]):
i) (s,s) when $x \geq u_{1,n}$.
ii) Player i plays s with probability $p_{i,n}(x)$ when $\max\{z_{1,n}, z_{2,n}\} \leq x \leq u_{1,n}$, where

$$p_{1,n}(x) = \frac{x - z_{2,n}}{\alpha_n x + (1-\alpha_n)u_{1,n} - z_{2,n}}; \quad p_{2,n}(x) = \frac{x - z_{1,n}}{(1-\alpha_n)x + \alpha_n u_{1,n} - z_{1,n}}.$$

iii) (c,s) when $z_{2,n} < x < z_{1,n}$.
iv) (s,c) when $z_{1,n} < x < z_{2,n}$.
v) (c,c) when $x \leq \min\{z_{1,n}, z_{2,n}\}$.
It can be shown by induction that $u_{2,n} - u_{1,n} \leq z_{i,n} \leq u_{1,n}$.

It should be noted that there are a multitude of non-cooperative solutions to such a game. The concept of a mixed equilibrium is used since the mixed equilibrium in symmetric games i.e., $\alpha_i = 0.5$, $i = 1, 2, \ldots, N$) is also symmetric. This is desirable, since by adopting the Shapley value we assume that the players have the same bargaining power.

Theorem 2.1. Let $\mathcal{H}$ denote a N-stage stopping game of the form given in which Player 1 always has priority. For any given realisation of the Markov process $\{X_i\}_{n=1}^N$, the objects taken at a Nash equilibrium of the game $\mathcal{H}$ are the same as the objects taken by an optimally behaving single player in the auxiliary problem.

Proof. Denote any non-cooperative value of this game when both players are searching and n objects are yet to appear by $(z_{1,n}^{\mathcal{H}}, z_{2,n}^{\mathcal{H}})$. Since Player 2 cannot prevent Player 1 from obtaining any object he/she wishes, it follows that $z_{1,n}^{\mathcal{H}} = u_{1,n}$. The payoff matrix of the subgame played when neither player has yet accepted an object, n objects are yet to appear and the value of the present object is x is given by:

$$\begin{array}{c c} & \begin{array}{cc} \text{s} & \text{c} \end{array} \\ \begin{array}{c} s \\ c \end{array} & \begin{pmatrix} (x, u_{1,n}) & (x, u_{1,n}) \\ (u_{1,n}, x) & (u_{1,n}, z^{\mathcal{H}}_{2,n}) \end{pmatrix} \end{array}.$$

This subgame will be referred to as $\mathcal{H}_n(x)$. It is assumed that if a player is indifferent between accepting an object and rejecting it, then he/she accepts it. Given this condition, (s,s) is the unique Nash equilibrium in this game when $x \geq u_{1,n}$. The unique Nash equilibrium is (c,s) when $z^{\mathcal{H}}_{2,n} \leq x < u_{1,n}$. The unique Nash equilibrium is (c,c) when $x < z^{\mathcal{H}}_{2,n}$.

If Player 1 has already accepted an object, then Player 2 accepts the next object of value $\geq u_{1,n}$. In order to show that the same objects are taken at a Nash equilibrium of this game as in the auxiliary problem, it suffices to show that $z^{\mathcal{H}}_{2,n} = u_{2,n} - u_{1,n}$. It follows from this that both in the auxiliary problem and at a Nash equilibrium of the game $\mathcal{H}$, the first object taken is the first object of value $\geq z^{\mathcal{H}}_{2,n}$ and the second object taken is the next object of value $\geq u_{1,n}$. In the game the first object chosen is taken by Player 1, iff its value is $\geq u_{1,n}$, otherwise the first object chosen is taken by Player 2.

When $N = 1$, it is obvious that in both cases the only object is taken. Considering the penultimate moment of the game, if the player in the auxiliary problem has not yet made a choice, then the penultimate object should always be taken, i.e., $u_{2,1} - u_{1,1} = 0$. In the game $\mathcal{H}$ Player 1 accepts this object iff $x_{N-1} \geq E[X_N|\mathcal{F}_{N-1}]$. Otherwise, Player 2 should take this object, since Player 1 will take the final object. It follows that $z^{\mathcal{H}}_{2,1} = 0$.

In general, considering the actions of the players in the 2-player game

$$z^{\mathcal{H}}_{2,n+1} = E[u_{1,n}\mathbb{I}_{X_{N-n}>u_{1,n}} + X_{N-n}\mathbb{I}_{z^{\mathcal{H}}_{n,2}<X_{N-n}<u_{1,n}} + z^{\mathcal{H}}_{n,2}\mathbb{I}_{X_{N-n}<z^{\mathcal{H}}_{n,2}}], \quad (3)$$

where $u_{1,n+1} = E[X_{N-n}\mathbb{I}_{X_{N-n}>u_{1,n}} + u_{1,n}\mathbb{I}_{X_{N-n}<u_{1,n}}]$.

In the auxiliary problem,

$$u_{2,n+1} = E[(X_{N-n} + u_{1,n})\mathbb{I}_{X_{N-n}>u_{2,n}-u_{1,n}} + u_{2,n}\mathbb{I}_{X_{N-n}<u_{2,n}-u_{1,n}}].$$

It follows that:

$$\begin{aligned} u_{2,n+1} - u_{1,n+1} = E[&u_{1,n}\mathbb{I}_{X_{N-n}>u_{1,n}} + X_{N-n}\mathbb{I}_{u_{2,n}-u_{1,n}<X_{N-n}<u_{1,n}} \\ &+ (u_{2,n} - u_{1,n})\mathbb{I}_{X_{N-n}<u_{2,n}-u_{1,n}}]. \end{aligned} \quad (4)$$

From the form of Eqs. (3) and (4), if $z^{\mathcal{H}}_{2,n} = u_{2,n} - u_{1,n}$, then $z^{\mathcal{H}}_{2,n+1} = u_{2,n+1} - u_{1,n+1}$. Since $z^{\mathcal{H}}_{2,1} = u_{2,1} - u_{1,1}$, it follows by induction that $z^{\mathcal{H}}_{2,n} = u_{2,n} - u_{1,n}$ for $1 \leq n \leq N$. □

A cooperative strategy must satisfy the following condition: if $x \geq u_{2,n} - u_{1,n} = z^{\mathcal{H}}_{2,n}$, then at least one of the players must accept the object. When $x \geq u_{1,n}$, the action s dominates the action c for both of the players. Hence, we define

$$v_n(\{1\};x)=\alpha_{N-n}x+(1-\alpha_{N-n})u_{1,n};\ \ v_n(\{2\};x)=\alpha_{N-n}u_{1,n}+(1-\alpha_{N-n})x.$$

When $\max\{z_{1,n},z_{2,n}\} < x < u_{1,n}$, then at the mixed equilibrium the players randomise between their actions c and s. Using the Bishop-Cannings theorem (see [4]), the expected payoffs when they do not form a coalition are

$$\begin{aligned} v_n(\{1\};x) &= [\alpha_{N-n}x+(1-\alpha_{N-n})u_{1,n}]p_{2,n}(x)+x[1-p_{2,n}(x)] \\ &= u_{1,n}p_{2,n}(x)+z_{1,n}[1-p_{2,n}(x)] \\ v_n(\{2\};x) &= [\alpha_{N-n}u_{1,n}+(1-\alpha_{N-n})x]p_{1,n}(x)+x[1-p_{1,n}(x)] \\ &= u_{1,n}p_{1,n}(x)+z_{2,n}[1-p_{1,n}(x)]. \end{aligned}$$

In both cases, the first (second) expression is the expected payoff when an individual plays s (c). At the mixed equilibrium these expected payoffs must be equal.

When $z_{2,n} < x < z_{1,n}$, at the mixed equilibrium the players play (c,s). Hence, we define

$$v_n(\{1\};x)=u_{1,n};\qquad v_n(\{2\};x)=x.$$

Similarly, when $z_{1,n} < x < z_{2,n}$, we define

$$v_n(\{1\};x)=x;\qquad v_n(\{2\};x)=u_{1,n}.$$

Suppose $u_{2,n}-u_{1,n} < x < \min\{z_{1,n},z_{2,n}\}$. In the cooperative game, one of the players should accept the object. At the mixed equilibrium, (c,c) is played. We define

$$v_n(\{1\};x)=z_{1,n};\qquad v_n(\{2\};x)=z_{2,n}.$$

In the case $x < u_{2,n}-u_{1,n}$ there is no conflict of interest. The players should reject the object both in the cooperative and the non-cooperative versions of the game. Hence, we assume that in this case when the players adopt a cooperative solution their expected reward from future search is given by the vector $(w_{1,n},w_{2,n})$. This is equivalent to assuming that

$$v_n(\{1\};x)=w_{1,n};\qquad v_n(\{2\};x)=w_{2,n}.$$

The Shapley value of such a subgame is given by the vector $(V_{1,n}(x),V_{2,n}(x))$, where

$$\begin{aligned} V_{1,n}(x) &= \frac{v_n(\{1,2\};x)+v_n(\{1\};x)-v_n(\{2\};x)}{2} \\ V_{2,n}(x) &= \frac{v_n(\{1,2\};x)+v_n(\{2\};x)-v_n(\{1\};x)}{2}. \end{aligned}$$

Hence,

$$V_{1,n}(x)=\begin{cases}\alpha_{N-n}x+(1-\alpha_{N-n})u_{1,n}, & x\geq u_{1,n}\\ xc_n(x)+u_{1,n}k_n(x), & \max\{z_{1,n},z_{2,n}\}\leq x<u_{1,n}\\ u_{1,n}, & z_{2,n}<x<z_{1,n}\\ x, & z_{1,n}<x<z_{2,n}\\ \frac{x+u_{1,n}+z_{1,n}-z_{2,n}}{2}, & u_{2,n}-u_{1,n}\leq x<\min\{z_{1,n},z_{2,n}\}\\ w_{1,n}, & x<u_{2,n}-u_{1,n}\end{cases}\tag{5}$$

$$V_{2,n}(x)=\begin{cases}\alpha_{N-n}x+(1-\alpha_{N-n})u_{1,n}, & x\geq u_{1,n}\\ x[1-c_n(x)]+u_{1,n}[1-k_n(x)], & \max\{z_{1,n},z_{2,n}\}\leq x<u_{1,n}\\ x, & z_{2,n}<x<z_{1,n}\\ u_{1,n}, & z_{1,n}<x<z_{2,n}\\ \frac{x+u_{1,n}+z_{2,n}-z_{1,n}}{2}, & u_{2,n}-u_{1,n}\leq x<\min\{z_{1,n},z_{2,n}\}\\ w_{2,n}, & x<u_{2,n}-u_{1,n}\end{cases},\tag{6}$$

where

$$c_n(x) = \frac{1-(1-\alpha_{N-n})p_{2,n}(x)+\alpha_{N-n}p_{1,n}(x)}{2}$$
$$k_n(x) = \frac{1+(1-\alpha_{N-n})p_{2,n}(x)-\alpha_{N-n}p_{1,n}(x)}{2}.$$

The Shapley value of the $n+1$-step stopping game is given by the expected Shapley value of the cooperative version of the subgame defined above. Thus, $(w_{1,n+1},w_{2,n+1}) = E[V_{1,n}(x),V_{2,n}(x)]$, where expectations are taken with respect to $\mathcal{F}_{N-n-1}$. Since $V_{1,n}(x)+V_{2,n}(x)=v_n(\{1,2\};x)$, $\forall x$, it follows that $w_{1,n+1}+w_{2,n+1}=E[\max\{X_{N-n}+u_{1,n},w_{1,n}+w_{2,n}\}]$. It is simple to show by induction that $w_{1,n}+w_{2,n}=u_{2,n}$ i.e., the sum of the payoffs to the players is maximised).

Let $\{v_n^2(\mathcal{S};x)\}_{\mathcal{S}=\{1\},\{2\},\{1,2\}}$ denote the following formulation of the characteristic function of the game $\mathcal{G}_n(x)$. According to this formulation, $v_n^2(\{1\};x)$ and $v_n^2(\{2\};x)$ are given by the minimax payoffs of the players, except in the case $x<u_{2,n}-u_{1,n}$. In this case, since there is no conflict, we assume that both players continue and cooperate obtaining expected rewards of $(w_{1,n},w_{2,n})$. Thus, we have

$$v_n^2(\{1\};x) = \begin{cases}\alpha_{N-n}x+(1-\alpha_{N-n})u_{1,n}, & x\geq u_{1,n}\\ x, & z_{1,n}\leq x<u_{1,n}\\ z_{1,n}, & u_{2,n}-u_{1,n}\leq x<z_{1,n}\\ w_{1,n}, & x<u_{2,n}-u_{1,n}\end{cases}$$
$$v_n^2(\{2\};x) = \begin{cases}\alpha_{N-n}u_{1,n}+(1-\alpha_{N-n})x, & x\geq u_{1,n}\\ x, & z_{2,n}\leq x<u_{1,n}\\ z_{2,n}, & u_{2,n}-u_{1,n}\leq x<z_{2,n}\\ w_{2,n}, & x<u_{2,n}-u_{1,n}\end{cases}.$$

Since the coalition of both players maximises the sum of their payoffs, we have $v_n^2(\{1,2\};x) = v_n(\{1,2\};x)$. The Shapley value of the subgame $\mathcal{G}_n(x)$ and of the n-step stopping game, $(w_{1,n}^2, w_{2,n}^2)$ are calculated as before.

Let $\{v_n^3(\mathcal{S};x)\}_{\mathcal{S}=\{1\},\{2\},\{1,2\}}$ denote the following formulation of the characteristic function of the game $\mathcal{G}_n(x)$. According to this formulation, $v_n^3(\{1\};x)$ and $v_n^3(\{2\};x)$ are given by the minimax payoffs of the players under the assumption that players only carry out realistic threats. In this case, when $\max\{z_{1,n}, z_{2,n}\} \leq x < u_{1,n}$, both players threat to play c, which minimises the payoff of the other player, is realistic. For example, given Player 1 carries out such a threat, the optimal action of Player 2 is to play s, which leads to Player 1 obtaining his/her greatest possible reward. However, if $z_{2,n} < x < z_{1,n}$, then Player 2's threat to play c is unrealistic. If he/she carries out such a threat, Player 2 will play his dominant strategy c, which leads to Player 2's payoff being minimised. In this case, it is assumed that Player 1's realistic threat to play c will lead to Player 2 playing s. Thus,

$$v_n^3(\{1\};x) = \begin{cases} \alpha_{N-n}x + (1-\alpha_{N-n})u_{1,n}, & x \geq u_{1,n} \\ x, & \max\{z_{1,n}, z_{2,n}\} \leq x < u_{1,n} \\ u_{1,n}, & z_{2,n} < x < z_{1,n} \\ x, & z_{1,n} < x < z_{2,n} \\ \frac{x+u_{1,n}+z_{1,n}-z_{2,n}}{2}, & u_{2,n} - u_{1,n} \leq x < \min\{z_{1,n}, z_{2,n}\} \\ w_{1,n}, & x < u_{2,n} - u_{1,n} \end{cases}$$

$$v_n^3(\{2\};x) = \begin{cases} \alpha_{N-n}x + (1-\alpha_{N-n})u_{1,n}, & x \geq u_{1,n} \\ x, & \max\{z_{1,n}, z_{2,n}\} \leq x < u_{1,n} \\ x, & z_{2,n} < x < z_{1,n} \\ u_{1,n}, & z_{1,n} < x < z_{2,n} \\ \frac{x+u_{1,n}+z_{2,n}-z_{1,n}}{2}, & u_{2,n} - u_{1,n} \leq x < \min\{z_{1,n}, z_{2,n}\} \\ w_{2,n}, & x < u_{2,n} - u_{1,n} \end{cases}.$$

The Shapley value of the subgame $\mathcal{G}_n(x)$ and of the n-step stopping game, $(w_{1,n}^3, w_{2,n}^3)$ are calculated as before.

Theorem 2.2. The Nash value of the game $\mathcal{H}$ is uniquely defined and equal to the Shapley value defined in any of the three ways given above.

Proof. First note that $(u_{1,0}, u_{2,0} - u_{1,0}) = (w_{1,0}, w_{2,0}) = (0,0)$. Assume that $(w_{1,n}, w_{2,n}) = (u_{1,n}, u_{2,n} - u_{1,n})$. The characteristic form of the cooperative game derived from the subgame $\mathcal{H}_n(x)$ according to any of the three definitions given above is given by:

$$\begin{aligned} v_n(\{1,2\};x) &= \max\{x + u_{1,n}, u_{2,n}\} \\ v_n(\{1\};x) &= \max\{x, u_{1,n}\} \\ v_n(\{2\};x) &= \begin{cases} u_{1,n}, & x \geq u_{1,n} \\ x, & u_{2,n} - u_{1,n} \leq x < u_{1,n} \\ u_{2,n} - u_{1,n}, & x < u_{2,n} - u_{1,n}. \end{cases} \end{aligned}$$

It follows that

$$V_{1,n}(x) = \begin{cases} x, & x \geq u_{1,n} \\ u_{1,n}, & x < u_{1,n}. \end{cases}$$

Hence, $w_{1,n+1} = u_{1,n+1}$. Since $w_{1,n+1} + w_{2,n+1} = u_{2,n+1}$, it follows that $w_{2,n+1} = u_{2,n+1} - u_{1,n+1}$. The proof follows by induction. □

Now consider a game in which priority is assigned randomly. Suppose that for some n in this game $u_{2,n} - u_{1,n} \leq w_{i,n} \leq u_{1,n}$, for $i \in \{1,2\}$. It follows that $v_n^{\mathcal{H}}(\{2\};x) \leq v_n(\{i\};x) \leq v_n^{\mathcal{H}}(\{1\};x)$, $\forall x$, $i = 1,2$, where $\{v_n^{\mathcal{H}}(\mathcal{S};x)\}_{S=\{1\},\{2\},\{1,2\}}$ is the characteristic form of the cooperative game derived from the matrix subgame $\mathcal{H}_n(x)$. Hence, $u_{2,n+1} - u_{1,n+1} \leq w_{i,n+1} \leq u_{1,n+1}$. Since, $w_{1,0} = w_{2,0} = u_{1,0} = u_{2,0}$, it follows by induction that $u_{2,n} - u_{1,n} \leq w_{i,n} \leq u_{1,n}$, for $i \in \{1,2\}$ and $0 \leq n \leq N$. This is true for all three of the formulations considered.

