Strategic Support of Cooperative Solutions in 2-Person Differential Games with Dependent Motions

Leon Petrosyan and Sergey Chistyakov

St. Petersburg University,
Faculty of Applied Mathematics and Control Processes,
35 Universitetsky prospekt, St. Petersburg, 198504, Russia
E-mail: spbuoasis7@peterlink.ru

Abstract The problem of strategically supported cooperation in 2-person differential games with integral payoffs is considered. Based on initial differential game the new associated differential game (CD-game) is designed. In addition to the initial game it models the players actions connected with transition from the strategic form of the game to cooperative with in advance chosen principle of optimality. The model provides possibility of refusal from cooperation at any time instant t for each player. As cooperative principle of optimality the Shapley value is considered. In the bases of CD-game construction lies the so-called imputation distribution procedure described earlier in (Petrosjan and Zenkevich, 2009). The theorem established by authors says that if at each instant of time along the conditionally optimal (cooperative) trajectory the future payments to each player according to the imputation distribution procedure exceed the maximal guaranteed value which this player can achieve in CD-game, then there exist a Nash equilibrium in the class of recursive strategies first introduced in (Chistyakov, 1981) supporting the cooperative trajectory. In the present paper the results similar to (Chistyakov and Petrosyan, 2011) are obtained without the requirement of independent motions and for the more general type of payoff functions.

Keywords: strong Nash equilibrium, time-consistency, core, cooperative trajectory.

1. Introduction

Similar to (Petrosjan and Zenkevich, 2009; Chistyakov and Petrosyan, 2011) in this paper the problem of strategically support of cooperation in differential 2-person game with prescribed duration T and dependent motions is considered.

$$\frac{dx}{dt} = f(t, x, u^{(1)}, u^{(2)}),\ i \in I = [1, 2], \tag{1}$$

$$x \in R^n, u^{(i)} \in P^{(i)} \subset CompR^{k(i)}, \quad i \in I$$

$$x(t_0) = x_0. \tag{2}$$

The payoffs of players $i \in I = [1, 2]$ have integral form

$$H^{(i)}_{t_0,x_0}(u^{(1)}(\cdot), u^{(2)}(\cdot)) = \int_{t_0}^{T} h^{(i)}\left(t, x(t), u^{(1)}(t), u^{(2)}(t)\right) dt, \tag{3}$$

where $u(\cdot) = (u^{(1)}(\cdot), u^{(2)}(\cdot))$ is a given vector-function of open loop controls, $x(t) = x(t, t_0, x_0, u^{(1)}(\cdot), u^{(2)}(\cdot))$ is the solution of the Cauchy problem (1) with corresponding initial conditions (2) and admissible open loop controls $u^{(1)}(\cdot), u^{(2)}(\cdot)$ of players.

Admissible open loop controls of players $i \in I$ are Lebesgue measurable open loop controls

$$u^{(i)}(\cdot) : t \mapsto u^{(i)}(t) \in R^{k(i)}, \quad i \in I = \{1, 2\}$$

such that

$$u^{(i)}(t) \in P^{(i)} \text{ for almost all } t \in [t_0, T], i \in I.$$

It is supposed that the function $f : R \times R^n \times P^{(1)} \times P^{(2)} \to R^n$ is continuous, locally Lipschitz with respect to x and satisfies the following condition: $\exists \lambda > 0$ such, that

$$||f(t, x, u^{(1)}, u^{(2)})|| \leq \lambda(1 + ||x||) \quad \forall x \in R^{k(i)}, \quad \forall u^{(1)} \in P^{(1)}, u^{(2)} \in P^{(2)}.$$

Each of the functions

$$h^{(i)} : R \times R^n \times P^{(1)} \times P^{(2)} \to R, \quad i \in I$$

are also continuous.

