Strongly Time-Consistent Solutions in Cooperative Dynamic Games

Leon A. Petrosyan

Abstract In the paper the evolution of dynamic game along the cooperative trajectory is investigated. Along cooperative trajectory at each time instant players find themselves in a new game which is a subgame of the originally defined game. In many cases the optimal solution of the initial game restricted to the subgame along cooperative trajectory fails to be optimal in the subgame. To overcome this difficulty we introduced (see Petrosyan and Danilov, Vestnik Leningrad Univ Mat Mekh Astronom 1:52–59, 1979; Petrosyan and Zaccour, J Econ Control 27(3):381–398, 2003; Yeung and Petrosyan, Subgame consistent economic optimization. Birkhauser, 2012) the special payment mechanism—imputation distribution procedure (IDP), or payment distribution procedure (PDP), but another serious question arises: under what conditions the initial optimal solution converted to any optimal solution in the subgame will remain optimal in the whole game. This condition we call strongly time-consistency condition of the optimal solution. If this condition is not satisfied players in reality may switch in some time instant from the previously selected optimal solution to any optimal solution in the subgame, and as result realize the solution which will be not optimal in the whole game. We propose different types of strongly time-consistent solutions for multicriterial control, cooperative differential, and cooperative dynamic games.

Keywords Differential game · Time consistency · Dynamic stability · Pareto optimality · Cooperation

1 What Is Strongly Time-Consistency?

What is strongly time-consistency? Try to explain this notion. Let $M \in R^n$ be a fixed point in R^n. Consider a classical control problem (with one player)

L. A. Petrosyan (✉)
St. Petersburg State University, Saint Petersburg, Russia
e-mail: l.petrosyan@spbu.ru

23

$$\dot{x} = f(x, u), x \in R^n, u \in U \subset Comp R^l$$
$$x(t_0) = x_0, \ t \in [t_0, T].$$

(1)

Find the control $\bar{u}(t)$, and corresponding trajectory $\bar{x}(t)$ such that at terminal instant the distance $\rho(\bar{x}(T), M)$ will be minimal.

Denote this problem by $\Gamma(x_0, T - t_0)$. And denote by $C(x_0, T - t_0)$ the reachability set of system (1) from initial point x_0 at terminal time T.

Suppose for simplicity that $M \notin C(x_0, T - t_0)$. The solution of this optimal control problem we can see on Fig. 1.

Consider the intermediate time instant $\tau \in [t_0, T]$, and the intermediate control problem $\Gamma(\bar{x}(\tau), T - \tau)$ with initial condition on the optimal trajectory with duration $T - \tau$. It is clear that the control $\bar{u}(t), t \in [\tau, T]$ will be optimal also in $\Gamma(\bar{x}(\tau), T - \tau)$, so will be also the trajectory $\bar{x}(t), t \in [\tau, T]$.

This is Bellman-optimality principle and also time-consistency of optimal control $\bar{u}(t), t \in [t_0, T]$. Suppose now that we have another optimal control $\bar{\bar{u}}(t), t \in [\tau, T]$ in the problem $\Gamma(\bar{x}(\tau), T - \tau)$. Then it is easy to see that the control

$$\hat{u}(t) = \begin{cases} \bar{u}(t), \ t \in [t_0, \tau] \\ \bar{\bar{u}}(t), \ t \in [\tau, T] \end{cases}$$

will be also optimal in the problem $\Gamma(x_0, T - t_0)$. In other words: "any optimal continuation of the original problem in the subproblem along optimal trajectory generates optimal solution of the original problem." This property we shall call *strongly time-consistency (strongly dynamic stability)* (see Fig. 1).

Fig. 1 Classical optimal control problem
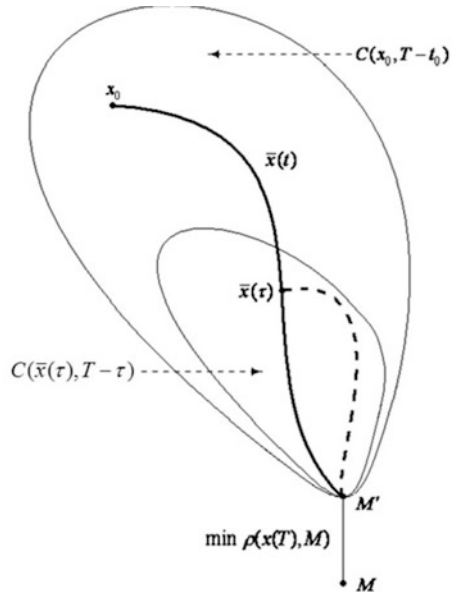
Consider now a slightly more complicated problem. The motion equations are the same (1), but the aim of control is different, it is necessary to come as close as possible to system of points $M_1, \ldots, M_k$, $M_i \in R^n$, $i \in \{1, \ldots k\}$.