3 The form of a cooperative strategy

When a game is defined in strategic form it is not enough to define the Shapley value of the game, one should also derive a cooperative strategy that achieves this value. In this context, a cooperative strategy should be understood as an agreement which describes any necessary side payments and the actions to be taken in any given state. We wish to define cooperative strategies, which achieve the Shapley value and are subgame consistent.

It should be noted that when $x < u_{2,n} - u_{1,n}$, the Shapley value of the subgame $\mathcal{G}_n(x)$ can only be achieved when both players reject an object. No side payments are required in this case. It follows that to achieve the Shapley value, at most one side payment is required and this side payment can be made when the first object is accepted. Also, when $x \geq u_{1,n}$, the Shapley value is achieved when both players play s (since s dominates c in the case of both players, it is sensible to assume that both players play s). Again, no side payments are required. When $u_{2,n} - u_{1,n} \leq x < u_{1,n}$, in order for the Shapley value to be achieved one of the players must accept the object. It is assumed that the agreement assigns the object to one of the players. The side payments which may be used to achieve the Shapley value according to the first formulation of the problem are described below:

1) $\max\{z_{1,n}, z_{2,n}\} \leq x < u_{1,n}$. The object may be assigned to Player 1, who receives a side payment of $u_{1,n}k_n(x) - x[1 - c_n(x)]$ from Player 2. Otherwise, the object is assigned to Player 2, who receives a side payment of $u_{1,n}[1 - k_n(x)] - xc_n(x)$ from Player 1.

2) $z_{2,n} \leq x < z_{1,n}$. The object is assigned to Player 2 and no side payment is necessary.

3) $z_{1,n} \leq x < z_{2,n}$. The object is assigned to Player 1 and no side payment is necessary.

4) $u_{2,n} - u_{1,n} \le x < \min\{z_{1,n}, z_{2,n}\}$. Suppose the object is assigned to Player 2. He receives a side payoff of $\frac{u_{1,n}+z_{2,n}-x-z_{1,n}}{2}$ from Player 1. If the object is assigned to Player 1, he receives a side payoff of $\frac{u_{1,n}+z_{1,n}-x-z_{2,n}}{2}$ from Player 2.

Such cooperative strategies can be defined for the other two formulations of the Shapley value of the stopping game. For example, in both of these formulations when $\max\{z_{1,n}, z_{2,n}\} \le x < u_{1,n}$, we have $v_n(\{1,2\}; x) = x + u_{1,n}, v_n(\{1\}; x) = v_n(\{2\}; x) = x$. The object may be assigned to either player and the player receiving the object obtains a side payment of $0.5(u_{1,n} - x)$ from the other.

Remark 3.1. It is possible to define Shapley values for such stopping games in other ways. However, there are certain characteristics that such a solution should satisfy, namely:
1) The Shapley value of the game in which Player 1 always has priority must be $(u_{1,n}, u_{2,n} - u_{1,n})$, since both players can ensure themselves these payoffs and no gain can be made by cooperating.
2) $w_{1,n} + w_{2,n} = u_{2,n}$ and for symmetric games $w_{1,n} = w_{2,n}$, $n = 1, 2, \ldots, N$.
3) A cooperative solution should be subgame consistent. That is to say that: (a) given both players are still searching and n objects are still to appear, then any renegotiation would lead to the same strategy being followed; and (b) there can be no subgame in which the expected reward of either player at the cooperative solution is less than his/her expected reward at the mixed equilibrium solution.

The three solutions described above satisfy the first two conditions. It will now be shown that only the first is always subgame consistent.

Theorem 3.1. The first cooperative strategy described above is subgame consistent. Suppose $N \ge 2$ and $0 < \alpha_i < 1$. The second cooperative strategy is only subgame consistent for symmetric games i.e., $\alpha_i = 0.5$, $i = 1, 2, \ldots N$). The third strategy is only subgame consistent when $\alpha_i = 0.5$ for $i = 1, 2, \ldots N - 1$.

Proof. The fact that renegotiation at any stage leads to the same cooperative strategy being followed results directly from the recursive definition of the cooperative strategies. The fact that the first cooperative strategy is subgame consistent results from the definition of $v_n(\{1\}; x)$ and $v_n(\{2\}; x)$.

Now we prove that the second cooperative strategy is not subgame consistent in asymmetric games. Suppose $z_{1,n} \ne z_{2,n}$ for some $n \ge 1$. It follows from this that the game is asymmetric. Without loss of generality we may assume that $z_{1,n} > z_{2,n}$. Let $z_{2,n} < x < z_{1,n} (< u_{1,n})$. The expected payoffs of the players at this cooperative solution are $\frac{x+u_{1,n}}{2} < u_{1,n}$. Since Player 1 obtains an expected reward of $u_{1,n}$ when the mixed equilibrium is played, this strategy is not subgame consistent.

It is simple to check that the conditions for subgame consistency are satisfied for the third cooperative strategy as long as $x \notin [\max\{z_{1,n}, z_{2,n}\}, u_{1,n}]$. Suppose

$\max\{z_{1,n}, z_{2,n}\} < x < u_{1,n}$. This interval has positive length for $n \geq 1$ (this is the interval where at the mixed equilibrium the players randomise, randomisation is never used at the final moment of the game if both players are still searching). Using either of the second or the third cooperative strategies, the expected payoff of both of the players is $\frac{x+u_{1,n}}{2}$. Considering Player 1's payoff, a necessary condition for consistency is

$$\frac{u_{1,n}+x}{2} \geq [\alpha_{N-n}x + (1-\alpha_{N-n})u_{1,n}]p_{2,n}(x) + x[1-p_{2,n}(x)].$$

This leads to

$$(u_{1,n}-x)[\alpha_{N-n}(u_{1,n}-x) + (2\alpha_{N-n}-1)(x-z_{1,n})] \geq 0.$$

Since $u_{1,n} - x \geq 0$, it follows that

$$\begin{aligned} \alpha_{N-n}(u_{1,n}-x) + (2\alpha_{N-n}-1)(x-z_{1,n}) &\geq 0 \\ \alpha_{N-n} &\geq \frac{x-z_{1,n}}{u_{1,n}+x-2z_{1,n}}. \end{aligned}$$

This inequality has to be satisfied for all $x \in (\max\{z_{1,n}, z_{2,n}\}, u_{1,n})$. Simple differentiation of the right-hand side shows that this expression is maximised for $x \in [\max\{z_{1,n}, z_{2,n}\}, u_{1,n}]$, when $x = u_{1,n}$, and thus we have $\alpha_{N-n} \geq 0.5$.

Considering Player 2's payoff, the other necessary condition for consistency is

$$\frac{u_{1,n}+x}{2} \geq [\alpha_{N-n}u_{1,n} + (1-\alpha_{N-n})x]p_{1,n}(x) + x[1-p_{1,n}(x)].$$

This leads to:

$$(u_{1,n}-x)[(1-\alpha_{N-n})u_{1,n} - \alpha_{N-n}x - (1-2\alpha_{N-n}z_{2,n})] \geq 0.$$

Since $u_{1,n} - x \geq 0$, it follows that

$$\begin{aligned} (1-\alpha_{N-n})u_{1,n} - \alpha_{N-n}x - (1-2\alpha_{N-n}z_{2,n}) &\geq 0 \\ \alpha_{N-n} &\leq \frac{u_{1,n}-z_{2,n}}{u_{1,n}+x-2z_{2,n}}. \end{aligned}$$

This inequality has to be satisfied for all $x \in (\max\{z_{1,n}, z_{2,n}\}, u_{1,n})$. Simple differentiation of the right-hand side shows that this expression is minimised for $x \in [\max\{z_{1,n}, z_{2,n}\}, u_{1,n}]$, when $x = u_{1,n}$, and thus we have $\alpha_{N-n} \leq 0.5$. It follows that for subgame consistency of the third cooperative strategy $\alpha_i = 0.5$, $i = 1, 2, \ldots N-1$. Since at any solution of such games both players accept the final object if they are still searching, it follows that the second cooperative strategy is subgame consistent for symmetric games. □

It may be possible to derive another formulation of the characteristic functions of the subgames such that the solution is subgame consistent in the sense given here. However, it seems that the first formulation given here would be more natural.

4 Example

Suppose the values of the objects are i.i.d. random variables from the uniform distribution on $[0, 1]$. The payoff to a player is the value of the object obtained (0 if no object is obtained) plus the value of the side payment obtained. It is assumed that when both players wish to accept an object, then Player 1 obtains it with probability α. Without loss of generality it may be assumed that $\alpha \geq \frac{1}{2}$. Let $(w_{1,n}, w_{2,n})$ be the value of the restricted game when both players are still searching, n objects have yet to appear, and the time consistent cooperative strategy is played.

The optimal strategy of a lone searcher at moment n is to accept an object iff the value of the object is greater than the threshold given by the expected reward from future search, $u_{1,n}$. Hence,

$$u_{1,n+1} = E[\max\{X_{N-n}, u_{1,n}\}] = \frac{1 + u_{1,n}^2}{2}, \tag{7}$$

where $u_{1,0} = 0$.

Suppose $\alpha = 0.5$. From the symmetry of the game and the identity $w_{1,n}+w_{2,n} = u_{2,n}$, it follows that $w_{1,n} = w_{2,n} = \frac{u_{2,n}}{2} = w_n$. The same expected payoffs are obtained at the other two cooperative solutions described. Equation (1) leads to:

$$u_{2,n+1}=\int_0^{u_{2,n}-u_{1,n}} u_{2,n}dx+\int_{u_{2,n}-u_{1,n}}^1 (u_{1,n}+x)dx=u_{1,n}+\frac{1}{2}(1+[u_{2,n}-u_{1,n}]^2).$$

Suppose $\alpha > 0.5$. In order to solve the cooperative game, we need first to derive the value of the game at the mixed equilibrium. It can be shown by induction that $z_{1,n} > z_{2,n}$. It follows that:

$$\begin{aligned}
z_{1,n+1}=&\int_0^{z_{2,n}} z_{1,n}dx + \int_{z_{2,n}}^{z_{1,n}} u_{1,n}dx + \int_{z_{1,n}}^{u_{1,n}} \{u_{1,n}p_{2,n}(x) + z_{1,n}[1-p_{2,n}(x)]\}\, dx \\
&+ \int_{u_{1,n}}^1 [\alpha x + (1-\alpha)u_{1,n}]dx. \\
z_{2,n+1}=&\int_0^{z_{2,n}} z_{2,n}dx + \int_{z_{2,n}}^{z_{1,n}} xdx + \int_{z_{1,n}}^{u_{1,n}} \{u_{1,n}p_{1,n}(x) + z_{2,n}[1-p_{1,n}(x)]\}\, dx \\
&+ \int_{u_{1,n}}^1 [\alpha u_{1,n} + (1-\alpha)x]dx.
\end{aligned}$$

Hence,

$$z_{1,n+1} = z_{1,n}z_{2,n} + u_{1,n}(1-\alpha+z_{1,n}-z_{2,n}) - u_{1,n}^2(1-\alpha/2) + \alpha/2 + \frac{(u_{1,n}-\alpha z_{1,n})(u_{1,n}-z_{1,n})}{1-\alpha} + \frac{(u_{1,n}-z_{1,n})^2\alpha\ln\alpha}{(1-\alpha)^2}$$

$$z_{2,n+1} = \frac{z_{1,n}^2 + z_{2,n}^2 + 1 - \alpha - (1+\alpha)u_{1,n}^2}{2} + \alpha u_{1,n} + \frac{[u_{1,n}-(1-\alpha)z_{2,n}](u_{1,n}-z_{1,n})}{\alpha} - \frac{(1-\alpha)(u_{1,n}-z_{2,n})}{\alpha^2} \ln\left(\frac{u_{1,n}-z_{2,n}}{\alpha z_{1,n}+(1-\alpha)u_{1,n}-z_{2,n}}\right).$$

Now we can calculate the cooperative value of the game

$$w_{1,n+1} = \int_0^{u_{2,n}-u_{1,n}} w_{1,n}dx + \int_{u_{2,n}-u_{1,n}}^{z_{2,n}} \frac{x+u_{1,n}+z_{1,n}-z_{2,n}}{2}dx + \int_{z_{2,n}}^{z_{1,n}} u_{1,n}dx + \int_{z_{1,n}}^{u_{1,n}} [xc_n(x)+u_{1,n}k_n(x)]dx + \int_{u_{1,n}}^{1} [\alpha x+(1-\alpha)u_{1,n}]dx.$$

This leads to:

$$\begin{aligned} w_{1,n+1} = {} & w_{1,n}[u_{2,n}-u_{1,n}] + u_{1,n}(z_{1,n}-z_{2,n}) \\ & + \frac{z_{1,n}z_{2,n}+\alpha u_{1,n}^2 - z_{1,n}u_{2,n} + z_{2,n}u_{2,n} + \alpha}{2} \\ & + u_{1,n}(1-\alpha) - \frac{z_{2,n}^2+u_{2,n}^2+z_{1,n}^2}{4} \\ & - \frac{\alpha(z_{1,n}-u_{1,n})[(1-\alpha)(u_{1,n}-z_{1,n})-(z_{1,n}-u_{1,n})\ln\alpha]}{2(1-\alpha)^2} \\ & + \frac{(1-\alpha)(z_{2,n}-u_{1,n})}{2\alpha^2} \\ & \left\{\alpha(u_{1,n}-z_{1,n}) + [z_{2,n}-u_{1,n}]\ln\frac{u_{1,n}-z_{2,n}}{\alpha z_{1,n}+(1-\alpha)u_{1,n}-z_{2,n}}\right\}. \end{aligned}$$

The values of $w_{1,n}$ can be calculated by induction. The Shapley value for Player 2 is given by $w_{2,n} = u_{2,n} - w_{1,n}$.

Table 1 gives values of these cooperative solutions for $1 \leq N \leq 10$ for $\alpha = 0.5$ and $\alpha = 0.75$.

Table 1: Shapley values of games based on the job search problem. The values of the objects are uniformly distributed on [0,1]

		$\alpha = 0.5$	$\alpha = 0.75$			
N	$u_{2,n}$	w_n	$z_{1,n}$	$z_{2,n}$	$w_{1,n}$	$w_{2,n}$
1	0.5000	0.2500	0.3750	0.1250	0.3750	0.1250
2	1.0000	0.5000	0.5711	0.4195	0.5758	0.4242
3	1.1953	0.5977	0.6538	0.5363	0.6564	0.5389
4	1.3203	0.6602	0.7071	0.6096	0.7089	0.6114
5	1.4091	0.7046	0.7451	0.6612	0.7465	0.6626
6	1.4761	0.7380	0.7738	0.7000	0.7750	0.7011
7	1.5287	0.7643	0.7964	0.7303	0.7974	0.7313
8	1.5712	0.7856	0.8147	0.7547	0.8156	0.7556
9	1.6064	0.8032	0.8298	0.7750	0.8306	0.7757
10	1.6360	0.8180	0.8426	0.7920	0.8433	0.7927

References

[1] R. Amer and F. Carreras, *Cooperation indices and weighted Shapley values,* Math. Oper. Res. **22**, (1997), 955–968.

[2] R.J. Aumann, *Subjectivity and correlation in randomized strategies,* J. Math. Econ. **1**, (1974), 67–96.

[3] J. F. Banzhaf, *Weighted voting doesn't work: a mathematical analysis,* Rutgers Law Rev. **19**, (1965), 317–343.

[4] D. T. Bishop and C. Cannings, *A generalized war of attrition,* J. Theor. Biol. **70**, (1978), 85–124.

[5] E. Dynkin, *Game variant of a problem on optimal stopping,* Soviet Math. Dokl. **10**, (1969), 270–274.

[6] E. Ferenstein, *A variation of the Dynkin's stopping game,* Math. Japonica **38**, (1993), 371–379.

[7] J. Filar and L. Petrosjan, *Dynamic Cooperative Games,* Int. Game Theory Rev. **2**, (2000), 47–65.

[8] M. Fushimi, *The secretary problem in a competitive situation,* J. Oper. Res. Soc. Jap. **24**, (1981), 350–358.

[9] P. Neumann, D. Ramsey and K. Szajowski, *Randomized stopping times in Dynkin games,* Z. Angew. Math. Mech. **82**, (2002), 811–819.

[10] A. S. Nowak and T. Radzik, *A solidarity value for n-person transferable utility games*, Int. J. Game Theory **23**, (1994), 261–282.

[11] Y. Ohtsubo, *A nonzero-sum extension of Dynkin's stopping problem,* Math. Oper. Res. **12**, (1987), 277–296.

