A discussion about Differential Games.

Adriano Festa

EEE Department, IC London.

21th January 2013

SADCO Young researchers workshop, Funchal, Madeira

ITN SADCO
Initial Training Network
Sensitivity Analysis for Deterministic Controller Design

Imperial College London

Outline of the next 2 hours

1. Meditating - chill out
   Main ideas in viscosity approach to differential games
   - a bit of History
   - Isaacs approach (verification theorems)
   - the key concept of strategies
   - link with HJ equations

2. Chatting - be active
   Open discussion
   - differential games as an abstract framework
   - limits of applicability
   - alternative ways to deal with

Let’s start from the end

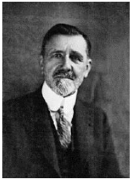
Open discussion (some suggestions)

- differential games as an abstract framework
  - Can we use the DG formulation for other kinds of applicative needs?
  - Which is the key relation between the two players?
- limits of applicability
  - How big are the HJ associated? (sometimes we are in troubles just with $> 3D$ eq)
  - A noise is a player? Does it really matter?
- alternative ways to deal with
  - Viscosity solutions v.s. piecewise regular solutions
  - Can we use some kind of adapted fast methods for optimal control problems?

A bit of History (Game theory)

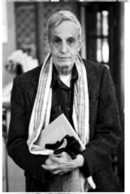
1713 In a letter, sir James Waldegrave provides a minimax mixed strategy solution to a two-person version of the card game le Her. James Madison made what we now recognize as a game-theoretic analysis of the ways states can be expected to behave under different systems of taxation.

1838 Recherches sur les principes mathematiques de la théorie des richesses Antoine Augustin Cournot considered a duopoly and presents a solution that is a restricted version of the Nash equilibrium.



1938 book *Applications aux Jeux de Hasard* and earlier notes, Émile Borel proved a minimax theorem for two-person zero-sum matrix games only when the pay-off matrix was symmetric.

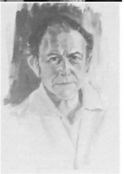
1928 John von Neumann published the (real) first paper in game theory. His paper was followed by his 1944 book *Theory of Games and Economic Behavior*. (with Oskar Morgenstern) This foundational work contains the method for finding mutually consistent solutions for two-person zero-sum games. During the following time period, work on game theory was primarily focused on cooperative game theory.



In 1950 John Nash developed a criterion for mutual consistency of players' strategies, known as Nash equilibrium.

In 1965, Reinhard Selten introduced his solution concept of subgame perfect equilibria,

In 1967, John Harsanyi developed the concepts of complete information and Bayesian games. Nash, Selten and Harsanyi became Economics Nobel Laureates in 1994. In 2005, game theorists Thomas Schelling and Robert Aumann early examples of evolutionary game theory.



Rufus Isaacs. He worked for the RAND Corporation from 1948 until 1955. His career after RAND was spent largely in the defense and avionics industries.

Isaacs, Rufus. *Differential Games*, John Wiley and Sons, 1965

What we need from Game theory

Definition (Nash equilibrium - informal)

- A set of strategies is a Nash equilibrium if no player can do better by unilaterally changing his or her strategy.
- Thus, each strategy in a Nash equilibrium is a best response to all other strategies in that equilibrium.

The Nash equilibrium may sometimes appear non-rational in a third-person perspective. This is because it may happen that a Nash equilibrium is not Pareto optimal.

	Cooperate	Defect
Cooperate	3,3	0,5
Defect	5,0	1,1

Table: Prisoner's dilemma.

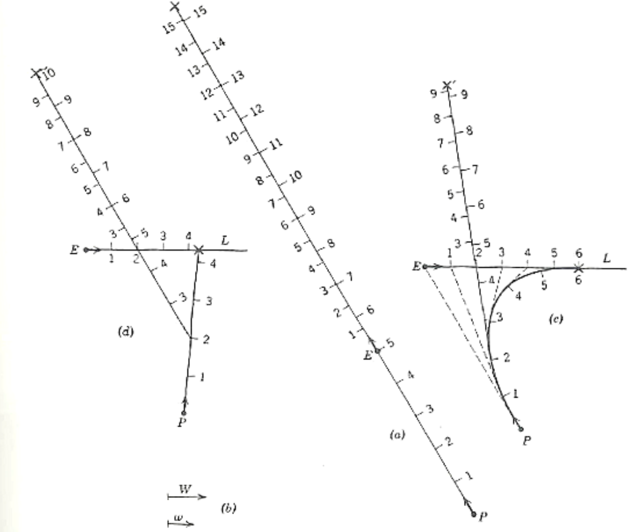
Differential games

Differential games theory: conflict problems which are driven by differential equations.