Denote as before the problem by $\Gamma(x_0, T - t_0)$ and by $C(x_0, T - t_0)$ the reachability set of (1) and suppose that $C(x_0, T - t_0) \cap \hat{M} = \emptyset$, where $\hat{M}$ is the convex hull of points $\{M_1, \ldots, M_k\}$. As optimal solution here we may consider Pareto-optimal set which coincides with arc AB, the projection (suppose that $C(x_0, T - t_0)$ is convex) of $\hat{M}$ on $C(x_0, T - t_0)$ (see Fig. 2).

Consider Pareto-optimal control $\bar{u}(t)$, $t \in [t_0, T]$ which connects the initial point $x_0 \in C(x_0, T - t_0)$ with the point M belonging to the Pareto-optimal set (M belongs to the arc AB which is projection of the set $\hat{M}$ on $C(x_0, T - t_0)$). And let $\bar{x}(t)$, $t \in [t_0, T]$ be the corresponding Pareto-optimal trajectory.

Consider a subproblem $\Gamma(\bar{x}(t), T - t)$ from initial position $\bar{x}(t)$ on the Pareto-optimal trajectory. We see that the Pareto-optimal set in $\Gamma(\bar{x}(t), T - t)$ (arc $A'B'$) is different from the Pareto-optimal set in $\Gamma(x_0, T - t_0)$ having only (in our example) one common point M. This means that the control $\bar{u}(t)$, $t \in [\tau, T]$ is Pareto-optimal in subproblem $\Gamma(\bar{x}(\tau), T - \tau)$, and the Pareto-optimal solution $\bar{u}(t)$, $t \in [t_0, T]$ is time-consistent (dynamic stable) [4, 5].

In the same time we can see that the control of the type

$$\hat{u}(t) = \begin{cases} \bar{u}(t), \ t \in [t_0, \tau] \\ \bar{\bar{u}}(t), \ t \in [\tau, T], \end{cases}$$
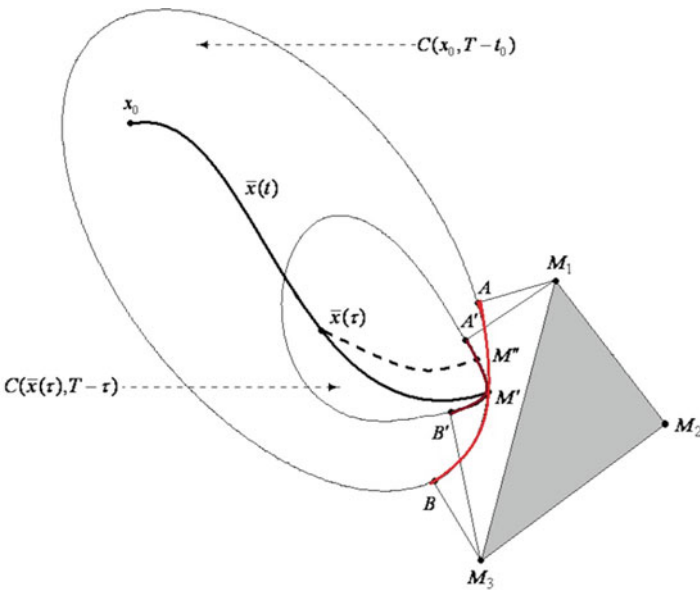

Fig. 2 Multicriterial optimal control problem

where $\bar{\bar{u}}(t)$ is an arbitrary Pareto-optimal control in subproblem $\Gamma(\bar{x}(\tau), T - \tau)$, may not be Pareto-optimal in $\Gamma(x_0, T - t_0)$.

Which means that in this case the optimal continuation of the motion in the subproblem with initial conditions on Pareto-optimal trajectory together with initial Pareto-optimal motion maybe not Pareto-optimal in the original problem. This means that the Pareto-optimal solution is time-consistent but not strongly time-consistent (see Fig. 2).