[12] L. Petrosjan, *Agreeable solutions in differential games,* Int. J. Math. Game Theory Alg. **7**, (1997), 165–177.

[13] L. Petrosyan, A. Sedakov and A. Syurin, *Multistage games with coalitional structure*, To be presented at ISDG 2006.

[14] L. Petrosjan and G. Zaccour, *Time-consistent Shapley value allocation of pollution cost reduction,* J. Econ. Dynam. Control **27**, (2003), 381–398.

[15] D. Ramsey and K. Szajowski, *Correlated Equilibria in Competitive Staff Selection Problem,* in Game Theory and Mathematical Economics, Banach Center Publications, Polish Academy of Sciences, Warsaw, Poland, 2006, pp. 253–265.

[16] D. Ramsey and K. Szajowski , *Correlated equilibrium in Markov stopping games*, Submitted to Eur. J. Oper. Res. (2006)

[17] L. S. Shapley, *A value for n-person games,* Contributions to the theory of games II. Princeton University Press, Princeton, NJ, 1953, pp. 307–317.

[18] K. Szajowski, *Markov stopping game with random priority,* ZOR - Math. Meth. Oper. Res. **39**, (1994), 69–84.

[19] M. Yasuda, *On a randomized strategy in Neveu's stopping problem,* Stochastic Proc. Appl. **21**, (1985) 159–166.

[20] D. W. K. Yeung and L. A. Petrosyan, *Cooperative Stochastic Differential Games,* Springer Series in Operations Research and Financial Engineering, Springer, Berlin, 2005.

[21] D. W. K. Yeung and L. A. Petrosyan, *Subgame Consistent Cooperative Solutions in Stochastic Differential Games,* J. Opt. Theory App. **120**, (2004), 651–666.

Pure Equilibrium Strategies for Stochastic Games via Potential Functions

Jos A.M. Potters
Radboud University Nijmegen
Nijmegen, The Netherlands
potters@math.ru.nl

T.E.S. Raghavan
University Illinois Chicago
Chicago, IL, USA
ter@uic.edu

Stef H. Tijs
University Tilburg
Tilburg, The Netherlands
S.H.Tijs@uvt.nl

Abstract

Strategic games with a potential function have quite often equilibria in pure strategies (Monderer and Shapley [4]). This is also true for stochastic games but the existence of a potential function is mostly hard to prove. For some classes of stochastic games with an additional structure, an equilibrium can be found by solving one or a finite number of finite strategic games. We call these games *auxiliary games*. In this paper, we investigate if we can derive the existence of equilibria in pure stationary strategies from the fact that the auxiliary games allow for a potential function. We will do this for zero-sum, two-person discounted stochastic games and non-zero-sum discounted stochastic games with additive reward functions and additive transitions (Raghavan et al. [8]) or with separable rewards and state independent transitions (Parthasarathy et al. [5]).

Key words. Stochastic games, strategic game with a potential function.

JEL-classification. C72, C73

1 Introduction

A potential game is a strategic game that allows a (potential) function on the set of strategy profiles such that potential differences under a *unilateral deviation* are the same as the differences in payoff for the deviating player. Therefore, the potential

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_22,

function can be used by deviating players as a leading principle in the sense that a profitable deviation increases the value of the potential function and, conversely, if the potential value increases under unilateral deviation, the deviation is profitable for the deviating player. This has two important consequences: an improvement process (a series of profitable unilateral deviations by various players) cannot cycle and a maximum of the potential function (if it exists) is automatically a Nash equilibrium. (Monderer and Shapley [4]) For finite strategic games this means that a potential game has a *pure* Nash equilibrium and every (neatly defined) improvement process is finite.

In the theory of stochastic games one finds several subclasses that can be solved by considering and solving one or more—what we will call—*finite auxiliary strategic games*. The solution of the auxiliary game(s) leads to a solution of the stochastic game. Examples are the zero-sum games occurring in the Shapley equations for zero-sum two-person stochastic games (Shapley [9],[10]), non-zero-sum stochastic game with an ARAT-structure (Additive Rewards and Additive Transitions; Raghavan et al. [8]) or a SER-SIT-structure (Separable Rewards and State Independent Transitions; Parthasarathy et al. [5]).

In this paper, we propose to investigate these classes of stochastic games and to see whether the existence of *pure stationary equilibria* (*saddle points* in the zero-sum case) can be derived from the existence of a potential function in the auxiliary game(s). The idea of using a pure stationary strategy is less problematic than the use of mixed actions in some states of the stochastic game.

Section 2 introduces the fundamental concepts in the theory of potential games (subsection 2.1) and the theory of β-discounted two-person stochastic games (subsection 2.2). Some of the results of the paper can be extended to n-person or undiscounted stochastic games but the paper focusses on β-discounted two-person games.

Section 3 discusses the zero-sum case and shows that a *zero-sum*, β-discounted, two-person stochastic game has a saddle point if the auxiliary games occurring in the Shapley equation (Shapley [4], have a potential function. *Zero-sum ARAT stochastic games* have these property but they form only a subset as we will show by an example.

Section 4 deals with non-zero-sum stochastic games. We consider the class of two-person stochastic games in which one player has additive rewards and the other player controls the transitions. These games are proved to have pure stationary equilibria. Another class we will investigate is formed by SER-SIT games (Parthasarathy et al. [5]). Here we need an additional condition to find a potential function in the auxiliary game introduced by Parthasarathy et al. By an example we show that the (or at least an) additional condition is needed.

2 The models and the tools

Subsection 2.1 contains the basic definitions in the theory of strategic games with a potential function and subsection 2.2 recalls the basic facts about stochastic games (cf. the book of Filar and Vrieze [2] for a comprehensive study of the subject).

2.1 Strategic Games with a Potential Function

Potential games are introduced in Monderer and Shapley [4].

An n-person strategic game $\langle A_1, \ldots, A_n, u_1, \ldots, u_n \rangle$ with (finite) action space A_i and utility function $u_i \colon A = \prod_{i=1}^{n} A_i \to I\!R$ for each player $i \in N$ is called a *potential game* if there is a *potential function* $F \colon A \to I\!R$ such that, for all $i \in N$, for all strategy profiles $(a_i)_{i \in N} \in A$ and every alternative action $b_i \in A_i$, we have $u_i(b_i, a_{-i}) - u_i(a) = F(b_i, a_{-i}) - F(a)$. Here, $a_{-i} = (a_1, \ldots a_{i-1}, a_{i+1}, \ldots, a_n)$ and $(b_i, a_{-i}) = (a_1, \ldots a_{i-1}, b_i, a_{i+1}, \ldots, a_n)$

Interesting properties of such a potential game are the existence of pure Nash equilibria (each element of argmax (F) is a Nash equilibrium) and the so-called finite improvement property: starting from any strategy profile a Nash equilibrium is reached after a finite number of unilateral improvements by various players.

In the next proposition we give two characterizations for two-person zero-sum potential games. One characterization says that, for each 2×2-subgame, the sum of the payoffs in the diagonal and the anti-diagonal are the same. The other characterization says that the utility function for each player is the sum of a part only dependent on the player's own action, and a part only dependent on the action of his opponent.

Lemma 2.1. *Given a strategic zero-sum game* $\langle A_1, A_2, u_1, u_2 \rangle$ *(*$u_1 + u_2 = 0$*) the following assertions are equivalent:*

(i) $\langle A_1, A_2, u_1, u_2 \rangle$ *is a potential game*

(ii) (diagonal property) For all $a_1, b_1 \in A_1$ *and* $a_2, b_2 \in A_2$ *we have*

$$u_1(a_1, a_2) + u_1(b_1, b_2) = u_1(a_1, b_2) + u_1(b_1, a_2)$$

(iii) (separation property) There are functions $g_1 \colon A_1 \to I\!R$ *and* $g_2 \colon A_2 \to I\!R$ *such that* $u_1(a_1, a_2) = g_1(a_1) + g_2(a_2)$ *for all* $(a_1, a_2) \in A_1 \times A_2$.

Proof: (a) Monderer and Shapley ([4], Theorem 2.8) proved that $\langle A_1, A_2, u_1, u_2 \rangle$ is a potential game if and only if, for all $a_i, b_i \in A_i$, we have:

$$(u_1(a_1, a_2) - u_1(b_1, a_2)) + (u_2(b_1, a_2) - u_2(b_1, b_2)) +$$
$$(u_1(b_1, b_2) - u_1(a_1, b_2)) + (u_2(a_1, b_2) - u_2(a_1, a_2)) = 0.$$

Substituting $u_2 = -u_1$ gives the equality

$2\,u_1(a_1, a_2) - 2\,u_1(b_1, a_2) - 2\,u_1(a_1, b_2) + 2\,u_1(b_1, b_2)) = 0$. This is assertion (ii).

(b) We assume the diagonal property. Take two points of reference $a_1^\circ \in A_1$ and $a_2^\circ \in A_2$ and define the functions g_1 and g_2 by $g_1(a_1) := u_1(a_1, a_2^\circ)$ and $g_2(a_2) := u_1(a_1^\circ, a_2) - u_1(a_1^\circ, a_2^\circ)$. Then,

$$u_1(a_1, a_2) = (u_1(a_1, a_2) - u_1(a_1^\circ, a_2)) + (u_1(a_1^\circ, a_2) - u_1(a_1^\circ, a_2^\circ)) + u_1(a_1^\circ, a_2^\circ) =$$

$$(u_1(a_1, a_2) + u_1(a_1^\circ, a_2^\circ) - u_1(a_1^\circ, a_2)) + g_2(a_2) = g_1(a_1) + g_2(a_2)$$

because of property (ii).

(c) If u_1 is separable, i.e., $u_1(a_1, a_2) = g_1(a_1) + g_2(a_2)$ for certain functions $g_i : A_i \to I\!R$ $(i = 1, 2)$, the diagonal property is easy to prove. ◁

Remark: For non-zero-sum strategic games the diagonal property and the separability property are not needed for the existence of a potential function, as the bimatrix game with the following payoff matrices and potential function shows:

$$(X, Y) = \begin{bmatrix} (1,5) & (1,0) \\ (3,10) & (4,6) \end{bmatrix} \quad \text{and} \quad F = \begin{bmatrix} 5 & 0 \\ 7 & 3 \end{bmatrix}.$$

Neither of the two payoff matrices has the diagonal property or the separability property.

For bimatrix games the equivalence of the diagonal property and the separability property remains true and these properties are still sufficient for the existence of a potential function. Necessary is that the difference of the payoff matrices satisfies the diagonal property.

2.2 Stochastic games with pure Nash equilibria

A *two-person stochastic game* is determined by the following data:

- There is a finite set S of *states*. Elements of S are denoted by $s, t, \ldots$
- There are two players, I and II.
- In each state $s \in S$ each of the players has a finite set of *actions*: A_s^1 and A_s^2.
- There are two functions r^1 and r^2 (the *reward functions*) defined on the set $T := \{(s, a^1, a^2) : s \in S, a^1 \in A_s^1 \text{ and } a^2 \in A_s^2\}$.
- There is a map $p : T \to \Delta(S)$ (the set of probability vectors on S). We write $p(t \mid s, a^1, a^2)$ for the t-th coordinate of $p(s, a^1, a^2)$. The map p gives the *Markov transition probabilities.*
- The game is played in a finite or countable sequence of *stages*. In each stage the players know what happened in the preceding stages and, in particular, the present state s and choose independently actions a^1 and a^2 in their own action spaces A_s^1 and A_s^2. Their action choice can be based on all information they have at the moment they make their decision: the history of the game. An history has the following form:

$$(s_0, a_0^1, a_0^2, s_1, a_1^1, a_1^2, \ldots, a_{n-1}^1, a_{n-1}^2, s_n).$$

The action choice results in an immediate reward to each of the players, $r^k(s, a^1, a^2)$ is the reward to player k and the next stage the process is in state t

with probability $p(t \mid s, a^1, a^2)$. The history (past states and past actions) of the game is common knowledge between the players.

By playing the game each of the players obtains a payoff every period and a stream of payoffs $\{r^k(s_n, a_n^1, a_n^2)\}_{n \in \mathbb{N}_0}$ results.

We assume that both players appreciate such a flow of rewards according to the utility function $U^k(\{r^k(s_n, a_n^1, a_n^2)\}_{n \in \mathbb{N}_0}) :\, = \sum_{n \in \mathbb{N}_0} \beta^n \, r^k(s_n, a_n^1, a_n^2)$. Here, $\beta \in (0,1)$ is a fixed *discount factor*. If there are only finitely many stages (a stochastic game with finite horizon) the summation is taken over the finitely many stages.

A stochastic game is called *zero-sum* if $r^1 + r^2 = 0$ on T.

A mixed *Markov component for player* k is a map $f^k \colon S \to \bigcup_{s \in S} \Delta(A_s^k)$ with the property $f^k(s) \in \Delta(A_s^k)$. So, a Markov component of player k determines a mixed action in A_s^k for each state $s \in S$ and the action choice is only dependent on the present state, the last entry of a history. A pure Markov component determines for every state s one admissible action. MC_k denotes the set of pure Markov components of player k. A *Markov strategy* consists of a sequence of Markov components $\{f_n^k\}_{n \in \mathbb{N}_0}$, one for each stage n. If f_n^k is the same for every stage, the strategy is called a *stationary (Markov) strategy*. If $f_n^k(s)$ is a pure action for every stage n and every state s, the (Markov or stationary) strategy is called a *pure strategy*. We denote Markov strategies for player k by $\bar{f}^k$ and the stationary strategy with $f_n^k = f^k$ for all n by $(f^k)^\infty$. If Markov strategies (or more general history dependent strategies) $\bar{f}^1$ and $\bar{f}^2$ are chosen, a stochastic process determines the probabilities to be in stage s_n at time n (as a function of the initial state s_0) and the stream of expected rewards $\{r^k(s_n, f_n^1(s_n), f_n^2(s_n))\}_{n \in \mathbb{N}_0}$ for both players. We define the payoff functions by:

$$V_\beta^k(\bar{f}^1, \bar{f}^2)_{s_0} :\, = \sum_{n \in \mathbb{N}_0} \beta^n \, r^k(s_n, f_n^1(s_n), f_n^2(s_n)) \;\; (k = 1, 2).$$

for all initial states $s_0 \in S$.

Then we have a strategic game (dependent on the set of strategies we allow). Nash equilibria of this game are called the optimal (or equilibrium) solutions of the β-discounted stochastic game.

For β-discounted two-person zero-sum stochastic games we have the following fundamental result of Shapley [9],[10].

(i) Both players have a—mostly mixed—stationary optimal strategy. So, a β-discounted two-person zero-sum stochastic games can be solved in mixed stationary strategies.

(ii) To find optimal stationary strategies $(f^1)^\infty$ and $(f^2)^\infty$ and the value function $v_\beta : S \to \mathbb{R}$ defined by $v_\beta(s) :\, = V_\beta^1((f^1)^\infty, (f^2)^\infty)$, a collection of *auxiliary* zero-sum games is considered.

For every state s and every vector $x \in \mathbb{R}^S$ we define the finite zero-sum game $\Gamma(s, x)$ with action space A_s^1 and A_s^2 and payoff matrix

$$\left[r^1(s, a^1, a^2) + \beta \sum_{t \in S} p(t \mid s, a^1, a^2) \, x_t\right]_{a^1 \in A_s^1, a^2 \in A_s^2}.$$

If a vector $x \in \mathbb{R}^S$ satisfies the *Shapley equations*

$x_s =$ value $(\Gamma(s,x))$ for every $s \in S$,

then $x = v_\beta$ and optimal strategy profiles $(f^1(s), f^2(s))$ in the auxiliary games $\Gamma(s, x = v_\beta)$ determine a stationary optimal solution $((f^1)^\infty, (f^2)^\infty)$ of the stochastic game.

Also, non-zero-sum, β-discounted, two-person stochastic games are solvable in stationary strategies. This result was proved by Fink [3] and Takahashi [11]. Both authors used a fixed point argument.

If in a stochastic game the Markov strategies $\bar{f}^1 = \{f^1_n\}_{n \in \mathbb{N}_0}$ and $\bar{f}^2 = \{f^2_n\}_{n \in \mathbb{N}_0}$ are applied and $x_0 \in \Delta(S)$ gives the initial probability distribution over the states, the probability distribution x_n over the states at time n is given by the recursive formula

$$x_n(t) = \textstyle\sum_{s \in S} x_{n-1}(s)\, p(t \mid s, f^1_{n-1}(s), f^2_{n-1}(s)) \text{ for all } t \in S.$$

If we introduce the $S \times S$-stochastic matrix

$P(f^1, f^2)$ by $P(f^1, f^2)_{s,t} : = p(t \mid s, f^1(s), f^2(s))$ for every pair of Markov components (f^1, f^2), the latter equation can be written as:

$$x_n = x_{n-1}\, P(f^1_{n-1}, f^2_{n-1}) \text{ for } n = 1, 2, \ldots.$$

The expected payoff in stage n, when $(\bar{f}^1, \bar{f}^2)$ is applied and the original distribution over the states is x_0, equals

$$\begin{aligned}&\textstyle\sum_{s \in S} x_n(s)\, r(s, f^1_n(s), f^2_n(s)) = \\ &\quad = \textstyle\sum_{s \in S} \left(x_0\, P(f^1_0, f^2_0) \cdots P(f^1_{n-1}, f^2_{n-1})\right)_s r(s, f^1_n(s), f^2_n(s)).\end{aligned}$$

If we define the S-vector $R^k(f^1, f^2) \in \mathbb{R}^S$ by $R^k(f^1, f^2)_s : = r^k(s, f^1(s), f^2(s))$ for every pair of Markov components (f^1, f^2), we find for the expected payoff in period n

$$\begin{aligned}&\textstyle\sum_{s \in S} x_0(s) \left(P(f^1_0, f^2_0) \cdots P(f^1_{n-1}, f^2_{n-1})\, R^k(f^1_n, f^2_n)\right)_s = \\ &\qquad \langle x_0, P(f^1_0, f^2_0) \cdots P(f^1_{n-1}, f^2_{n-1})\, R^k(f^1_n, f^2_n)\rangle.\end{aligned}$$

For stationary strategies the β-discounted payoff vectors $V^k_\beta((f^1)^\infty, (f^2)^\infty)$ $(k = 1, 2)$ equals

$$\begin{aligned}&V^k_\beta((f^1)^\infty, (f^2)^\infty) = R^k(f^1, f^2) + \beta\, P(f^1, f^2)\, R^k(f^1, f^2) + \cdots + \\ &+ \beta^{n-1} P(f^1, f^2)^{n-1} R^k(f^1, f^2) + \cdots = \ [I - \beta\, P(f^1, f^2)]^{-1} R^k(f^1, f^2).\end{aligned}$$

Notice that the matrix $[I - \beta\, P(f^1, f^2)]^{-1}$ is a nonnegative matrix for every pair of Markov components (f^1, f^2).