- a kind of continuous games
- finding an optimal strategy in a differential game is **closely related to the optimal control theory**
- the open-loop strategies are found using the Pontryagin Maximum Principle.
- the closed-loop strategies are found using Bellman's Dynamic Programming method.

The simplest pursuit game

Let P and E in a plane with speed W and w with $W > w$.
The payoff is the time of capture.



The general framework
two players zero-sum differential games

We are given the system

$$\begin{cases} y'(t) = f(y(t), a(t), b(t)), & t > 0 \\ y(0) = x, \end{cases} \tag{1}$$

where $y(t) \in \Omega \subset \mathbb{R}^N$ is the state, and $a \in \mathcal{A}$ and $b \in \mathcal{B}$ are the controls.

We assume

$$\begin{cases} f : \mathbb{R}^N \times A \times B \to \mathbb{R}^N & \textit{Lipschitz} \\ A, B \text{ are compact metric spaces} \end{cases} \tag{2}$$

and a cost functional

$$J(x, a(\cdot), b(\cdot)) = \int_0^{\tau_x} l(y_x(t), a(t), b(t)) e^{-\lambda t} dt \tag{3}$$

and τ_x is the exit time for a target T

Strategies I: feedback strategies

Definition

A feedback strategy for the first player (resp. second player) is a map $\tilde{a} : \mathbb{R}^+ \times \mathbb{R}^N \to A$ (resp. $\tilde{b} : \mathbb{R}^+ \times \mathbb{R}^N \to B$)

Definition

A pair $(\tilde{A}, \tilde{B})$ of sets of strategies is admissible if

1. all the Lebesgue measurable maps a (b) belong to $\tilde{A}$ ($\tilde{B}$)
2. for any pair $(\tilde{a}, \tilde{b})$ eq. (10) has a unique solution
3. (concatenation) if $a_1, a_2 \in \tilde{A}$, then $a_3 := a_1$ or a_2 at different times belongs to $\tilde{A}$
4. (shift) if $\tilde{a} \in \tilde{A}$, then for a $T > 0$, $\tilde{a}_1(t) := \tilde{a}(t + T)$ belongs to $\tilde{A}$.

Value of the game

Definition

The lower (upper) value of the game is defined by

$$v_f^-(x) := \sup_{\tilde{a}\in\tilde{A}} \inf_{\tilde{b}\in\tilde{B}} J(x,\tilde{a},\tilde{b}) \quad \left(v_f^+(x) := \inf_{\tilde{b}\in\tilde{B}} \sup_{\tilde{a}\in\tilde{A}} J(x,\tilde{a},\tilde{b})\right)$$

We define

$$H^-(x,p) = \min_{\tilde{a}\in\tilde{A}} \max_{\tilde{b}\in\tilde{B}} \{-f(x,\tilde{a},\tilde{b})\cdot p - l(x,\tilde{a},\tilde{b})\}$$

$$H^+(x,p) = \max_{\tilde{b}\in\tilde{B}} \min_{\tilde{a}\in\tilde{A}} \{-f(x,\tilde{a},\tilde{b})\cdot p - l(x,\tilde{a},\tilde{b})\}$$

Definition

The Isaacs conditions are verified when
$H(x,p) = H^-(x,p) = H^+(x,p)$.

Isaacs' Verification Theorem

Theorem

Let us assume that the target is closed and that Isaacs' condition holds. Suppose that there is a nonnegative map $v : \mathbb{R}^N$ of class C^1 over $\mathbb{R}^N \setminus T$ with $v = 0$ on T and satisfying the Isaacs' equation

$$\lambda v(x) + H(x, Dv(x)) = 0 \quad \forall x \in \mathbb{R}^N \setminus T$$

then the game has a value and v is the value of the game.

striking result! :) the resolution of a game is to solve a PDE
poor applicability.. :(the value function should be smooth enough.

we need a different notion of strategies...

Strategies II: Non anticipating strategies

The idea under the definition is that the players play in continuous time and observe each other continuously. (It's not an easy task and, actually, no completely satisfactory definition has found..)