For all $t \in R^+$, $x \in R^n$, $\ell \in R^n$

$$\max_{u^{(1)} \in P^{(1)}} \min_{u^{(2)} \in P^{(2)}} \left(< \ell, f(t, x, u^{(1)}, u^{(2)}) > + h^{(1)}(t, x, u^{(1)}, u^{(2)})\right) = \min_{u^{(2)} \in P^{(2)}} \max_{u^{(1)} \in P^{(1)}} \left(< \ell, f(t, x, u^{(1)}, u^{(2)}) > + h^{(1)}(t, x, u^{(1)}, u^{(2)})\right)$$

and

$$\max_{u^{(2)} \in P^{(2)}} \min_{u^{(1)} \in P^{(1)}} \left(< \ell, f(t, x, u^{(1)}, u^{(2)}) > + h^{(2)}(t, x, u^{(1)}, u^{(2)})\right) = \min_{u^{(1)} \in P^{(1)}} \max_{u^{(2)} \in P^{(2)}} \left(< \ell, f(t, x, u^{(1)}, u^{(2)}) > + h^{(2)}(t, x, u^{(1)}, u^{(2)})\right),$$

here $< \cdot, \cdot >$ is scalar product in R^n.

It is supposed that at each time instant $t \in [t_0, T]$ the players have information about the current position $(t, x(t))$ on the time interval $[t_0, t]$ and use recursive strategies (Chistyakov, 1977; Chistyakov, 1999).

2. Recursive strategies

Recursive strategies were first introduced in (Chistyakov, 1977) for justification of dynamic programming approach in zero sum differential games, known as method of open loop iterations in non regular differential games with non smooth value function. The ε-optimal strategies constructed with the use of this method are universal in the sense that they remain ε-optimal in any subgame of the previously defined differential game (for every $\varepsilon > 0$). Exploiting this property it became possible to prove the existence of ε-equilibrium (Nash equilibrium) in non zero sum differential games (for every $\varepsilon > 0$) using the so called "punishment strategies" (Chistyakov, 1981).

The basic idea is that when one of the players deviates from the conditionally optimal trajectory other players after some small time delay start to play against the deviating player. As result the deviating player is not able to get much more than he could get using the conditionally optimal trajectory. The punishment of the deviating player at each time instant using one and the same strategy is possible because of the universal character of ε-optimal strategies in zero sum differential games.

In this paper the same approach is used to testify the stability of cooperative agreements in the game $\Gamma(t_0, x_0)$ and as in mentioned case the principal argument is the universal character of ε-optimal recursive strategies in specially defined zero sum games $\Gamma_i(t_0, x_0), i \in I = [1, 2]$ associated with the non-zero sum game $\Gamma(t_0, x_0)$.

The recursive strategies lie somewhere in-between piecewise open loop strategies (Petrosyan, 1993) and ε-strategies introduced by B. N. Pshenichny (Pschenichny, 1973). The difference from piecewise open loop strategies consists in the fact that like in the case of ε-strategies of B. N. Pshenichny the moments of correction of open loop controls are not prescribed from the beginning of the game but are defined during the game process. In the same time they differ from ε-strategies of B. N. Pshenichny by the fact that the formation of open loop controls happens in finite number of steps.

Recursive strategies $U_i^{(n)}$ of player i with maximal number of control corrections n is a procedure for the admissible open loop formation by player i in the game $\Gamma(t_0, x_0), (t_0, x_0) \in D$.

At the beginning of the game $\Gamma(t_0, x_0)$ player i using the recursive strategy $U_i^{(n)}$ defines the first correction instant $t_1^{(i)} \in (t_0, T]$ and his admissible open loop control $u^{(i)} = u^{(i)}(t)$ on the time interval $[t_0, t_1^{(i)}]$. Then if $t_1^{(i)} < T$ having the information about state of the game at time instant $t_1^{(i)}$ he chooses the next moment of correction $t_2^{(i)}$ and his admissible open loop control $u^{(i)} = u^{(i)}(t)$ on the time interval $(t_1^{(i)}, t_2^{(i)}]$ and so on. Then whether on k-th step $(k \leq n-1)$ the admissible control will be formed on the time interval $[t_k, T]$ or on the step n player i will end up with the process by choosing at time instant $t_{n-1}^{(i)}$ his admissible control on the remaining time interval $(t_{n-1}^{(i)}, T]$.

3. Associated games and corresponding solutions

For each given state $(t_*, x_*) \in D$ and $i \in I = [1, 2]$ consider zero sum differential game $\Gamma_i(t_*, x_*)$ between player i and $I \backslash \{i\}$ with the same dynamics as in $\Gamma(t_*, x_*)$ and payoff of player i equal to:

$$H_{t_* x_*}^{(i)}(u^{(1)}(\cdot), u^{(2)}(\cdot)) = \int_{t_0}^{T} h^{(i)}(t, x(t), u^{(1)}(t), u^{(2)}(t))dt.$$

The game $\Gamma_i(t_*, x_*)$, $i \in I$, $(t_*, x_*) \in D$, as $\Gamma(t_*, x_*), (t_*, x_*) \in D$ we consider in the class of recursive strategies. Under the above formulated conditions each of the games $\Gamma_i(t_*, x_*)$, $i \in I$, $(t_*, x_*) \in D$ has a value

$$val\Gamma_i(t_*, x_*),$$

and optimal strategies (saddle point).