The main topic of this paper is to find conditions under which two-person β-discounted stochastic games have pure stationary equilibria. In the literature there are several special classes of stochastic games that can be solved by using ‘auxiliary games’ of various nature. We will investigate for which classes the existence of pure stationary strategies can be derived from the fact that the auxiliary game has a potential function.

We will repeat the definition of the classes we will consider.

ARAT-games (Additive Rewards and Additive Transitions)

A two-person stochastic game is an *ARAT-game* if, for all $s \in S$, all $a^1 \in A^1_s$ and all $a^2 \in A^2_s$,

$$r^1(s, a^1, a^2) = r^{11}(s, a^1) + r^{12}(s, a^2),$$
$$r^2(s, a^1, a^2) = r^{21}(s, a^1) + r^{22}(s, a^2),$$
$$p(t \mid s, a^1, a^2) = p^1(t \mid s, a^1) + p^2(t \mid s, a^2)$$

for some functions r^{ij} and p^j. So, the reward functions as well as the transition probabilities are sums of two functions, one only dependent on the state and the action of player I, the other only dependent on the state and the action of player II. If p^1 or p^2 is identically zero, we talk about *single control games*. If, for every element $s \in S$ one of the transition functions $p^1(s, a^1, a^2)$ or $p^2(s, a^1, a^2)$ vanishes, we call the stochastic game a stochastic game with *switching control.*

SER-SIT-games (Separable Rewards and State Independent Transitions)

In a SER-SIT stochastic game the actions spaces are the same in each state:

$$A^k_s = A^k_t (=: A^k) \text{ for } s, t \in S \text{ and } k = 1, 2.$$

A stochastic game is a *SER-SIT-game* if, for all states $s \in S$, all $a^1 \in A^1$ and all $a^2 \in A^2$,

$$r^1(s, a^1, a^2) = r^{10}(s) + r^{11}(a^1, a^2)$$
$$r^2(s, a^1, a^2) = r^{20}(s) + r^{21}(a^1, a^2)$$
$$p(t \mid s, a^1, a^2) = p(t \mid a^1, a^2) \text{ is not dependent on } s.$$

So, the reward function is the sum of two parts, one only dependent on the present state and the other only dependent on the action profile. The transitions are not dependent on the present state.

If $p(t \mid s, a^1, a^2) = p(t \mid s)$ is not dependent on the actions, we call the stochastic game a stochastic game with *action independent transitions* (AIT).

3 Zero-sum stochastic games with pure optimal strategies

In this section we prove that two-person, zero-sum, β-discounted stochastic games have a *pure* optimal stationary strategy if each of the auxiliary games $\Gamma(s, v_\beta)$ $(s \in S)$ has a potential function. If all auxiliary games $\Gamma(s, x)$ $(s \in S,\ x \in \mathbb{R}^S)$ have a potential function, then the stochastic game is an ARAT-game. We provide an example showing that not every stochastic game of the first category is an ARAT-game.

Theorem 3.1. If the auxiliary games $\{\Gamma(s, v_\beta)\}_{s \in S}$ of a two-person, zero-sum, β-discounted stochastic game have a potential function, then the stochastic game has an optimal stationary strategy in pure strategies.

Proof: By Monderer and Shapley [4], every auxiliary game has a pure equilibrium and by Shapley [9] these pure strategies form an optimal stationary strategy in the stochastic game. ◁

In the second lemma we prove that Theorem 3.1 can be applied to zero-sum ARAT-games.

Lemma 3.2 *If a two-person zero-sum game is an ARAT-game, all auxiliary games* $\Gamma(s,x)$ *have a potential function and, conversely, if all auxiliary games* $\Gamma(s,x)$ *have a potential function, the stochastic game is an ARAT-game.*

Proof: The payoff in any game $\Gamma(s,x)$ can be written as:

$$r^1(s,a^1,a^2)+\beta\sum_{t\in S}p(t\,|\,s,a^1,a^2)\,x_t =$$
$$[r^{11}(s,a^1)+\beta\sum_{t\in S}p^1(t\,|\,s,a^1)\,x_t]+[r^{12}(s,a^2)+\beta\sum_{t\in S}p^2(t\,|\,s,a^2)\,x_t].$$

So, the zero-sum game $\Gamma(s,x)$ satisfies the separability property and, therefore, it admits a potential function by Theorem 3.1.

Conversely, if $\Gamma(s,x)$ has a potential function for all $s\in S$ and all $x\in \mathbb{R}^S$, one can take $x=0$ and find by Proposition 1, that the rewards are additive: $r^1(s,a^1,a^2)=r^{11}(s,a^1)+r^{12}(s,a^2)$.

If we take any $s,t\in S$ and $x=e_t$, the separation condition for the auxiliary game $\Gamma(s,e_t)$ gives $r^{11}(s,a^1)+r^{12}(s,a^2)+\beta\,p(t\,|\,s,a^1,a^2)=\bar{r}^{11}(s,a^1)+\bar{r}^{12}(s,a^2)$ for certain functions $\bar{r}^{11}$ and $\bar{r}^{12}$.

From this equation we can find the components

$$p^k(t\,|\,s,a^k)=\frac{\bar{r}^{1k}(s,a^k)-r^{1k}(s,a^k)}{\beta}.$$ ◁

Corollary. *Two-person zero-sum ARAT-games have a pure optimal stationary strategy.*

The next example shows that the ARAT-games do not exhaust the class of stochastic games where Lemma 2.1 applies.

Example 1. Let $S=\{s_0,s_1,s_2,s_3\}$. The states $s_i\neq s_0$ are absorbing states: both players have one action and the reward is $r(s_i)=u_i$ for $i=1,2,3$. In state s_0 both players have two actions and the rewards and transitions are given by:

$$r^1(s_0)=\begin{bmatrix}-u_1 & -u_2\\ -u_3 & 0\end{bmatrix}\quad\text{and}\quad p(s_0)=\begin{bmatrix}\to s_1 & \to s_2\\ \to s_3 & \to s_0\end{bmatrix}.$$

Then for the state s_i $(i\geq 1)$ we have $v_\beta(s_i)=\dfrac{u_i}{1-\beta}$ and for state s_0 the auxiliary game is

$$\Gamma(s_0,v_\beta):\quad\begin{bmatrix}u_1(-1+\frac{\beta}{1-\beta}) & u_2(-1+\frac{\beta}{1-\beta})\\ u_3(-1+\frac{\beta}{1-\beta}) & \beta\,v_\beta(s_0)\end{bmatrix}.$$

Then $v_\beta(s_0)=0$ solves the Shapley equations if

$$u_3(-1+\frac{\beta}{1-\beta})>0 \text{ and } u_2(-1+\frac{\beta}{1-\beta})<0.$$

If we take $u_1=u_2+u_3$ we have the diagonal property. If $\beta>0.5$ we must have $u_3 > 0$, $u_2 < 0$ and $u_1 = u_2 + u_3$. Clearly, the rewards are additive but the transitions are not.

4 Non-zero-sum stochastic games with equilibria in pure strategies

We start with two rather simple classes of non-zero-sum stochastic games with pure Nash equilibria. The first class is the class of *coordination games*. A n-person stochastic game is a *coordination game* if all players have the *same reward function*. The second class is the class of stochastic games with action independent transitions in which all n-person games $a = (a_1, \dots a_n) \mapsto (r^1(s,a), \dots, r^n(s,a))$ have a potential function.

Theorem 4.1. (i) If in an n-person stochastic game the reward functions are the same for all players, then the stochastic game has a Nash equilibrium in pure strategies.

(ii) If an n-person stochastic game has action independent transitions (AIT) and for each state $s \in S$ the finite n-person game with strategy sets $\{A_s^k\}_{k=1,\dots,n}$ and payoff functions $a \mapsto \{r^k(s,a)\}_{k=1,\dots,n}$ has a potential function, then the stochastic game has a Nash equilibrium in pure strategies.

Proof: (i) We prove that a Nash equilibrium is obtained by solving the following dynamical programming problem. The set of states $\bar{S}$ is the same as in the stochastic game. The actions in state $s \in S$ is $\bar{A}_s : = \prod_{i=1}^n A_s^i$. Finally, the reward functions and transition functions are also the same as in the stochastic game (although the interpretation is slightly different): $\bar{r} = r = r^i$ for all players i and $\bar{p} = p$. It is well known (see Blackwell [1]) that dynamic programming problems have optimal stationary strategies in pure actions. Let $\bar{f} : s \in S \mapsto \bar{f}(s) \in \bar{A}_s$ be such an optimal strategy. If each player i chooses the i-th component f_i of $\bar{f}$, this is a Nash equilibrium in the stochastic game. If any player deviates from his strategy in any stage of the game, his expected payoff cannot increase as the same deviation in the dynamical program would give the same increase of expected payoff.

(ii) If the players have no influence on the transitions, they will make the best out of the state they are. By the result of Shapley and Monderer there is a pure Nash equilibrium in each *auxiliary game*. That is, for each state $s \in S$ there is a Nash equilibrium $(f^1(s), \dots, f^n(s)) \in \prod_{i=1}^n A_s^i$ of the stage game. This defines a Markov component f^i for each player i. Deviation of any player will not increase the payoff at any time in any state and has no influence on later expected payoffs. No player can gain by deviating at any time in any state. $\triangleleft$

For non-zero-sum stochastic games the ARAT-structure is not sufficient for the existence of a pure stationary equilibrium (see Raghavan et al.[8] and Thuijsman and Raghavan [12] for an example). Moreover, the auxiliary games as considered in the case of zero-sum games make little sense. If, however, the rewards for one player are *additive* and the transitions are *single controlled* by the other player, there are pure stationary equilibria.

Theorem 4.2. If in a non-zero-sum, two-person, stochastic game player I has an additive reward function and the transitions are controlled by (the actions of) player II, then there is a pure stationary equilibrium.

Proof: The additivity of the reward function of player I means that

$$r^1(s, a^1, a^2) = r^{11}(s, a^1) + r^{12}(s, a^2).$$

The single control says that $p(t \mid s, a^1, a^2) = p^2(t \mid s, a^2)$.

As player I has no influence on the transitions it seems natural to assume that, in every state he will maximize the part of the reward $r^{11}(s, a^1)$ under his control.

So, we define a pure Markov component f_0^1 with the property that

$$r^{11}(s, f_0^1(s)) = \max_{a^1 \in A_s^1} r^{11}(s, a^1).$$

If we fix f_0^1 player II will consider the dynamic decision problem with reward function $\bar{r} : T^2 := \{(s, a^2) : s \in S, a^2 \in A_s^2\} \to I\!R$ defined by $\bar{r}(s, a^2) := r^2(s, f_0^1(s), a^2)$ and transition probabilities $\bar{p} = p^2 : T^2 \to \Delta(S)$ as before.

Dynamical programming problems can be solved in pure stationary strategies. One solves first the *Bellman equations*:

$$x \in I\!R^S : \quad x_s = \max_{a^2 \in A_s^2} [\bar{r}(s, a^2) + \beta \sum_{t \in S} \bar{p}(t \mid s, a^2)\, x_t]$$

and takes $f_0^2(s) \in A_s^2$ such that

$$f_0^2(s) \in \arg\max \{a^2 \mapsto [\bar{r}(s, a^2) + \beta \sum_{t \in S} \bar{p}(t \mid s, a^2)\, x_t]\}.$$

Then $(f_0^2)^\infty$ is a pure stationary strategy and a best response to $(f_0^1)^\infty$, because it is an optimal strategy in the dynamic program.

The pure stationary strategy $(f_0^1)^\infty$ is a best response to every stationary strategy of player II. This follows from the formula $V_\beta^1(f_0^1, f^2) = [I - \beta\, P(f^2)]^{-1}\, (R^{11}(f_0^1) + R^{12}(f^2))$. If player I deviates from f_0^1 to f^1 all components of $R^{11}(f^1)$ are less or equal to the components of $R^{11}(f_0^1)$ and as the matrix $[I - \beta\, P(f^2)]^{-1} \geq 0$ also

$$[I - \beta\, P(f^2)]^{-1}\, (R^{11}(f_0^1) - R^{11}(f^1) \geq 0.$$

The term $[I - \beta\, P(f^2)]^{-1}\, (R^{12}(f^2)$ does not change by a deviation of player I and $(f_0^1)^\infty$ is a best response to every stationary strategy of player II. ◁

Remark: In Nowak and Raghavan (1993) it is proved that, if the transitions are single control, a Nash equilibrium can be found by solving the following *auxiliary* bimatrix game. The strategy space of each player consists of the finite set of all Markov components $\{f_k^1\}_{k \in MC^1}$ and $\{f_l^2\}_{l \in MC^2}$.

If player I chooses the Markov component f_k^1 and player II chooses the Markov component f_l^2, the payoff to player I is $\sum_{s \in S} R^1(f_k^1, f_l^2)_s$ and the payoff to player II is $\sum_{s \in S} V_\beta^2(f_k^1, f_l^2)_s$. If $\{\xi_k\}_{k \in MC^1}$ and $\{\eta_l\}_{l \in MC^2}$ form a Nash equilibrium of the auxiliary game, the stationary strategies $f_*^1 := \sum_{k \in MC^1} \xi_k\, f_k^1$ and $f_*^2 := \sum_{l \in MC^2} \eta_l\, f_l^2$ defines an equilibrium in stationary strategies for the β-discounted stochastic game. Therefore, the stochastic game has a pure stationary strategy, if the auxiliary bimatrix game has a pure Nash equilibrium.

If we have a SER-SIT-stochastic game and the bimatrix game

$[r^k(s, a^1, a^2)]_{a^1 \in A^1, a^2 \in A^2}$

is a potential game for every state $s \in S$, the β-discounted stochastic game need not have a pure stationary equilibrium. We need an additional condition.

Theorem 4.3. If in a stochastic game with separable reward functions and state independent transitions

(i) the partial reward game $[r^{k1}(a^1, a^2)]_{a^1 \in A^1, a^2 \in A^2}$ has a potential function and

(ii) the matrix game $Q\colon (a^1, a^2) \to \langle p(a^1, a^2), r^{10} - r^{20} \rangle$ has the diagonal property,

then the stochastic game has an equilibrium in pure stationary strategies.

Proof: Parthasarathy et al. [5] provides a method to solve SER-SIT stochastic games. It is sufficient to solve the auxiliary bimatrix game with strategy spaces A^1 and A^2 and payoff matrices

$[r^{k1}(a^1, a^2) + \beta \sum_{t \in S} p(t \mid a^1, a^2)\, r^{k0}(t)]_{a^1 \in A^1, a^2 \in A^2}.$

If $(\xi, \eta) \in \Delta(A^1) \times \Delta(A^2)$ is a Nash equilibrium of the auxiliary game, then the (even state independent) stationary strategy $(\xi^\infty, \eta^\infty)$ is an equilibrium in the β-discounted SER-SIT stochastic game. If the auxiliary game has a potential function, the stochastic game has a pure stationary equilibrium.

Let $F_1\colon A^1 \times A^2 \to I\!R$ be a potential function of the reward game with payoffs

$[r^{k1}(a^1, a^2)]_{a^1 \in A^1, a^2 \in A^2}.$

We have to prove that the bimatrix game $[\beta \sum_{t \in S} p(t \mid a^1, a^2)\, r^{k0}(t)]$ has a potential function. By Monderer and Shapley [4], (Theorem 2.8) this is true if and only if the matrix

$Q = [\beta \sum_{t \in S} p(t \mid a^1, a^2)\, (r^{10}(t) - r^{20}(t))]_{a^1 \in A^1, a^2 \in A^2}$ has the diagonal property (see the proof of Proposition 2.1). ◁

Example 2. The following example shows that the condition

$Q\colon \quad (a^1, a^2) \to \langle p(a^1, a^2), r^{10} - r^{20} \rangle$ has the diagonal property,

is not superfluous. Let $S = \{s_1, s_2\}$. The action spaces A^1 and A^2 consist of two actions:

$$r^{k1}(a^1, a^2)\colon = \begin{bmatrix} (1,0) & (0,0) \\ (2,2) & (0,1) \end{bmatrix} \qquad r^{k0}\binom{s_1}{s_2}\colon = \begin{bmatrix} (3 & -1) \\ (-3 & 1) \end{bmatrix}$$

$$p(s_1 \mid a^1, a^2) : p(s_2 \mid a^1, a^2)\colon = \begin{bmatrix} (1:0) & (0:1) \\ (0.5:0.5) & (0.5:0.5) \end{bmatrix}.$$

A potential function F_1 for $[r^{k1}]$ and the function $Q\colon (a^1, a^2) \to \langle p(a^1, a^2), r^{10} - r^{20} \rangle$ have the following values

$$F_1(a^1, a^2) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \qquad Q(a^1, a^2) = \begin{bmatrix} 4 & -4 \\ 0 & 0 \end{bmatrix}$$ (not satisfying the diagonal property).