Definition

A non anticipating strategy for the first player is a map $\alpha : \mathcal{B} \to \mathcal{A}$ if , for any $t > 0$ and $b, \bar{b} \in \mathcal{B}$, $b(s) = \bar{b}(s)$ for all $s \leq t$ implies $\alpha[b](s) = \alpha[b](s)$ for all $s \leq t$.
We denote Γ the set of the non anticipating strategies for the first player, and Δ the one for the second player (defined in the same way)

Value of a game II

Definition

The lower (upper) value of the game is defined by

$$v^-(x) := \inf_{\beta \in \Delta} \sup_{a \in \mathcal{A}} J(x, a, \beta[a]) \quad \left(v^+(x) := \sup_{\alpha \in \Gamma} \inf_{b \in \mathcal{B}} J(x, a, \beta[a]) \right)$$

we can prove the following:

$$v_f^- \leq v^- \leq v_f^+ \quad \textit{and} \quad v_f^- \leq v^+ \leq v_f^+$$

because of the more restrictive notion of non anticipating strategy.

Theorem (Dynamic programming principle)

Assumed l, f regular enough, we have for all $x \in \mathbb{R}^N$ and $t > 0$

$$v^- = \inf_{\beta \in \Delta} \sup_{a \in \mathcal{A}} \left\{ \int_0^t l(y_x(s), a(s), \beta[a](s)) e^{-\lambda s} ds + v^-(y_x(t; a, \beta[a])) e^{-\lambda t} \right\}, \quad (4)$$

$$v^+ = \sup_{a \in \Gamma} \inf_{b \in \mathcal{B}} \left\{ \int_0^t l(y_x(s), \alpha[b](s), b(s)) e^{-\lambda s} ds + v^+(y_x(t; \alpha[b], b)) e^{-\lambda t} \right\}, \quad (5)$$

Verification theorem II

Using the DPP as in the Optimal Control case we find out

Theorem

Let assume the target close ad l, f regular enough. Then the lower value v^- is a viscosity solution of

$$\lambda v^- + H^-(x, Dv^-) = 0 \quad in\ \mathbb{R}^N \setminus T$$

and ***the upper value v^+ is a viscosity solution*** *of*

$$\lambda v^+ + H^+(x, Dv^+) = 0 \quad in\ \mathbb{R}^N \setminus T$$

A 'comparison' result

Theorem

Assumed e the target close ad l, f regular enough. Then

$$v^-(x) \leq v^+(x) \quad \forall x \in \mathbb{R}^N$$

if in addition,

$$H^-(x,p) = H^+(x,p) \quad \forall x \in \mathbb{R}^N$$

then $v(x) = v^-(x) = v^+(x)$ *is the unique value of the game for all initial points.*

Why the value function?

Solving the HJ equation (in this case considering verified the Isaacs condition), and therefore getting the value function of the game, we can get the optimal behavior for every player from the starting point x_0 as

$$a(t) = S(y_{x_0}(t))$$

$$S(z) \in \operatorname{argmax}_{a\in A} \min_{b\in B}\{-f(x,a,b)\cdot Dv(x)\} \quad (6)$$

$$b(t) = W(y_{x_0}(t))$$

$$W(z) \in \operatorname{argmin}_{b\in B} \max_{a\in A}\{-f(x,a,b)\cdot Dv(x)\}. \quad (7)$$

Optimal control and optimal strategy

We consider the case of the lower value

Definition

A control $a^* \in \mathcal{A}$ is optimal at x if

$$v^-(x) = \inf_{\beta \in \Delta} J(x, a^*, \beta[a^*]).$$

Definition

A **strategy** $\beta^* \in \Delta$ **is optimal** at x if

$$v^-(x) = \sup_{a \in \mathcal{A}} J(x, a, \beta^*[a]).$$

Note that

$$J(x, a, \beta^*[a]) \leq v^-(x) = J(x, a^*, \beta^*[a^*]) \leq J(x, a^*, \beta[a^*])$$

Come back to the beginning..

Theorem

Assumed standard Hyp. then there are equivalent:

- $\beta^* \in \Delta$ *is optimal at* x_0
- *there exists* $u \in BUC(\mathbb{R}^N)$ *subsolution of HJI eq. such that*

$$\sup_{a \in \mathcal{A}} J(x_0, a, \beta^*[a]) \leq u(x_0)$$

as well as the following statements

- $a^* \in \mathcal{A}$ *is optimal at* x_0
- *there exists* $v \in BUC(\mathbb{R}^N)$ *supersolution of HJI eq. such that*

$$v(x_0) \leq \inf_{\beta \in \Delta} J(x_0, a^*, \beta[a^*])$$

Thus, the pair (a^*, β^*) is a Nash equilibrium.