Consider also the following optimization problem $\Gamma_I(t_*, x_*)$:

$$\max_{u^{(1)}(\cdot), u^{(2)}(\cdot)} \sum_{i=1}^{2} H_{t_0, x_0}^{(i)} \left(u^{(1)}(\cdot), u^{(2)}(\cdot) \right),$$

denoting the resulting maximal value as $v_I(t_0, x_0)$. We suppose that this optimization problem has an optimal open-loop solution.

The corresponding trajectory — solution of (1), (2) on the time interval $[t_0, T]$ we denote by $x_0(\cdot)$ and call "*conditionally optimal cooperative trajectory*". This trajectory may not be necessary unique. Thus on the set D the mapping

$$v(\cdot) : D \to R^3$$

is defined with coordinate functions

$$v_I(\cdot), v_1(\cdot), v_2(\cdot) : D \to R,$$

$$v_i(t_*, x_*) = val\Gamma_i(t_*, x_*), i \in I, v_I(t_*, x_*).$$

This mapping correspond to each state $(t_*, x_*) \in D$ a characteristic function $v(t_*, x_*) : 2^I \to R$ of non zero-sum game $\Gamma(t_*, x_*)$ and thus 2-person classical cooperative game $(I, v(t_*, x_*))$.

Let $E(t_*, x_*) = \{\alpha = (\alpha_1, \alpha_2) : \alpha_i \geq v_i(t_*, x_*), \alpha_1 + \alpha_2 = v_I(t_*, x_*)\}$ be the set of all imputations in the game $(I, v(t_*, x_*))$. Multivalue mapping

$$M : (t_*, x_*) \mapsto M(t_*, x_*) \subset E(t_*, x_*) \subset R^2,$$

$$M(t_*, x_*) \neq \Lambda \quad \forall (t_*, x_*) \in D,$$

is called "*optimality principle*" (defined over the family of games $\Gamma(t_*, x_*)$, $(t_*, x_*) \in D$) and the set $M(t_*, x_*)$ "*cooperative solution of the game $\Gamma(t_*, x_*)$ corresponding to this principle*".

As it follows from (Fridman, 1971) under the above imposed conditions the following Lemma holds.

Lemma 1. *The functions $v_I(\cdot), v_1(\cdot), v_2(\cdot) : D \to R$, are locally Lipschitz.*

Since the solution of the Cauchy problem (1), (2) in the sense of Caratheodory is absolutely continuous, from Lemma 1 it follows.

Theorem 1. *For every solution of the Cauchy problem (1), (2) in the sense of Caratheodory $x(\cdot)$ corresponding to the open loop controls $u(\cdot) = (u^{(1)}(\cdot), u^{(2)}(\cdot))$ functions*

$$\varphi_i : [t_0, T] \to R, \quad i \in I, \quad \varphi_i(t) = v_i(t, x(t)), \varphi_I(t) = v_I(t, x(t))$$

are absolutely continuous functions on the time interval $[t_0, T]$.

As defined let $E(t_*, x_*)$ be the set of imputations in the game $\Gamma(t_*, x_*)$, and let

$$\xi(t_*, x_*) = \{\xi_1(t_*, x_*), \xi_2(t_*, x_*)\} \in E(t_*, x_*).$$

Then we have

$$\xi_i(t_*, x_*) \geq v_i(t_*, x_*).$$

4. Realization of cooperative solutions

The realization of the solution of the game $\Gamma(t_0, x_0)$ we shall connect with the known "*imputation distribution procedure*" (IDP) (Petrosjan and Danilov, 1979; Petrosjan, 1995).