If we compute the auxiliary bimatrix game, we get the matrices:

$$\begin{bmatrix}(1,0) & (0,0)\\(2,2) & (0,1)\end{bmatrix} + \beta \begin{bmatrix}(3,-1) & (-3,1)\\(0,0) & (0,0)\end{bmatrix} = \begin{bmatrix}(1+3\beta,-\beta) & (-3\beta,\beta)\\(2,2) & (0,1)\end{bmatrix}.$$

If $\beta > \frac{1}{3}$, the auxiliary game has only one *completely mixed* Nash equilibrium.

References

[1] Blackwell D. Discounted dynamic programming. *Annals of Mathematical Statistics,* **33,** (1962), 719–726.

[2] Filar JA and Vrieze OJ. Competitive Markov decision processes. Springer Verlag, Berlin (1996).

[3] Fink AM. Equilibrium in stochastic n-person game. *J. Science of the Hiroshima University,* Series A–I, **28,** (1964), 89–93.

[4] Monderer D and Shapley LS. Potential games. *Games and Economic Behavior,* **14,** (1996), 124–143.

[5] Parthasarathy T, Tijs SH and Vrieze OJ. Stochastic games with state independent transitions and separable rewards. In: Selected topics in OR and Mathematical Economics (Eds: G. Hammer and D.Pallaschke) Springer Verlag, Berlin, Lecture Note Series 226, (1984), 262–271.

[6] Nowak AS and Raghavan TES. A finite step algorithm via a bimatrix game to a single controller non-zero sum stochastic game, *Mathematical Programming* **59**, (1993), 249–259.

[7] Thuijsman F and Raghavan TES. Perfect information stochastic games and related classes, *International J. Game Theory,* **26,** (1997), pp. 403–408.

[8] Raghavan TES, Tijs SH and Vrieze OJ. On stochastic games with additive reward and additive transition structure. *J. Optimization Theory and Applications*, **47,** (1985), 451–464.

[9] Shapley LS. Stochastic games. *Proceedings of the National Academy of Science USA*, **39,** (1953), 1095–1100.

[10] Shapley LS. Some topics in two-person games. In: Advances in game theory (Eds. M Dresher, LS Shapley and AW Tucker). Princeton University Press, Princeton NJ, (1964), 1–28.

[11] Takahashi M. Equilibrium points of stochastic non-cooperative n-person game. *J. Science of the Hiroshima University,* Series A–I, **28,** (1964), 95–99.

[12] Thuijsman F and Raghavan TES. Perfect information stochastic games and related classes, *International J. Game Theory,* **26,** (1997), pp. 403–408.

Fibonacci Numbers and Equilibria in Large “Neighborhood” Games

Andrzej Wieczorek
Institute of Computer Science
Polish Academy of Sciences
Warszawa, Poland
Andrzej.Wieczorek@ipipan.waw.pl

Abstract

We deal with a game-theoretic framework involving a finite number of infinite populations, members of which have a finite number of available strategies. The payoff of each individual depends on her own action and distributions of actions of individuals in all populations. A method to find all equilibria is discussed which requires the search through all nonempty subsets of the types' strategy sets, assigning equilibria to each of them. The method is then used to find equilibria in two types of “neighborhood” games in which there is one type of player who has strategies in $V = \{1, \ldots, k\}$ and payoff functions $\Phi(j; \mathbf{p}) = \alpha \cdot \mathbf{p}_{j-1} + \mathbf{p}_j + \alpha \cdot \mathbf{p}_{j+1}$ for $j = 2, \ldots, k-1$ and: in the case of “chain” games $\Phi(1; \mathbf{p}) = \mathbf{p}_1 + \alpha \cdot \mathbf{p}_2;\ \Phi(k; \mathbf{p}) = \alpha \cdot \mathbf{p}_{k-1} + \mathbf{p}_k$; in the case of “circular” games $\Phi(1; \mathbf{p}) = \alpha \cdot \mathbf{p}_k + \mathbf{p}_1 + \alpha \cdot \mathbf{p}_2;\ \Phi(k; \mathbf{p}) = \alpha \cdot \mathbf{p}_{k-1} + \mathbf{p}_k + \alpha \cdot \mathbf{p}_1$; (in both cases $0 \leq \alpha \leq \frac{1}{2}$; $\mathbf{p}$ is a distribution on V). The Fibonacci numbers are used to determine the coordinates of equilibria in the case $\alpha = \frac{1}{3}$, for other values of α we need to construct numerical Fibonacci-like sequences which determine, in an analogous manner, coordinates of equilibria. An alternative procedure makes use of some numerical Pascal-like triangles, specially constructed for this purpose.

Key words. Fibonacci numbers, numerical triangles, large game, “neighborhood” game, equilibrium

AMS Subject Classifications. 91A10, 91A13, 91D25, 11B39

1 Equilibria in simple large games

Formally, a *game* of the type considered in this paper is defined as a system consisting of positive integers $n, k_1, k_2, \ldots, k_n$ (n is the number of *populations* and, for $i = 1, \ldots, n$, k_i is the number of *strategies* available for individuals of *type* i) and functions $\Phi^i : V^i \times \Delta_{k_1} \times \cdots \times \Delta_{k_n} \to \mathbb{R}$, $i = 1, \ldots, n$ (we denote the set $\{1, 2, \ldots, k_i\}$ of strategies of type i by V^i; Δ_{k_i} denotes the standard simplex in $\mathbb{R}^{k_i}$); $\Phi^i\left(j; \mathbf{p}^1, \ldots, \mathbf{p}^n\right)$ is the *payoff* of a player of type i choosing the strategy j

P. Bernhard et al. (eds.), *Advances in Dynamic Games and Their Applications*,
Annals of the International Society of Dynamic Games 10, DOI: 10.1007/978-0-8176-4834-3_23,

while, for $i = 1, \ldots, n$, the prevailing distributions of strategies of the respective types are $\mathbf{p}^1, \ldots, \mathbf{p}^n$.

An *equilibrium* for a game $\Gamma = \left(n, k_1, k_2, \ldots, k_n; \Phi^1, \ldots, \Phi^n\right)$ is a system of distributions of strategies of the players of respective types $\left(\mathbf{p}^1, \ldots, \mathbf{p}^n\right) \in \Delta_{k_1} \times \cdots \times \Delta_{k_n}$ such that, for $i = 1, \ldots, n$,

$$\max\nolimits_{j=1,\ldots,k_i} \Phi^i\left(j; \mathbf{p}^1, \ldots, \mathbf{p}^n\right) = \sum\nolimits_{j=1}^{k_i} \mathbf{p}^i_j \cdot \Phi^i\left(j; \mathbf{p}^1, \ldots, \mathbf{p}^n\right).$$

So an equilibrium is a system of distributions of strategies such that for each type i, the average payoff of players of this type is equal to the maximal possible payoff; this, somewhat informally, means that there are no players of any type who could increase their payoffs.

The most natural technique to find all equilibria in a game Γ seems to reduce the problem to finding solutions of the system of equations:

$$\begin{cases} \max\nolimits_{j=1,\ldots,k_1} \Phi^1\left(j; \mathbf{p}^1, \ldots, \mathbf{p}^n\right) = \sum\nolimits_{j=1}^{k_1} \mathbf{p}^1_j \cdot \Phi^1\left(j; \mathbf{p}^1, \ldots, \mathbf{p}^n\right) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \max\nolimits_{j=1,\ldots,k_n} \Phi^n\left(j; \mathbf{p}^1, \ldots, \mathbf{p}^n\right) = \sum\nolimits_{j=1}^{k_n} \mathbf{p}^n_j \cdot \Phi^n\left(j; \mathbf{p}^1, \ldots, \mathbf{p}^n\right) \\ \mathbf{p}^1 \in \Delta_{k_1}, \ldots, \mathbf{p}^n \in \Delta_{k_n}. \end{cases}$$

However, in practice, finding all solutions of the above system, or even finding a single solution, is difficult because of the "$\max$" operator which always turns a smooth system into a nonsmooth one and, in general, it changes the class of the considered problems.

We therefore propose the following algorithm to find all equilibria in Γ.

1. Select, for each type $i = 1, \ldots, n$, a nonempty set of strategies $W^i \subseteq V^i$;

2. Find all solutions of the system of equations:

$$\begin{cases} \Phi^1\left(j; \tilde{\mathbf{p}}^1, \ldots, \tilde{\mathbf{p}}^n\right) = \Phi^1\left(j'; \tilde{\mathbf{p}}^1, \ldots, \tilde{\mathbf{p}}^n\right), \quad j, j' \in W^1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \Phi^n\left(j; \tilde{\mathbf{p}}^1, \ldots, \tilde{\mathbf{p}}^n\right) = \Phi^n\left(j'; \tilde{\mathbf{p}}^1, \ldots, \tilde{\mathbf{p}}^n\right), \quad j, j' \in W^n \\ \tilde{\mathbf{p}}^1 \in \Delta_{k_1}, \ldots, \tilde{\mathbf{p}}^n \in \Delta_{k_n}, \end{cases}$$

where, for $i = 1, \ldots, n$, $\tilde{\mathbf{p}}^i_j = \mathbf{p}^i_j$ whenever $j \in W^i$ and otherwise $\tilde{\mathbf{p}}^i_j$ is fixed as $\mathbf{0}$ (note that the system consists of $\#W^1 + \cdots + \#W^n - n$ equations which contain $\#W^1 + \cdots + \#W^n$ variables; if for technical reasons the number of equations should be increased, the natural candidates are the equations $\sum_{j=1}^{k_i} \tilde{\mathbf{p}}^i_j = 0$, $i = 1, \ldots, n$);

3. Eliminate the found solutions $\left(\hat{\mathbf{p}}^1, \ldots, \hat{\mathbf{p}}^n\right)$ which are not equilibria, i.e., those for which there exist $i = 1, \ldots, n$, $j \in W^i$and $j' \notin W^i$ such that $\Phi^i\left(j'; \hat{\mathbf{p}}^1, \ldots, \hat{\mathbf{p}}^n\right) > \Phi^i\left(j; \hat{\mathbf{p}}^1, \ldots, \hat{\mathbf{p}}^n\right)$.

This procedure should be repeated for all possible selections $\left(W^1, \ldots, W^n\right)$, as described in step (1).

Some remarks are in order now.

We write in point (2): "find all solutions"; probably in some specific cases one will not be able to find "all" solutions. The phrase "find all solutions" should rather be understood as: "we reduce the problem of finding equilibria to an optimization problem; if all its solutions can be found, it is fine; otherwise, we have at least exported the problem to another branch of mathematics".

The proposed procedure can be usually simplified because of symmetries or other regularities in the data; this will be illustrated in the example below.

Some equilibria may be identified several times, while executing this procedure for different potential supports (which is the case where a found equilibrium has some coordinates equal to 0); if we want to avoid these phenomena, we may simply add a step in the algorithm eliminating such equilibria.

We illustrate the proposed algorithm by an example.

Example 1.1. Suppose that we are given a game $\Gamma = (1, 4; \Phi)$ with one population and four available strategies $1, 2, 3, 4$; the payoff function is defined by $\Phi(j; \mathbf{p}) = \mathbf{p}_j + \frac{1}{3}\left(\sum_{r:|r-j|=1} \mathbf{p}_r\right)$ *(*the upper script ^{1} at Φ^1 may be skipped in this case*)*. We need to consider 15 possible supports W of an equilibrium: $\{1\}, \{2\}, \{3\}, \{4\}, \ldots, \{1, 2, 3, 4\}$. In general, all the 15 cases should be considered separately, but due to the occurring symmetries this number reduces to 9 *(*except *(7)*, *(8)*, and *(15)*, every case has a symmetric counterpart; see the last column in Table 1*)*.

For instance, for $W = \{1\}$ we get the payoff 1 corresponding to the distribution $(1, 0, 0, 0)$ and the use of the first strategy. We only need to check whether the payoffs corresponding to a choice of strategies 2, 3, or 4 do not exceed 1; of course, this is not the case, so $(1, 0, 0, 0)$ is an equilibrium. The case of $W = \{4\}$ is symmetric, so both cases get the same label *(*1*)*.
For $W = \{1, 3\}$ we have to solve, for $\mathbf{p}_1, \mathbf{p}_3$, the simultaneous equations $\mathbf{p}_1 + \frac{1}{3} \cdot \mathbf{p}_2 = \frac{1}{3} \cdot \mathbf{p}_2 + \mathbf{p}_3 + \frac{1}{3} \cdot \mathbf{p}_4$, $\mathbf{p}_1 + \mathbf{p}_3 = 1$. We find the solution $\left(\frac{1}{2}, 0, \frac{1}{2}, 0\right)$; while using a strategy 1 or 3, a player obtains the payoff $\frac{1}{2}$*;* otherwise, using the strategy 2 he obtains the payoff $\frac{1}{3}$ and using the fourth strategy he gets the payoff $\frac{1}{6}$. The case $W = \{2, 4\}$ is symmetric.

The remaining cases are skipped, because their discussion is similar.

As understood in the present paper, a *neighborhood* game is one with one type of players and k strategies $1, 2, \ldots, k$ which are vertices in an undirected graph without loops $G = (\{1, 2, \ldots, k\}, E)$; the payoff function is $\Phi(j; \mathbf{p}) = \mathbf{p}_j + \alpha \cdot \sum_{r:\{r,j\}\in E} \mathbf{p}_r$, α is real. So the payoff for a player who chose a strategy j is $\mathbf{p}_j$ plus the sum of all $\mathbf{p}_r$ such that r is a neighbor of j, multiplied by a coefficient α.

In the present paper, we give a complete analysis of equilibria in neighborhood games in which G is a chain or a cycle while $0 \leq \alpha \leq \frac{1}{2}$.

Table 1: A simple "neighborhood" game.

#	*Support* W	*Equilibrium* $Eq(W)$	*Payoffs for resp. strategies*	*Actual payoff*	*Symmetries*
1	$\{1\}$	$(1,0,0,0)$	$\left(1,\frac{1}{3},0,0\right)$	1	(1)
2	$\{2\}$	$(0,1,0,0)$	$\left(\frac{1}{3},1,\frac{1}{3},0\right)$	1	(2)
3	$\{3\}$	$(0,0,1,0)$	$\left(0,\frac{1}{3},1,\frac{1}{3}\right)$	1	(2)
4	$\{4\}$	$(0,0,0,1)$	$\left(0,0,\frac{1}{3},1\right)$	1	(1)
5	$\{1,2\}$	$\left(\frac{1}{2},\frac{1}{2},0,0\right)$	$\left(\frac{2}{3},\frac{2}{3},\frac{1}{6},0\right)$	$\frac{2}{3}$	(5)
6	$\{1,3\}$	$\left(\frac{1}{2},0,\frac{1}{2},0\right)$	$\left(\frac{1}{2},\frac{1}{3},\frac{1}{2},\frac{1}{6}\right)$	$\frac{1}{2}$	(6)
7	$\{1,4\}$	$\left(\frac{1}{2},0,0,\frac{1}{2}\right)$	$\left(\frac{1}{2},\frac{1}{6},\frac{1}{6},\frac{1}{2}\right)$	$\frac{1}{2}$	(7)
8	$\{2,3\}$	$\left(0,\frac{1}{2},\frac{1}{2},0\right)$	$\left(\frac{1}{6},\frac{2}{3},\frac{2}{3},\frac{1}{6}\right)$	$\frac{2}{3}$	(8)
9	$\{2,4\}$	$\left(0,\frac{1}{2},0,\frac{1}{2}\right)$	$\left(\frac{1}{6},\frac{1}{2},\frac{1}{3},\frac{1}{2}\right)$	$\frac{1}{2}$	(6)
10	$\{3,4\}$	$\left(0,0,\frac{1}{2},\frac{1}{2}\right)$	$\left(0,\frac{1}{6},\frac{2}{3},\frac{2}{3}\right)$	$\frac{2}{3}$	(5)
11	$\{1,2,3\}$	$\left(\frac{2}{5},\frac{1}{5},\frac{2}{5},0\right)$	$\left(\frac{7}{15},\frac{7}{15},\frac{7}{15},\frac{2}{15}\right)$	$\frac{7}{15}$	(11)
12	$\{1,2,4\}$	$\left(\frac{3}{10},\frac{3}{10},0,\frac{4}{10}\right)$	$\left(\frac{12}{30},\frac{12}{30},\frac{7}{30},\frac{12}{30}\right)$	$\frac{12}{30}$	(12)
13	$\{1,3,4\}$	$\left(\frac{4}{10},0,\frac{3}{10},\frac{3}{10}\right)$	$\left(\frac{12}{30},\frac{7}{30},\frac{12}{30},\frac{12}{30}\right)$	$\frac{12}{30}$	(12)
14	$\{2,3,4\}$	$\left(0,\frac{2}{5},\frac{1}{5},\frac{2}{5}\right)$	$\left(\frac{2}{15},\frac{7}{15},\frac{7}{15},\frac{7}{15}\right)$	$\frac{7}{15}$	(11)
15	$\{1,2,3,4\}$	$\left(\frac{3}{10},\frac{2}{10},\frac{2}{10},\frac{3}{10}\right)$	$\left(\frac{11}{30},\frac{11}{30},\frac{11}{30},\frac{11}{30}\right)$	$\frac{11}{30}$	(15)

2 A few tricks

This section has an introductory character; it contains intuitive examples related to Fibonacci numbers; the general theory will be presented in Sec. 4. The proofs of all statements in this section are skipped, because they will be proved in Sec. 4.