PE games

The dynamic system is modeled as

$$\begin{cases} y'(t) = f(y, a(t), b(t)) \\ y(0) = x \end{cases} \tag{8}$$

where $y(t) \in R^N$ is the state, and a and b are the controls. We also have a target $\mathcal{T} := B(0, \rho)$, And a cost functional

$$J(x, a, b) := \int_0^{\tau_x} dt = \min\{t | y_x(t) \in \mathcal{T}\}$$

Thus, the HJ equation associated (after a Kruzkov transform) is

$$v(x) + \min_{b \in B} \max_{a \in A} \{-f(x, a, b) \cdot Dv(x)\} = 1 \quad \text{in } \mathbb{R} \setminus \mathcal{T}$$

For the simplest PE game in a plane

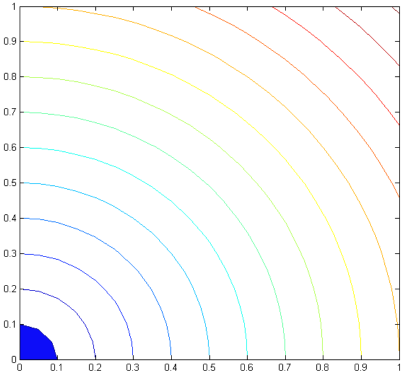
With the evader slower ($v = 0.8$) than the pursuer ($v = 1$), $A = B = B(0, 1)$

$$\begin{cases} y_1'(t) = 0.8 * a_1(t) - b_1(t)) \\ y_2'(t) = 0.8 * a_2(t) - b_2(t) y(0) = x \end{cases} \tag{9}$$

and so

$$v(x) + \min_{b \in B} \max_{a \in A} \{-(0.8 * a - b) \cdot Dv(x)\} = 1 \quad \text{in } \mathbb{R}^2 \setminus \mathcal{T}$$

PE in a plane



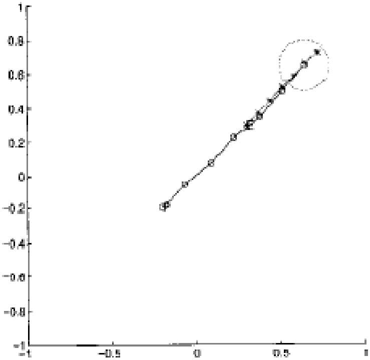


Figure: Value function and one optimal trajectory

Homicidal chauffeur

It is a more complicated Pursuit Evasion game

$$\begin{cases} y_1'(t) = v_e \sin a(t) - v_p \sin\theta(t) \\ y_2'(t) = v_e \cos a(t) - v_p \cos\theta(t) \\ \theta'(t) = \omega b(t) \\ y_1(0) = x_1 \\ y_2(0) = x_2 \\ \theta(0) = 0 \end{cases} \tag{10}$$

We remark

- In this case the problem is 3D (also in this reduced system)
- very simply the HJ equation become hard to solve

HC results

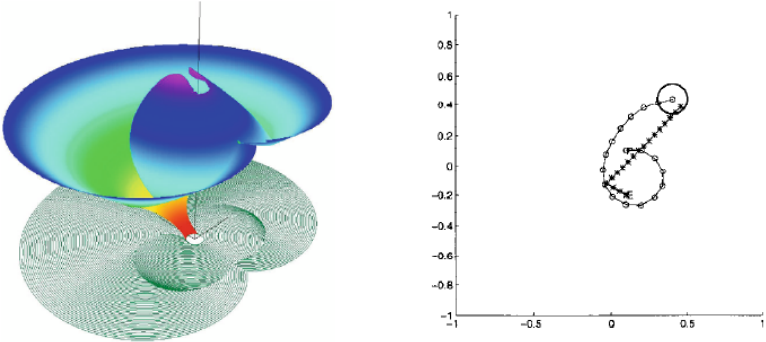


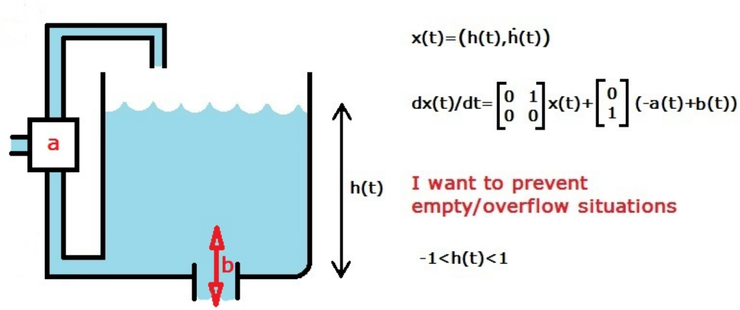
Figure: Value function and one optimal trajectory

That's all at the moment.