Under IDP of the imputation $\xi(t_0, x_0)$ from the solution $M(t_0, x_0)$ of the game $\Gamma(t_0, x_0)$ along conditionally optimal trajectory $x_0(\cdot)$ we understand such function

$$\beta(t) = (\beta_1(t), \beta_2(t)), \quad t \in [t_0, T], \tag{4}$$

that

$$\xi(t_0, x_0) = \int_{t_0}^{T} \beta(t) dt \tag{5}$$

and

$$\int_{t}^{T} \beta(t) dt \in E(t, x_0(t)) \quad \forall t \in [t_0, T] \tag{6}$$

where $E(t, x_0(t))$ is the set of imputations in the game $(I, v(t, x_0(t)))$.

The IDP $\beta(t), \quad t \in [t_0, T]$ of the imputation $\xi(t_0, x_0) \in M(t_0, x_0)$ of the game $\Gamma(t_0, x_0)$ is called *dynamically stable (time-consistent)* along the conditionally optimal trajectory $x_0(\cdot)$ if

$$\int_{t}^{T} \beta(t) dt \in M(t, x_0(t)) \quad \forall t \in [t_0, T] \tag{7}$$

The solution $M(t_0, x_0)$ of the game $\Gamma(t_0, x_0)$ is *dynamically stable (time-consistent)* if for all $\xi(t_0, x_0) \in M(t_0, x_0)$ along at least one conditionally optimal trajectory the dynamically stable IDP exist.

If $M(t, x_0(t)) = E(t, x_0(t))$, $t \in [t_0, T]$, then $M(t, x_0(t)) \neq \emptyset$ ($M(t, x_0(t))$ is the set of imputations in the subgame $\Gamma(t, x_0(t))$ with initial conditions on conditionally optimal cooperative trajectory with duration $T - t$), and $\xi(t, x_0(t)) \in M(t, x_0(t))$ can be selected as absolutely continuous function of t. Then the following theorem holds.

Theorem 2. *For any conditionally optimal trajectory $x_0(\cdot)$ the following IDP of the solution $\xi(t_0, x_0) \in M(t_0, x_0)$ of the game $\Gamma(t_0, x_0)$*

$$\beta(t) = -\frac{d}{dt}\xi(t, x_0(t)), \quad t \in [t_0, T], \tag{8}$$

is the dynamically stable IDP along this trajectory. Therefore the solution $M(t_0, x_0)$ of the game $\Gamma(t_0, x_0)$ is dynamically stable.

As $\xi_i(t_0, x_0)$ we can take the Shapley value:

$$\xi_i(t_0, x_0) = Sh_i(t_0, x_0) = v_i(t_0, x_0) + \frac{v_I(t_0, x_0) - \sum_{i=1}^{2} v_i(t_0, x_0)}{2}$$

and for subgame along cooperative trajectory

$$\xi_i(t, x_0(t)) = Sh_i(t, x_0(t)) = v_i(t, x_0(t)) + \frac{v_I(t, x_0(t)) - \sum_{i=1}^{2} v_i(t, x_0(t))}{2}.$$

From Theorem 1 it follows that the function $Sh_i(t, x_0(t))$ is absolutely continuous and thus differentiable along $x_0(t)$. This shows that IDP $\beta(t)$ for $\xi_i(t, x_0(t)) = Sh_i(t, x_0(t))$ can be computed by (8) according to Theorem 2.

5. About the strategically support of the imputation $\xi(t_0, x_0)$

If in the game the cooperative agreement is reached and each player gets his payoff according to the IDP (8), then it is natural to suppose that those who violate this agreement are to be punished. The effectiveness of the punishment (sanctions) comes to question of the existence of Nash Equilibrium in the following differential game $\Gamma^{\xi}(t_0, x_0)$ which differs from $\Gamma(t_0, x_0)$ only by payoffs of players.

The payoff of player i in $\Gamma^{\xi}(t_0, x_0)$ is equal to

$$H^{(\xi,\, i)}_{t_0, x_0}(u(\cdot)) = -\int_{t_0}^{t(u(\cdot))} \frac{d}{dt}\xi_i(t, x_0(t))dt + \int_{t(u(\cdot))}^{T} h^{(i)}(t, x(t, t_0, x_0, u(\cdot)))dt$$

where $t(u(\cdot))$ is the last time instant $t \in [t_0, T]$ for which

$$x_0(\tau) = x(\tau, t_0, x, u(\cdot)) \quad \forall \tau \in [t_0, t].$$

Theorem 3. *In the game $\Gamma^{\xi}(t_0, x_0)$ for each $\varepsilon > 0$ there exist ε-Nash equilibrium with outcomes (payoffs) of players in this equilibrium equal to*

$$\xi(t_0, x_0) = \{\xi_1(t_0, x_0), \xi_2(t_0, x_0)\} \in E(t_0, x_0).$$

The idea of the proof is following. Since $\xi(t_0, x_0)$ belongs to the imputation set of the game $\Gamma(t_0, x_0)$ we have

$$\xi_i(t, x_0(t)) \geq v_i(t, x_0(t)) \quad \forall i \in I \quad \forall t \in [t_0, T] \tag{9}$$