Example 2.1. We take the Fibonacci sequence $(F_1, F_2, \ldots) = (1,1,2,3,5,8,13,21,34,\ldots)$ and choose a positive integer k. We form a sequence $\left(f_1^k,\ldots,f_k^k\right)$ of length k letting $\left(f_1^k,\ldots,f_k^k\right) = (F_1F_k, F_2F_{k-1}, F_3F_{k-2}, \ldots, F_{k-1}F_2, F_kF_1)$. We find that, for each k, the value $\frac{1}{3}f_{i-1}^k + f_i^k + \frac{1}{3}f_{i+1}^k$ (for $i=1$ and $i=k$ we supplement the formula taking $f_0^k = f_{k+1}^k = 0$) is constant for all $i = 1,\ldots,k$, although this value is different for different k. Obviously, the sequence $\left(f_1^k,\ldots,f_k^k\right)$ is *symmetric*, i.e., $f_i^k = f_{k+1-i}^k$ for all i.

For instance, for $k=5$ we have $\left(f_1^5, f_2^5, f_3^5, f_4^5, f_5^5\right) = (5,3,4,3,5)$ and the term $\frac{1}{3}f_{i-1}^5 + f_i^5 + \frac{1}{3}f_{i+1}^5$, for $i=2,3,4$ (at boundaries, $f_1^5 + \frac{1}{3}f_2^5$ or $\frac{1}{3}f_4^5 + f_5^5$) is constant and equal to $\frac{18}{3}$.

Another example: for $k=8$, we have $\left(f_1^8, f_2^8, \ldots, f_8^8\right) = (21,13,16,15,15,16,13,21)$ and the term $\frac{1}{3}f_{i-1}^8 + f_i^8 + \frac{1}{3}f_{i+1}^8$, for $i = 2,\ldots,7$, (at boundaries, $f_1^8 + \frac{1}{3}f_2^8$ or $\frac{1}{3}f_7^8 + f_8^8$) is constant and equal to $\frac{76}{3}$.

Let us now form a numerical triangle (Table 2) whose k-th row is composed of elements of the sequence $\left(f_1^k,\ldots,f_k^k\right)$, visually arranged as the Pascal triangle. On the right, we put the corresponding to each row values—equal for all i numbers $\frac{1}{3}f_{i-1}^k + f_i^k + \frac{1}{3}f_{i+1}^k$ (at boundaries $f_1^k + \frac{1}{3}f_2^k$ or $\frac{1}{3}f_{k-1}^k + f_k^k$).

Observe that this triangle may be obtained by the following rule: the first two rows are given; suppose that the first k rows are already constructed; then the row of number $k+1$ is constructed so that its terms are situated as in the Pascal triangle and each of them is defined as a sum of two elements in the two preceding rows, along a diagonal, no matter which. If one element along a diagonal is missing, it is substituted by 0; if two are missing, this diagonal is neglected (this is the case for the border entries). The rule is illustrated by two entries in the triangle, $\boxed{8}$ and $\boxed{16}$, and the entries whose sum they are, along a diagonal, are marked by *, or $^\circ$.

Note that, in the same triangle (Table 3), the entries at the right border, marked by * (as well as those at the left border) form the Fibonacci sequence; the same is true for the entries marked by $^\circ$. The entries marked by $^\times$ form the sequence of Fibonacci numbers multiplied by 3 (the fourth Fibonacci number); those marked by $^\triangleright$ form the sequence of Fibonacci numbers multiplied by 8, the sixth Fibonacci number, etc.

Table 2: A recursive rule to construct the Fibonacci triangle.

#																*Value*
1								1								$3/3$
2							1		1							$4/3$
3						2		1		2						$7/3$
4					3		2		2		3					$11/3$
5				5		3°		4		3		5				$18/3$
6			8^*		5°		6		6^*		5		8°			$29/3$
7		13		$\boxed{8}$		10		9		10^*		8°		13		$47/3$
8	21		13		16		15		15		$\boxed{16}$		13		21	$76/3$
...		...		...		...		...		...		...		...		...

Table 3: Products of Fibonacci numbers.

#																Value
1								1^*								$3/3$
2							1°		1^*							$4/3$
3						2		1°		2^*						$7/3$
4					$3^\times$		2		2°		3^*					$11/3$
5				5		$3^\times$		4		3°		5^*				$18/3$
6			$8^\triangleright$		5		$6^\times$		6		5°		8^*			$29/3$
7		13		$8^\triangleright$		10		$9^\times$		10		8°		13^*		$47/3$
8	21		13		$16^\triangleright$		15		$15^\times$		16		13°		21^*	$76/3$
...		...		...		...		...		...		...		...		...

Example 2.2. We take, instead of the Fibonacci sequence, the sequence $\left(F_1^2, F_2^2, \ldots\right) = (1, 1, 3, 4, 11, 15, 41, 56, 153, \ldots)$, formed by the following rule: $F_1^2 = F_2^2 = 1$; if F_{k-2}^2 and F_{k-1}^2 are already defined then we define $F_k^2 = F_{k-1}^2 + F_{k-2}^2$ if k is even and $F_k^2 = 2F_{k-1}^2 + F_{k-2}^2$ if k is odd (the notation F_i^2 is used to stress similarity to ordinary Fibonacci numbers F_i and also

for the reason of consistency with notation to be introduced in Sec. 4). Choose a positive integer k. We form the sequence $\left(f_1^{2,k},\dots,f_k^{2,k}\right)$ of length k letting $\left(f_1^{2,k},\dots,f_k^{2,k}\right)=\left(F_1^2F_k^2,F_2^2F_{k-1}^2,F_3^2F_{k-2}^2,\dots,\ F_{k-1}^2F_2^2,F_k^2F_1^2\right)$ if k is even and $\left(f_1^{2,k},\dots,f_k^{2,k}\right)=\left(F_1^2F_k^2,2F_2^2F_{k-1}^2,F_3^2F_{k-2}^2,\dots,\ 2F_{k-1}^2F_2^2,F_k^2F_1^2\right)$ if k is odd (entries at odd places i have the form $F_i^2F_{k+1-i}^2$; entries at even places i are $2F_i^2F_{k+1-i}^2$).

Note that, for each k, the value $\frac{1}{4}f_{i-1}^{2,k}+f_i^{2,k}+\frac{1}{4}f_{i+1}^{2,k}$ is constant for all $i=1,\dots,k$, although this value is different for different k (for $i=1$ and $i=k$ we supplement the formula assuming $f_0^{2,k}=f_{k+1}^{2,k}=0$). Obviously, the sequence $\left(f_1^{2,k},\dots,f_k^{2,k}\right)$ is symmetric.

For instance, for $k=5$ we have $\left(f_1^{2,5},f_2^{2,5},f_3^{2,5},f_4^{2,5},f_5^{2,5}\right)=(11,8,9,8,11)$ and the term $\frac{1}{4}f_{i-1}^{2,5}+f_i^{2,5}+\frac{1}{4}f_{i+1}^{2,5}$ (at boundaries, $f_1^{2,k}+\frac{1}{4}f_2^{2,k}$ or $\frac{1}{4}f_{k-1}^{2,k}+f_k^{2,k}$) is constant over i and equal to $\frac{52}{4}$.

For $k=8$ we have $\left(f_1^{2,8},f_2^{2,8},\dots,f_8^{2,8}\right)=(56,41,45,44,44,45,41,56)$ and the term $\frac{1}{4}f_{i-1}^{2,8}+f_i^{2,8}+\frac{1}{4}f_{i+1}^{2,8}$ (at boundaries, $f_1^{2,8}+\frac{1}{4}f_2^{2,8}$ or $\frac{1}{4}f_7^{2,8}+f_8^{2,8}$) is constant over i and equal to $\frac{265}{4}$.

Let us now form a numerical triangle (Table 4) whose k-th row is composed of the elements of the sequence $\left(f_1^{2,k},\dots,f_k^{2,k}\right)$, visually arranged as the Pascal triangle. On the right, we put the corresponding to each row values $\frac{1}{4}f_{i-1}^{2,k}+f_i^{2,k}+\frac{1}{4}f_{i+1}^{2,k}$ (again, all of them equal, at boundaries $f_1^{2,k}+\frac{1}{4}f_2^{2,k}$ or $\frac{1}{4}f_{k-1}^{2,k}+f_k^{2,k}$).

Table 4: A Fibonacci-like triangle.

#																Value
1								1								4/4
2							1		1							5/4
3						3		2		3						14/4
4					4		3°		3		4					19/4
5				11^*		8°		9^*		8		11°				52/4
6			15		$\boxed{11}$		12		*12**		*11*$^\circ$		15			71/4
7		41		30		33		32		$\boxed{33}$		30		41		194/4
8	56		41		45		44		44		45		41		56	265/4
...		...		...		...		...		...		...		...		...

Observe that this triangle may be obtained by the following rule: the first two rows are given; suppose that the first $k-1$ rows are already constructed; then the row of number k is constructed so that its terms are situated as in the Pascal triangle and each of them is defined as follows:

- for even k, as a sum of two elements in the two preceding rows, along a diagonal, no matter which;

- for odd k, as a sum of two elements in the two preceding rows (the one in the row right above multiplied by 2), along a diagonal, no matter which.

If one element along a diagonal is missing, it is substituted by 0; if two are missing, this diagonal is neglected (this is the case for the border entries). The rule is illustrated by two entries in the triangle, $\boxed{11}$ and $\boxed{33}$, and the entries whose (adjusted) sum they are, along a diagonal, are marked by *, or $^\circ$; the entry to be multiplied by 2 is italicized.

3 Some linear algebra

For real α and a positive integer k we define $k \times k$ matrices $D_k^\alpha = (d_{ij})$ and $E_k^\alpha = (e_{ij})$ by formulas

$$d_{ij} = \begin{cases} 1 & \text{if } i = j; \\ \alpha & \text{if } |i-j| = 1; \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad e_{ij} = \begin{cases} 1 & \text{if } i = j; \\ \alpha & \text{if } |i-j| = 1 \text{ or } |i-j| = k-1; \\ 0 & \text{otherwise.} \end{cases}$$

So the matrices D_k^α and E_k^α look like:

$$D_k^\alpha = \begin{bmatrix} 1 & \alpha & 0 & 0 & \vdots & 0 & 0 & 0 \\ \alpha & 1 & \alpha & 0 & \vdots & 0 & 0 & 0 \\ 0 & \alpha & 1 & \alpha & \vdots & 0 & 0 & 0 \\ 0 & 0 & \alpha & 1 & \vdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \ddots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \vdots & 1 & \alpha & 0 \\ 0 & 0 & 0 & 0 & \vdots & \alpha & 1 & \alpha \\ 0 & 0 & 0 & 0 & \vdots & 0 & \alpha & 1 \end{bmatrix},$$

$$E_k^\alpha = \begin{bmatrix} 1 & \alpha & 0 & 0 & \vdots & 0 & 0 & \alpha \\ \alpha & 1 & \alpha & 0 & \vdots & 0 & 0 & 0 \\ 0 & \alpha & 1 & \alpha & \vdots & 0 & 0 & 0 \\ 0 & 0 & \alpha & 1 & \vdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \ddots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \vdots & 1 & \alpha & 0 \\ 0 & 0 & 0 & 0 & \vdots & \alpha & 1 & \alpha \\ \alpha & 0 & 0 & 0 & \vdots & 0 & \alpha & 1 \end{bmatrix}.$$

Theorem 3.1. *For any* $-\frac{1}{2} \leq \alpha \leq \frac{1}{2}$ *and positive integer* k*,* $\det D_k^\alpha > 0$.

Lemma 3.1. *Let* $0 \leq \beta \leq \frac{1}{4}$*. The sequence* (x_k) *defined by* $x_0 = x_1 = 1$ *and, recursively for* $k = 2, 3, \ldots$*,* $x_k = x_{k-1} - \beta x_{k-2}$*, has all terms positive.*

Proof. By definition, x_0 and x_1 are positive and $x_1 > \frac{1}{2}x_0$. We prove, by induction for $k = 2, 3, \ldots$, that x_k is positive and $x_k > \frac{1}{2}x_{k-1}$ (the inequality is necessary to perform the inductive step). So suppose that for some positive k we already know that x_k is positive and $x_k > \frac{1}{2}x_{k-1}$. Then we have (we use below both inductive assumptions, the second in the form $-x_{k-1} > -2x_k$)

$$x_{k+1} = x_k - \beta x_{k-1} \geq x_k - \frac{1}{4}x_{k-1} > x_k - \frac{1}{4} \cdot 2x_k = \frac{1}{2}x_k,$$

and therefore x_{k+1} is positive. $\square$

Proof of Theorem 3.1. Taking the Laplace expansion of D_k^α along the first row, we get a recursive formula which holds for $k = 3, 4, \ldots$:

$$\det D_k^\alpha = \det D_{k-1}^\alpha - \alpha^2 \det D_{k-2}^\alpha.$$

Let $\beta = \alpha^2$ and define a sequence (x_k) letting $x_0 = 1$ and, for $k = 1, 2, \ldots$, $x_k = \det D_k^\alpha$. We have, applying the displayed formula, $x_k = x_{k-1} - \beta x_{k-2}$. Hence, the lemma applies and it follows that all terms of the sequence are positive, which means that $\det D_k > 0$. $\square$

Theorem 3.2. *For any* $-\frac{1}{2} < \alpha < \frac{1}{2}$ *and positive integer* k*, the matrix* E_k^α *is nonsingular. The matrix* $E_k^{-1/2}$ *has rank* $k - 1$ *for all* k*. The matrix* $E_k^{1/2}$ *is nonsingular for odd* k*; for even* k*,* $k \geq 4$*, it has the rank* $k - 1$.

Proof. Fix $-\frac{1}{2} < \alpha < \frac{1}{2}$ and a positive integer k. Suppose that E_k^α is singular. Then there exists a nonzero vector $\mathbf{x}$ such that $E_k^\alpha \cdot \mathbf{x} = \mathbf{0}$, i.e., $x_j = -\alpha x_{j-1} - \alpha x_{j-1}$ for $j = 1, \ldots, k$ (for $j = 1$ and $j = k$ some indices at the right side should be understood modulo k). For at least one j, $x_j \neq 0$ in which case we have $|x_j| < \frac{1}{2}|x_{j-1}| + \frac{1}{2}|x_{j+1}|$; for all remaining j we have $|x_j| \leq \frac{1}{2}|x_{j-1}| + \frac{1}{2}|x_{j+1}|$. Summing up the sides of all these inequalities over $j = 1, \ldots, k$ we get a contradiction $\sum_{j=1}^k |x_j| < \sum_{j=1}^k |x_j|$.

The matrix $E_k^{-1/2}$ is singular because $E_k^{-1/2} \cdot [1, 1, \ldots, 1]^\mathsf{T} = \mathbf{0}$. The rank of $E_k^{-1/2}$ is $k - 1$ because $E_k^{-1/2}$ has a submatrix $D_{k-1}^{-1/2}$ which, by Theorem 3.1, is nonsingular.

For even $k \geq 4$, the matrix $E_k^{1/2}$ is singular, because $E_k^{1/2} \cdot [1, -1, 1, -1, \ldots, 1, -1]^\mathsf{T} = \mathbf{0}$. The rank of $E_k^{1/2}$ is $k - 1$ because $E_k^{1/2}$ has a submatrix $D_{k-1}^{1/2}$ which, by Theorem 3.1, is nonsingular.

Finally, we prove that $E_k^{1/2}$ is nonsingular for odd k. Suppose not, then there exists a nonzero vector $\mathbf{x}$ such that $E_k^{1/2} \cdot \mathbf{x} = \mathbf{0}$. Let j be such that the absolute value $|x_j|$ is maximal among $|x_1|, \ldots, |x_k|$. Because of symmetry we may assume, without loss off generality, that j is equal to 1 and $x_1 = 1$. We have then $1 = -\frac{1}{2}x_k - \frac{1}{2}x_2$, which is possible only if $x_k = -1$ and $x_2 = -1$. Repeating this argument, we prove inductively, for $r = 2, 3, \ldots, k$, that $x_r = 1$ for odd r and $x_r = -1$ for even r; hence also $x_k = 1$, a contradiction. □

4 Exact sequences

Let α be real and let k be a positive integer. A sequence of reals $(x_1, \ldots, x_k)$ is α-*exact* if for each $i = 1, \ldots, k$ and $j = 1, \ldots, k$, $\alpha x_{i-1} + x_i + \alpha x_{i+1} = \alpha x_{j-1} + x_j + \alpha x_{j+1}$ (at boundary entries we supplement by $x_0 = x_{k+1} = 0$).