Open discussion (some suggestions)

- differential games as an abstract framework
  - Can we use the DG formulation for other kinds of applicative needs?
  - Which is the key relation between the two players?
- limits of applicability
  - How big are the HJ associated? (sometimes we are in troubles just with $> 3D$ eq)
  - A noise is a player? Does it really matter?
- alternative ways to deal with
  - Viscosity solutions v.s. piecewise regular solutions
  - Can we use some kind of adapted fast methods for optimal control problems?

The Surge Tank Control Problem



Falugi, Kountouriotis, Vinter. *Differential Games Controllers That Confine a System to a Safe Region in the State Space, With Applications to Surge Tank Control. IEEE Trans. Automat. Contr. 57(11): 2778-2788 (2012)*

Dupuis and McEneaney, Risk-sensitive and robust escape criteria, SIAM J. Control Optim., vol. 35, pp. 2021-2049, (1997).

The Surge Tank Control Problem

$$\frac{dx(t)}{dt} = f(x(t), a(t)) + \sigma(x(t))b(t)$$
$$= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} (-a(t) + b(t)) \; .$$

The constraints that the surge tank must neither overflow or empty are expressible (in normalized units) as

$$-1 \; < \; x_1(t) \; < \; +1 \; .$$

so $x \in \Omega := (-1, 1) \times \mathbb{R}$.
Permitted tolerances on the Max Rate of Change of Outflow (MROC) index are captured by the additional constraint on the outflow:

$$-1 \; \leq \; a(t) \; \leq \; +1 \; .$$

Surge Tank as a Differential Game

$$\begin{aligned} \mathcal{A} &:= \{a(.) : [0,\infty) \to \mathbb{R} \mid a(t) \in [-1,1]\} . \\ \mathcal{B} &:= \{b(.) : [0,\infty) \to \mathbb{R}\} . \end{aligned}$$

The space Φ of closed loop controls for the *a* player is

$$\Phi := \{\text{non-anticipative mappings } \phi(.) : \mathcal{B} \to \mathcal{A}\} .$$

The Differential game is: find

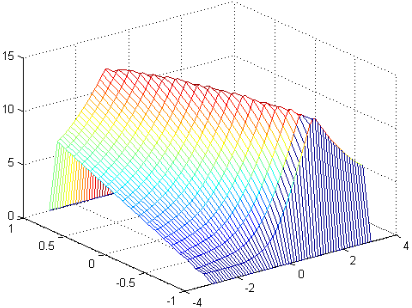
$$v(x) = \sup_{\phi\in\Phi} \inf_{b\in\mathcal{B}} J\left(x, \phi(b(.)), a(.)\right)$$

where the payoff function is

$$J(x,a,b) := \int_0^{\tau_x} \left(\frac{1}{2}|b(t)|^2 + \theta\right) dt.$$

with $\theta \geq 0$ (design parameter) and τ_x first exit time from Ω.

Surge tank



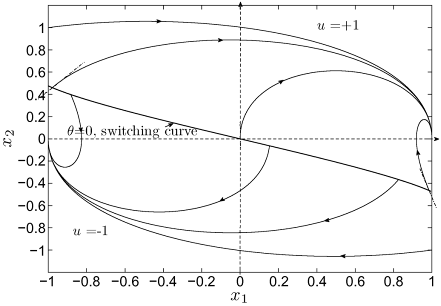


Figure: Value function and trajectories.

Tag-Chase: example 1

(case1 $c = 0.95$)

Tag-Chase: example 2

(case2 $c = 0.95$)

Tag-Chase: example 3

(case3 $c = 0.95$)

Other proposed concepts of strategies

Definition (Non anticipating feedback strategies)

A non anticipating feedback strategy for the first player is a map $\tilde{a} : \mathbb{R}^+ \times \mathbb{R}^N \to A$ which is playable (exixts a solution of the dynamic system) and non aticipating (for any $t > 0$ if $y_1 \equiv y_2$ on $[0, t]$, then $a[y_1] \equiv a[y_2]$ on $[0, t]$)

Definition (Delay strategies)

A non anticipating strategy with delay forfirst Player is a map $\alpha : \mathcal{B} \to \mathcal{A}$ for which there is a delay $\tau > 0$ such that, for any two controls $b_1, b_2 \in \mathcal{B}$ and for any $t \geq 0$, if $b_1 \equiv b_2$ on $[0, t]$, then $\alpha[b_1] \equiv \alpha[b_2]$ on $[0, t + \tau]$