This means that at each time instant $t \in [t_0, T]$ moving along conditionally optimal trajectory $x_0(\cdot)$ no player $i \in I$ can guarantee himself the payoff $[t, T]$ more than according to IDP (8), i.e. more than

$$\int_t^T \beta(\tau)d\tau = -\int_t^T \frac{d}{dt}\xi(\tau, x_0(\tau))d\tau = \xi_i(t, x_0(t))$$

since if player i deviates from cooperative trajectory at some time instant t, this will be immediately seen by his opponent $3 - i$ (since both players know $x(t)$ at each time instant t, and deviation of one player will cause the change of $x(t)$) and he will use punishment strategy in the zero-sum game $\Gamma_{3-i}(t, x_0(t))$ (his optimal strategy in zero-sum game $\Gamma_{3-i}(t, x_0(t))$). Therefore, the player i will get no more then $v_i(t + \delta, x_0(t + \delta)) \leq \xi_i(t, x_0(t) + \varepsilon$.

In the same time on the time interval $[t_0, t]$ according to the IDP she already got the payoff equal to

$$\int_{t_0}^t \beta_i(\tau)d\tau = -\int_{t_0}^t \frac{d}{dt}\xi_i(\tau, x_0(\tau))d\tau = \xi_i(t_0, x_0) - \xi_i(t, x_0(t))$$

Consequently no player can guarantee in the game $\Gamma^{\xi}(t_0, x_0)$ the payoff more than $\xi_i(t_0, x_0)$.

According to the cooperative solution $x_0(\cdot)$ but moving always in the game $\Gamma^{\xi}(t_0, x_0)$ along conditionally optimal trajectory each player will get his payoff according to the imputation $\xi(t_0, x_0)$. Thus no player can benefit from the deviation from the conditionally optimal trajectory which in this case is natural to call "equilibrium trajectory".

References

Petrosjan, L. A., Zenkevich, N. A. (2009). *Principles of stable cooperation*. The mathematical games theory and its applications, **1**, 1, 102–117 (in Russian).

Chistyakov, S. V. (1977). *To the solution of game problem of pursuit*. Prikl. Math. i Mech., **41**, 5, 825–832, (in Russian).

Chistyakov, S. V. (1999). *Operatory znacheniya antagonisticheskikx igr (Value Operators in Two-Person Zero-Sum Differential Games)*. St. Petersburg: St. Petersburg Univ. Press.

Chistyakov, S. V., Petrosyan, L. A. (2011). *Strong Strategic Support of Cooperative solutions in Differential Games*. Contributions to Game Theory and Management, Vol. 4, pp. 105–111.

Chentsov, A. G. (1976). *On a game problem of convering at a given instant time*. Math. USSR Sbornic, **28**, 3, 353–376.

Chistyakov, S. V. (1981). *O beskoalizionnikx differenzial'nikx igrakx (On Coalition-Free Differential Games)*. Dokl. Akad. Nauk, **259(5)**, 1052–1055; English transl. in Soviet Math. Dokl. 24, 1981, no. 1, pp. 166–169.

Petrosjan, L. A. (1993). *Differential Games of Pursuit*. World Scientific, Singapore.

Pschenichny, B. N. (1973). *E-strategies in differential games*. Topics in Differential Games. New York, London, Amsterdam, pp. 45–56

Fridman, A. (1971). *Differential Games*. John Wiley and Sons, New York, NY.

Petrosjan, L. A., Danilov, N. N. (1979). *Stability of Solutions in nonzero-sum Differential Games with Integral Payoffs*. Viestnik Leningrad University, N1, pp. 52–59.

Petrosjan, L. A. (1995). *The Shapley Value for Differential Games*. Annals of the International Society of Dynamic Games, Vol.3, Geert Jan Olsder Editor, Birkhauser, pp. 409–417.