4.1 Generalized Fibonacci sequences

Let A be real. A *generalized Fibonacci sequence with parameter* A (for short, *Fibonacci* A-*sequence*) is the sequence of reals $\left(F_1^A, F_2^A, F_3^A, \ldots\right)$ such that $F_1^A = F_2^A = 1$ and $F_i^A = A \cdot F_{i-1}^A + F_{i-2}^A$ for odd i while $F_i^A = F_{i-1}^A + F_{i-2}^A$ for even i.

The *standard* (A, k)-*sequence* is the sequence $f^{Ak} = \left(f_i^{Ak}\right)_{i=1}^{k}$ defined as follows:

for odd k and odd $i = 1, 3, 5, \ldots, k$, $f_i^{Ak} = F_i^A \cdot F_{k+1-i}^A$;

for odd k and even $i = 2, 4, \ldots, k-1$, $f_i^{Ak} = A \cdot F_i^A \cdot F_{k+1-i}^A$;

for even k and $i = 1, 2, \ldots, k$, $f_i^{Ak} = F_i^{Ak} \cdot F_{k+1-i}^{Ak}$.

More visually:

a standard $(A, 6)$-sequence looks like:

$$\left(F_1^A \cdot F_6^A, F_2^A \cdot F_5^A, F_3^A \cdot F_4^A, F_4^A \cdot F_3^A, F_5^A \cdot F_2^A, F_6^A \cdot F_1^A\right);$$

a standard $(A, 7)$-sequence looks like:

$$\left(F_1^A \cdot F_7^A, A \cdot F_2^A \cdot F_6^A, F_3^A \cdot F_5^A, A \cdot F_4^A \cdot F_4^A, F_5^A \cdot F_3^A, A \cdot F_6^A \cdot F_2^A, F_7^A \cdot F_1^A\right).$$

Clearly, every standard sequence is *symmetric*, i.e., $f_i^{Ak} = f_{k+1-i}^{Ak}$, for all A, k and $i = 1, 2, \ldots, k$. The ordinary Fibonacci sequence is the same as the generalized Fibonacci 1-sequence.

Remark 4.1. Every constant sequence is 0-exact and there are no other 0-exact sequences. □

Proposition 4.1. *If $A > 0$ then, for each k, all elements of the standard (A, k)-sequence are positive. For each even k, all elements of the standard $(0, k)$-sequence are positive; for odd k, the standard $(0, k)$-sequence is* $(1, 0, 1, 0, \ldots, 1, 0, 1)$. □

Theorem 4.1. *Let $A \neq -2$ and let k be a positive integer. The standard (A, k)-sequence is $\left(\frac{1}{A+2}\right)$-exact.*

Proof. Choose A and k, as indicated. Denote $\frac{1}{A+2} = \alpha$. We need to prove that

$$\alpha \cdot f_{t-1}^{Ak} + f_t^{Ak} + \alpha \cdot f_{t+1}^{Ak} = \alpha \cdot f_t^{Ak} + f_{t+1}^{Ak} + \alpha \cdot f_{t+2}^{Ak}, \text{ for } t = 2, 3, \dots, k-1, \quad (1)$$

$$f_1^{Ak} + \alpha \cdot f_2^{Ak} = \alpha \cdot f_1^{Ak} + f_2^{Ak} + \alpha \cdot f_3^{Ak} \quad (2)$$

and

$$\alpha \cdot f_{k-2}^{Ak} + f_{k-1}^{Ak} + \alpha \cdot f_k^{Ak} = \alpha \cdot f_{k-1}^{Ak} + f_k^{Ak}. \quad (3)$$

To reduce the number of cases to be considered, we augment the original Fibonacci A-sequence with an initial term $F_0^A = 0$, so instead we get the sequence $(F_0^A, F_1^A, F_2^A, \dots)$; note that this sequence is defined by $F_0^A = 0, F_1^A = 1$ and exactly the same recursive formula as in the definition of a Fibonacci A-sequence: $F_i^A = AF_{i-1}^A + F_{i-2}^A$ for odd i and $F_i^A = F_{i-1}^A + F_{i-2}^A$ for even i. The standard sequence is then augmented to $(f_0^{Ak}, f_1^{Ak}, \dots, f_k^{Ak}, f_{k+1}^{Ak})$, defined in the same manner as in the original definition: the elements f_i^{Ak}, $i = 1, \dots, k$, remain unchanged while $f_0^{Ak} = f_{k+1}^{Ak} = 0$ (for even k we have $f_0^{Ak} = F_0^A \cdot F_{k+1}^A = 0$, $f_{k+1}^{Ak} = F_{k+1}^A \cdot F_0^A = 0$; for odd k we have $f_0^{Ak} = A \cdot F_0^A \cdot F_{k+1}^A = 0$, $f_{k+1}^{Ak} = A \cdot F_{k+1}^A \cdot F_0^A = 0$). Hence, instead of proving (1), (2), and (3) we actually have to prove that (1) holds for $i = 1, 2, 3, \dots, k$, for the augmented standard sequence.

Case 1. k is even. Then we have the augmented standard sequence

$$\left(F_0^A \cdot F_{k+1}^A, F_1^A \cdot F_k^A, F_2^A \cdot F_{k-1}^A, F_3^A \cdot F_{k-2}^A, \dots, F_{k-1}^A H \cdot F_2^A, F_k^A \cdot F_1^A, F_{k+1}^A \cdot F_0^A\right),$$

and we have to prove that, for $i = 1, 2, 3, \dots, k$,

$$\begin{aligned} &\alpha \cdot F_{i-1}^A \cdot F_{k+2-i}^A + F_i^A \cdot F_{k+1-i}^A + \alpha \cdot F_{i+1}^A \cdot F_{k-i}^A \\ &\quad = \alpha \cdot F_i^A \cdot F_{k+1-i}^A + F_{i+1}^A \cdot F_{k-i}^A + \alpha \cdot F_{i+2}^A \cdot F_{k-1-i}^A. \end{aligned} \quad (4)$$

Formally, we should consider two cases: where i is even and where i is odd. The two cases are symmetric, so we shall only deal with the first one.

We denote $F_{i-1}^A = a, \quad F_i^A = b, \quad F_{k-1-i}^A = c, \quad F_{k-i}^A = d.$

Then we have (keep in mind that now $k - i$ is even):

$$\begin{aligned} F_{i+1}^A &= A \cdot F_i^A + F_{i-1}^A = a + A \cdot b; \\ F_{i+2}^A &= F_{i+1}^A + F_i^A = (A+1) \cdot F_i^A + F_{i-1}^A = a + (A+1) \cdot b; \\ F_{k+1-i}^A &= A \cdot F_{k-i}^A + F_{k-1-i}^A = c + A \cdot d; \\ F_{k+2-i}^A &= F_{k+1-i}^A + F_{k-i}^A = (A+1) \cdot F_{k-i}^A + F_{k-1-i}^A = c + (A+1) \cdot d. \end{aligned}$$

Substituting in (4), also $\alpha = \frac{1}{A+2}$, we get:

$$\frac{1}{A+2} \cdot a \cdot [c + (A+1) \cdot d] + b \cdot (c + A \cdot d) + \frac{1}{A+2} \cdot (a + A \cdot b) \cdot d =$$

$$= \frac{1}{A+2} \cdot b \cdot (c + A \cdot d) + (a + A \cdot b) \cdot d + \frac{1}{A+2} \cdot [a + (A+1) \cdot b] \cdot c,$$

which is a tautology that holds for all real numbers a, b, c, d and $A \neq -2$.

Case 2. k is odd. Then we have the augmented standard sequence

$$\big(A \cdot F_0^A \cdot F_{k+1}^A, F_1^A \cdot F_k^A, A \cdot F_2^A \cdot F_{k-1}^A, F_3^A \cdot F_{k-2}^A, \dots,$$

$$A \cdot F_{k-1}^A \cdot F_2^A, F_k^A \cdot F_1^A, A \cdot F_{k+1}^A \cdot F_0^A\big) .$$

For even $i = 2, 4, \dots, k-1$ we have to prove that

$$\alpha \cdot F_{i-1}^A \cdot F_{k+2-i}^A + A \cdot F_i^A \cdot F_{k+1-i}^A + \alpha \cdot F_{i+1}^A \cdot F_{k-i}^A =$$

$$= \alpha \cdot A \cdot F_i^A \cdot F_{k+1-i}^A + F_{i+1}^A \cdot F_{k-i}^A + \alpha \cdot A \cdot F_{i+2}^A \cdot F_{k-1-i}^A. \qquad (5)$$

For odd $i = 1, 3, \dots, k$ we have to prove that

$$\alpha \cdot A \cdot F_{i-1}^A \cdot F_{k+2-i}^A + F_i^A \cdot F_{k+1-i}^A + \alpha \cdot A \cdot F_{i+1}^A \cdot F_{k-i}^A =$$

$$= \alpha \cdot F_i^A \cdot F_{k+1-i}^A + A \cdot F_{i+1}^A \cdot F_{k-i}^A + \alpha \cdot F_{i+2}^A \cdot F_{k-1-i}^A. \qquad (6)$$

Again, the two cases are symmetric and we shall only deal with the first one. As in the previous case, we denote $F_{i-1}^A = a, \quad F_i^A = b, \quad F_{k-1-i}^A = c, \quad F_{k-i}^A = d.$

Then we have (keep in mind that now $k - i$ is odd):
$F_{i+1}^A = A \cdot F_i^A + F_{i-1}^A = a + A \cdot b$;
$F_{i+2}^A = F_{i+1}^A + F_i^A = (A+1) \cdot F_i^A + F_{i-1}^A = a + (A+1) \cdot b$;
$F_{k+1-i}^A = F_{k-i}^A + F_{k-1-i}^A = c + d$;
$F_{k+2-i}^A = A \cdot F_{k+1-i}^A + F_{k-1-i}^A = (A+1) \cdot F_{k-i}^A + A \cdot F_{k-1-i}^A = A \cdot c + (A+1) \cdot d.$

Substituting in (5), also $\alpha = \frac{1}{A+2}$, we get:

$$\frac{1}{A+2} \cdot a \cdot [A \cdot c + (A+1) \cdot d] + A \cdot b \cdot (c+d) + \frac{1}{A+2} \cdot (a + A \cdot b) \cdot d =$$

$$= \frac{1}{A+2} \cdot A \cdot b \cdot (c+d) + (a + A \cdot b) \cdot d + \frac{1}{A+2} \cdot A \cdot [a + (A+1) \cdot b] \cdot c,$$

which is a tautology that holds for all real numbers a, b, c, d and $A \neq -2$.
This completes the proof of the theorem. □

Corollary 4.1. *For each positive integer k and $\alpha \neq 0$, the sequence f^{Ak}, where $A = \frac{1}{\alpha} - 2$, is α-exact. For any $-\frac{1}{2} \leq \alpha \leq \frac{1}{2}$, $\alpha \neq 0$, any other α-exact sequence of length k is proportional to f^{Ak}.*

Proof. The last statement follows from Theorem 3.1. □

Table 5: The Fibonacci-like triangle – general case.

				f_1^{A1} [1]				
			f_1^{A2} [1]		f_2^{A2} [1]			
		f_1^{A3} $[A+1]$		f_2^{A3} $[A]$		f_3^{A3} $[A+1]$		
	f_1^{A4} $[A+2]$		f_2^{A4} $[A+1]$		f_3^{A4} $[A+1]$		f_4^{A4} $[A+2]$	
...		...		...		...		...

4.2 Pascal-like triangles

We now turn back to the subject of Pascal-like triangles, already tackled in Sec. 2. We show the triangle (Table (5) whose k-th row consists of consecutive elements of the standard (A, k)-sequence. Right below, in square brackets, we give the explicit formula.

Theorem 4.2. *Let $k \geq 5$ be an odd positive integer. Then:*

a. $f_1^{Ak} = A \cdot f_1^{A,k-1} + f_1^{A,k-2}$;
b. $f_2^{Ak} = A \cdot f_1^{A,k-1} = A \cdot f_2^{A,k-1} + f_2^{A,k-2}$;
c. $f_i^{Ak} = A \cdot f_{i-1}^{A,k-1} + f_{i-2}^{A,k-2} = A \cdot f_i^{A,k-1} + f_i^{A,k-2}$, *for* $i = 3, 4, \ldots, k-2$;
d. $f_{k-1}^{Ak} = A \cdot f_{k-2}^{A,k-1} + f_{k-3}^{A,k-2} = A \cdot f_{k-1}^{A,k-1}$;
e. $f_k^{Ak} = A \cdot f_{k-1}^{A,k-1} + f_{k-2}^{A,k-2}$.

Theorem 4.3. *Let $k \geq 4$ be an even positive integer. Then:*

a. $f_1^{Ak} = f_1^{A,k-1} + f_1^{A,k-2}$;
b. $f_2^{Ak} = f_1^{A,k-1} = f_2^{A,k-1} + f_2^{A,k-2}$;
c. $f_i^{Ak} = f_{i-1}^{A,k-1} + f_{i-2}^{A,k-2} = f_i^{A,k-1} + f_i^{A,k-2}$, *for* $i = 3, 4, \ldots, k-2$;
d. $f_{k-1}^{Ak} = f_{k-2}^{A,k-1} + f_{k-3}^{A,k-2} = f_{k-1}^{A,k-1}$;
e. $f_k^{Ak} = f_{k-1}^{A,k-1} + f_{k-2}^{A,k-2}$.

Analogous formulas to those in Theorems 4.2 and 4.3 are also true for $k = 3$ and $k = 2$; they are (immediately checked):
4.2.f ($k = 3$) $f_1^{A3} = A \cdot f_1^{A2} + f_1^{A1}$, $f_2^{A3} = A \cdot f_1^{A2} = A \cdot f_2^{A2}$, $f_3^{A3} = A \cdot f_2^{A2} + f_1^{A1}$;
4.3.f ($k = 2$) $f_1^{A2} = f_1^{A1}$, $f_2^{A2} = f_1^{A1}$. □

Proof of Theorem 4.2. Because of the symmetry of the triangle we only need to prove that:

(X) for $i = 1, 2, \ldots, k-2$, $f_i^{Ak} = A \cdot f_i^{A,k-1} + f_i^{A,k-2}$

and

(Y) $f_{k-1}^{Ak} = A \cdot f_{k-1}^{A,k-1}$.

We shall consider separately two cases of (X): (XO) where i is odd and (XE) where i is even.

To prove (XO), we substitute (according to the definition of standard sequences) in (X) $f_i^{Ak} = F_i^A \cdot F_{k+1-i}^A, f_i^{A,k-1} = F_i^A \cdot F_{k-i}^A, f_i^{A,k-2} = F_i^A \cdot F_{k-1-i}^A$ and we get $F_i^A \cdot F_{k+1-i}^A = A \cdot F_i^A \cdot F_{k-i}^A + F_i^A \cdot F_{k-1-i}^A$ which is obviously true, because, by definition of an (A,k)-sequence, $F_{k+1-i}^A = A \cdot F_{k-i}^A + F_{k-1-i}^A$ (clearly, here $k+1-i$ is odd).

To prove (XE), we substitute (according to the definition of standard sequences) in (X) $f_i^{Ak} = A \cdot F_i^A \cdot F_{k+1-i}^A, f_i^{A,k-1} = F_i^A \cdot F_{k-i}^A, f_i^{A,k-2} = A \cdot F_i^A \cdot F_{k-1-i}^A$ and we get $A \cdot F_i^A \cdot F_{k+1-i}^A = A \cdot F_i^A \cdot F_{k-i}^A + A \cdot F_i^A \cdot F_{k-1-i}^A$ which is obviously true, because, by definition of an (A,k)-sequence, $F_{k+1-i}^A = F_{k-i}^A + F_{k-1-i}^A$ (clearly, here $k+1-i$ is even).

To prove (Y), we substitute in (Y), according to the definition of an (A,k)-sequence, $f_{k-1}^{Ak} = A \cdot F_2^A \cdot F_{k-1}^A$ and $f_{k-1}^{A,k-1} = F_1^A \cdot F_{k-1}^A$, so we get $A \cdot F_2^A \cdot F_{k-1}^A = A \cdot F_1^A \cdot F_{k-1}^A$ which is obviously true, because $F_1^A = F_2^A$. □

Proof of Theorem 4.3. Because of the symmetry of the triangle we only need to prove that:

(Z) for $i = 1, 2, \ldots, k-2$, $f_i^{Ak} = f_i^{A,k-1} + f_i^{A,k-2}$

and

(T) $f_{k-1}^{Ak} = f_{k-1}^{A,k-1}$.

We shall consider separately two cases of (Z): (ZO) where i is odd and (ZE) where i is even.

To prove (ZO), we substitute (according to the definition of standard sequences) in (Z) $f_i^{Ak} = F_i^A \cdot F_{k+1-i}^A, f_i^{A,k-1} = F_i^A \cdot F_{k-i}^A, f_i^{A,k-2} = F_i^A \cdot F_{k-1-i}^A$ and we get $F_i^A \cdot F_{k+1-i}^A = F_i^A \cdot F_{k-i}^A + F_i^A \cdot F_{k-1-i}^A$ which is obviously true, because, by definition of an (A,k)-sequence, $F_{k+1-i}^A = F_{k-i}^A + F_{k-1-i}^A$ (clearly, here $k+1-i$ is even).

To prove (ZE), we substitute (according to the definition of standard sequences) in (Z) $f_i^{Ak} = F_i^A \cdot F_{k+1-i}^A, f_i^{A,k-1} = A \cdot F_i^A \cdot F_{k-i}^A, f_i^{A,k-2} = F_i^A \cdot F_{k-1-i}^A$ and we get $F_i^A \cdot F_{k+1-i}^A = A \cdot F_i^A \cdot F_{k-i}^A + F_i^A \cdot F_{k-1-i}^A$ which is obviously true, because, by definition of an (A,k)-sequence, $F_{k+1-i}^A = A \cdot F_{k-i}^A + F_{k-1-i}^A$ (clearly, here $k+1-i$ is odd).

To prove (T), we substitute in (T), according to the definition of an (A,k)-sequence, $f_{k-1}^{Ak} = F_2^A \cdot F_{k-1}^A$ and $f_{k-1}^{A,k-1} = F_1^A \cdot F_{k-1}^A$ and we get $F_2^A \cdot F_{k-1}^A = F_1^A \cdot F_{k-1}^A$ which is obviously true, because $F_1^A = F_2^A$. □

Theorems 4.2 and 4.3 actually describe a recursive procedure, extending this already mentioned in Sec. 2, allowing for an alternative construction of standard sequences: we first fix the first row of the triangle consisting of a single element 1, the second row consisting of terms of the sequence $(1,1)$, and then we use the recursive formulas to construct the consecutive rows (standard sequences): the newly defined element is a sum of two preceding elements, along (any) diagonal, the immediate predecessor multiplied by A or 1, according to whether the number of the row is odd or even.

4.3 Value and volume of a sequence

Proposition 4.2. *1. Sequences $(0,0,\ldots,0)$, (t) and (t,t) (t is real) are α-exact for every real α.*
2. If any other sequence is α-exact and β-exact then $\alpha=\beta$.

Proof. (1) is obvious. To prove (2), suppose that a sequence $\mathbf{x}=(x_1,\ldots,x_k)$ is α- and β-exact. Then, in particular, we have $x_1+\alpha x_2=\alpha x_1+x_2+\alpha x_3$ and $x_1+\beta x_2=\beta x_1+x_2+\beta x_3$ which implies $(\alpha-\beta)(x_1-x_2+x_3)=0$. If $x_1-x_2+x_3=0$ then $\mathbf{x}$ must have the form $(a,b,b-a,\ldots)$. Equality $a+\alpha b=\alpha a+b+\alpha(b-a)$ (following from α-exactness) implies $a=b$; the sequence of the form $(a,a,0,\ldots)$ is either one of those in (1) or else $\alpha,\beta\neq 0$ and $x_4=\frac{a}{\alpha}=\frac{a}{\beta}$, hence $\alpha=\beta$). If $x_1-x_2+x_3\neq 0$ then also $\alpha=\beta$. □

The *value* of an α-exact sequence $\mathbf{x}=(x_1,\ldots,x_k)$ (denoted by $\mathrm{Val}(\mathbf{x})$) is the number $x_1+\alpha x_2$ (if the length of the sequence is 1 then $\mathrm{Val}(\mathbf{x})=x_1$). For sequences of the form (t,t) (t - real) the term "value" should rather be specified to "α-*value*" (for different α, their α-values are different). The values of sequences $(0,0,\ldots,0)$ are obviously 0; the value of (t) is t. Otherwise, in view of Proposition 9, the α of an α-exact sequence is unique and therefore the term "value" is sufficient.

Equalization of an exact sequence $\mathbf{x}=(x_1,\ldots,x_k)$ with nonzero value is the sequence $\left(x_1\cdot\mathrm{Val}(\mathbf{x})^{-1},\ldots,x_k\cdot\mathrm{Val}(\mathbf{x})^{-1}\right)$. In the case of sequences of the form (t,t), $t\neq 0$, we should rather speak of their α-*equalizations*.

Proposition 4.3. *Let $A\neq -2$ and let k be a positive integer. The value of the standard sequence f^{Ak} is equal to $F_k^A+\frac{1}{A+2}\cdot F_{k-1}^A$ if k is even and $F_k^A+\frac{A}{A+2}\cdot F_{k-1}^A$ if k is odd.* □

The *volume* of any sequence $\mathbf{x}=(x_1,\ldots,x_k)$ of reals is the sum of its elements. It is denoted by $\mathrm{Vol}(\mathbf{x})$.

Proposition 4.4. *Let $A\neq -2$ and $A\neq -4$ and let k be a positive integer.*
1. If k is odd then the volume of the standard sequence f^{Ak} is equal to

$$\frac{kA+2k+2}{A+4}\cdot F_k^A+\frac{k}{A+4}\cdot F_{k-1}^A.$$

2. If k is even then the volume of the standard sequence f^{Ak} is equal to

$$\frac{kA+2k+2}{A+4}\cdot F_k^A+\frac{kA}{A+4}\cdot F_{k-1}^A.$$

Proof. We have

$$\left(f_1^{Ak}+\frac{1}{A+2}f_2^{Ak}\right)+\left(\frac{1}{A+2}f_1^{Ak}+f_2^{Ak}+\frac{1}{A+2}f_1^{Ak}\right)+\cdots$$
$$+\left(\frac{1}{A+2}f_{k-1}^{Ak}+f_k^{Ak}\right)=k\cdot\mathrm{Val}\left(f^{Ak}\right),$$

hence

$$\mathrm{Vol}\left(f^{Ak}\right) + \frac{2}{A+2}\mathrm{Vol}\left(f^{Ak}\right) = k \cdot \mathrm{Val}\left(f^{Ak}\right) + \frac{1}{A+2}f_1^{Ak} + \frac{1}{A+2}f_k^{Ak}.$$

Equivalently,

$$\frac{A+4}{A+2} \cdot \mathrm{Vol}\left(f^{Ak}\right) = k \cdot \mathrm{Val}\left(f^{Ak}\right) + \frac{2}{A+2}f_1^{Ak}$$

or

$$\mathrm{Vol}\left(f^{Ak}\right) = \frac{A+2}{A+4} \cdot k \cdot \mathrm{Val}\left(f^{Ak}\right) + \frac{2}{A+4}f_1^{Ak}.$$

Substituting $f_1^{Ak} = F_k^A$ and, $\mathrm{Val}\left(f^{Ak}\right) = F_k^A + \frac{1}{A+2} \cdot F_{k-1}^A$ for even k or $\mathrm{Val}\left(f^{Ak}\right) = F_k^A + \frac{A}{A+2} \cdot F_{k-1}^A$ for odd k, we get the hypothesis. □

5 Equilibria of "neighborhood" games

In this section we shall deal with two types of "neighborhood" games, where the underlying graph is a chain or a cycle. In the present paper we restrict our attention to the case where the parameter α is in the interval $\left[0, \frac{1}{2}\right]$.

5.1 "Chain" games

The "chain" games are formally defined as ones of the form $\Gamma = (1, k; \Phi)$ with the payoff function

$$\Phi(j; \mathbf{p}) = \begin{cases} \mathbf{p}_1 + \alpha \cdot \mathbf{p}_2, & \text{if } j = 1; \\ \alpha \cdot \mathbf{p}_{j-1} + \mathbf{p}_j + \alpha \cdot \mathbf{p}_{j+1}, & \text{if } j = 2, \ldots, k-1; \\ \alpha \cdot \mathbf{p}_{k-1} + \mathbf{p}_k, & \text{if } j = k, \end{cases}$$

for some real number α. So, a "chain" game is uniquely determined by the pair (k, α).

Let us now fix a positive integer k and $\alpha \in \left[0, \frac{1}{2}\right]$, so we fix the game. For a nonempty set $S \subseteq \{1, \ldots, k\}$, a *block* is a maximal (w. r. t. inclusion) subset B of S such that for every $j, j' \in B$ and $j \leq r \leq j'$, also $r \in B$. Every nonempty set $S \subseteq \{1, \ldots, k\}$ is a union of disjoint blocks; we denote, for $j = 1, \ldots, k$, by $\chi_j(S)$ the number of blocks in S with exactly j elements. For each nonempty set $S \subseteq \{1, \ldots, k\}$ we define a vector $\mathbf{p}^S = \left(\mathbf{p}_1^S, \ldots, \mathbf{p}_1^S\right)$ as follows:

Case 1. $\alpha > 0$. Write $A = \frac{1}{\alpha} - 2$. We denote $H(S) = \sum_{r=1}^k \chi_r(S) \cdot \mathrm{Val}\left(f^{Ar}\right)^{-1} \cdot \mathrm{Vol}\left(f^{Ar}\right)$. If $t \notin S$ we set $\mathbf{p}_t^S = 0$; otherwise we define $\mathbf{p}_t^S = f_v^{Au} \cdot \mathrm{Val}\left(f^{Au}\right)^{-1} \cdot H(S)^{-1}$, where u and v are such that t is the v-th consecutive element of a block with u elements.

Case 2. $\alpha = 0$. We define $\mathbf{p}_i^S = (\#S)^{-1}$ if $i \in S$ and $\mathbf{p}_i^S = 0$ otherwise.

Theorem 5.1. *Let Γ be a "chain" game with parameters k and $\alpha \in \left[0, \frac{1}{2}\right)$. For each nonempty set $S \subseteq \{1, \ldots, k\}$, the vector $\mathbf{p}^S$ is the unique equilibrium in Γ whose support is equal to S. So, the game has exactly $2^k - 1$ equilibria.*

Proof. For $\alpha = 0$ the hypothesis is obvious. In the opposite case, write $A = \frac{1}{\alpha} - 2$. Choose $S \subseteq \{1, \ldots, k\}$, $S \neq \emptyset$. By Theorems 4.1 and 3.1, the quantities assigned to consecutive elements in each block in S with r elements must be proportional to entries of the sequence f^{Ar}. The payoffs for strategies in different blocks must be equal, by the definition of equilibrium, so the sequences of the assigned quantities should be equalized. Finally, all assigned quantities must be normalized to sum up to 1. This is exactly what has been performed while constructing distributions $\mathbf{p}^S$. By Proposition 4, the coordinates of $\mathbf{p}_j^S$ are positive if $j \in S$ and they are 0 otherwise.

Let $j \notin S$. If the two neighbors of j belong to S then the payoff at j, equal to $\alpha \cdot \mathbf{p}_{j-1}^S + \alpha \cdot \mathbf{p}_{j+1}^S$, is less than $\max\{\mathbf{p}_{j-1}^S, \mathbf{p}_{j+1}^S\}$, hence it is less than the payoff for at least one of the strategies $j-1$ and $j+1$. If j has only one or zero neighbors in S, the inequality holds as well which means that $\mathbf{p}^S$ is actually an equilibrium.

By construction, the equilibrium with the support S is unique. □

5.2 "Circular" games

The "circular" games are defined as ones of the form $\Gamma = (1, k; \Phi)$ with the payoff function

$$\Phi(j; \mathbf{p}) = \begin{cases} \alpha \cdot \mathbf{p}_k + \mathbf{p}_1 + \alpha \cdot \mathbf{p}_2, & \text{if } j = 1; \\ \alpha \cdot \mathbf{p}_{j-1} + \mathbf{p}_j + \alpha \cdot \mathbf{p}_{j+1}, & \text{if } j = 2, \ldots, k-1; \\ \alpha \cdot \mathbf{p}_{k-1} + \mathbf{p}_k + \alpha \cdot \mathbf{p}_1, & \text{if } j = k, \end{cases}$$

for some real number α. So, a "circular" game is uniquely determined by the pair (k, α).

Let us now fix a positive integer k and $\alpha \in \left[0, \frac{1}{2}\right]$, so we fix the game. In this case, we define a "block" in a somewhat different manner. For a nonempty set $S \subseteq \{1, \ldots, k\}$, a *C-block* is a maximal (w. r. t. inclusion) subset B of S such that for every $j, j' \in B$ and $j \leq r \leq j'$, also $r \in B$; a *C-block* is also a set of the form $\{r, r+1, \ldots, k-1, k, 1, \ldots, s\}$, if $s < r - 1$. By definition, the set $\{1, 2, \ldots, k\}$ is not a C-block. Every nonempty set $S \subset \{1, \ldots, k\}$ is a union of disjoint C-blocks; we denote, for $j = 1, \ldots, k$, by $\dot{\chi}_j(S)$ the number of C-blocks in S with exactly j elements. For each nonempty set $S \subseteq \{1, \ldots, k\}$ we define a vector $\dot{\mathbf{p}}^S = \left(\dot{\mathbf{p}}_1^S, \ldots, \dot{\mathbf{p}}_1^S\right)$ as follows:

Case 1. $\alpha > 0$. *a.* If $S = \{1, 2, \ldots, k\}$ then we define $\dot{\mathbf{p}}^S = \left(k^{-1}, k^{-1}, \ldots, k^{-1}\right)$.

b. Suppose that $S \neq \{1, 2, \ldots, k\}$; denote $A = \frac{1}{\alpha} - 2$. We also denote $\dot{H}(S) = \sum_{r=1}^{k} \dot{\chi}_r(S) \cdot \mathrm{Val}\left(f^{Ar}\right)^{-1} \cdot \mathrm{Vol}\left(f^{Ar}\right)$. If $t \notin S$ we set $\mathbf{p}_t^S = 0$; otherwise we define $\dot{\mathbf{p}}_t^S = f_v^{Au} \cdot \mathrm{Val}\left(f^{Ar}\right)^{-1} \cdot \dot{H}(S)^{-1}$, where u and v are such that t is the v-th consecutive element of a C-block with u elements.

Case 2. $\alpha = 0$. We define $\dot{\mathbf{p}}_i^S = (\#S)^{-1}$ if $i \in S$ and otherwise $\dot{\mathbf{p}}_i^S = 0$.

Theorem 5.2. *Let* Γ *be a "circular" game with parameters* k *and* $\alpha \in \left[0, \frac{1}{2}\right)$. *For each nonempty set* $S \subseteq \{1, \ldots, k\}$, *the vector* $\dot{\mathbf{p}}^S$ *is the unique equilibrium in* Γ *whose support is equal to* S. *So, the game has exactly* $2^k - 1$ *equilibria.*

Proof. The case $\alpha = 0$ is trivial. Otherwise, choose $S \neq \emptyset$; if $S = \{1, 2, \ldots, k\}$ then, by Theorem 3.2, the uniqueness in this case holds. For $S \neq \{1, 2, \ldots, k\}$ the proof is exactly the same as that of Theorem 5.1. □

5.3 The case $\alpha = \frac{1}{2}$

We say that a nonempty set $S \subseteq \{1, \ldots, k\}$ is *regular* if its all blocks are singletons or else they have an even number of elements. We say that a nonempty set $S \subset \{1, \ldots, k\}$ is C-*regular* if its all C-blocks are singletons or else they have an even number of elements.

Theorem 5.3. *Let* Γ *be a "chain" game with parameters* k *and* $\alpha = \frac{1}{2}$. *For each nonempty regular set* $S \subseteq \{1, \ldots, k\}$, *the vector* $\mathbf{p}^S$ *is the unique equilibrium in* Γ, *whose support is equal to* S. *There are no other equilibria in this game.*

Proof. If a set of strategies S is not regular then the standard sequence corresponding to a block with an odd number of elements r, $r > 1$, has the form $(1, 0, 1, \ldots, 0, 1)$, hence the support of the equilibrium found for S is properly included in S; this equilibrium will be found again for another S', properly included in S. □

Theorem 5.4. *Let* Γ *be a "circular" game with parameters* k *and* $\alpha = \frac{1}{2}$. *For each nonempty set* $S \subseteq \{1, \ldots, k\}$, *equal to* $\{1, \ldots, k\}$ *or C-regular, the vector* $\dot{\mathbf{p}}^S$ *is the unique equilibrium in* Γ, *whose support is equal to* S. *There are no other equilibria in this game.*

Proof. Almost the same as the proof of Theorem 5.3. □

6 Final remarks

Generalizations of Fibonacci sequences are too numerous to quote or discuss them here. An interested reader may consult articles in *The Fibonacci Quarterly* [1]. The games, as described in Sec. 1, have been introduced by Wieczorek in [2], earlier reported in [3]. The spatial problems related to those in the present paper have been studied by Wieczorek and Wiszniewska-Matyszkiel in [5] (see also Wieczorek [4]). The games with parameter α outside the domain $\left[0, \frac{1}{2}\right]$ have not been discussed in the present paper; they are also interesting from game-theoretic point of view, but their interpretation is different and in this case some extra complications occur, which require a separate discussion. Computing the number of equilibria in "neighborhood" games with parameter $\alpha = \frac{1}{2}$ may be a nice exercise for students at an undergraduate combinatorics course.

Acknowledgements

The author is grateful to participants of the Seminar on Game and Decision Theory at the Institute of Computer Science (Warsaw) for their inspiring comments and suggestions.

The research financed from Poland's funds for science in 2005–2008, grant # 1 H02B 016 29.

References

[1] *The Fibonacci Quarterly*, **1-45**, 1963-2007

[2] A. Wieczorek, *Large Games with Only Small Players and Finite Strategy Sets*, Applicationes Mathematicae **31**, 79-96, 2004

[3] A. Wieczorek, *Elementary Large Games and an Application to Economies with Many Agents*, Report **805**, Institute of Computer Science, Polish Academy of Sciences, 1996

[4] A. Wieczorek, *Large Games with Only Small Players and Strategy Sets in Euclidean Spaces*, Applicationes Mathematicae **32**, 183-193, 2005

[5] A. Wieczorek, A. Wiszniewska (Wiszniewska-Matyszkiel), *A Game-Theoretic Model of Social Adaptation in an Infinite Population*, Applicationes Mathematicae **25**, 417-430, 1999