In this special problem there is one approach for constructing strongly time-consistent solutions on the bases of Pareto-optimal solutions. The idea of this approach is to consider all possible outcomes which may occur if at each time instant t on the time interval $[t_k, t_k + \delta)$ the control $u(\tau)$ will be selected leading to one of Pareto-optimal points in the subproblem $\Gamma(x(t_k), T - t_k)$. Let $t_0 < t_1 < \ldots < t_k < t_{k+1} < \ldots < t_n = T$ be the decomposition of the time interval $[t_0, T]$, $t_{k+1} - t_k = \delta > 0$. The resulting trajectory will be not Pareto-optimal, but we shall call it conditionally Pareto-optimal. Denote by $P(x(t_k), t_k)$ the set of end-points of these trajectories for all possible controls selected in a described manner. It is clear that

$$P(x(t_0), t_0) \supset P(x(t_1), t_1) \supset \ldots \supset P(x(t_k), t_k) \supset \ldots \supset P(x(T), T).$$

And the set $P(x(t_0), t_0)$ is δ-strongly time-consistent if we allow possible changes of controls only in points $t_k, k = 0, \ldots, n$.
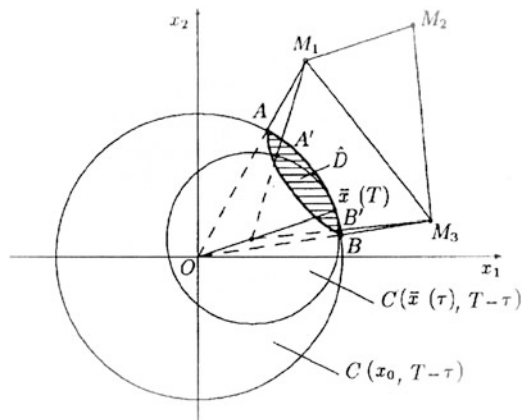
For the system

$$\dot{x} = u_1 + u_2 + u_3, x(t_0) = x_0$$

$$|u_i| \leq 1, x \in R^2, t \in [t_0, T],$$

the set $P(x(t_0), t_0)$ is denoted by $\hat{D}$ on the Fig. 3 (dashed region).

Fig. 3 Example of strongly time-consistent solution

1.1 Cooperative Differential Game

Consider now cooperative differential games with player set N. Motion equations have the form

$$\dot{x} = f(x, u_1, \ldots, u_k), x \in R^n, u_i \in U_i \subset Comp R^l \tag{2}$$

$$x(t_0) = x_0 \tag{3}$$

and the payoffs of players are defined as

$$K_i(x_0, T - t_0; u_1, \ldots, u_k) = \int_{t_0}^{T} h_i(x(t))dt, h_i > 0, i \in N.$$

Denote this game by $\Gamma(x_0, T - t_0)$. Cooperative trajectory $\overline{x}(t)$, $\overline{x}(t_0) = x_0$, $t \in [t_0, T]$ is defined as

$$\max_{u_1,\ldots,u_k} \sum_{i=1}^{n} K_i(x_0, T - t_0; u_1, \ldots, u_n) = \sum_{i=1}^{n} K_i(x_0, T - t_0; \overline{u}_1, \ldots, \overline{u}_n) =$$

$$= \sum_{i=1}^{n} \int_{t_0}^{T} h_i(\overline{x}(t))dt = v(x_0, T - t_0; N). \tag{4}$$

We suppose that max in (4) is attained. Let $v(x_0, T - t_0; S)$, $S \subset N$ be the characteristic function defined in classical way as value of zero-sum game between coalition S as first player and $N \backslash S$ as second (see [6]), and $E(x_0, T - t_0)$, the set of imputations

$$E(x_0, T - t_0) = \{\xi = \{\xi_i\} : \sum_{i=1}^{n} \xi_i = v(x_0, T - t_0; N),$$

$$\xi_i \geq v(x_0, T - t_0; \{i\}), i \in N\}. \tag{5}$$

Denote by $C(x_0, T - t_0)$ reachability set of the system (1), for $y \in C(x_0, t - t_0)$, $t \in [t_0, T]$ define a subgame $\Gamma(y, T - t)$ of $\Gamma(x_0, T - t_0)$ with characteristic function $v(y, T - t; S)$, $S \subset N$ and imputation set $E(y, T - t)$.

Optimality principle (solution) is a subset of imputation set

$$C(y, T - t) \subset E(y, T - t)$$

(Core, NM-solution,...).

Consider the family of subgames along the cooperative trajectory $\Gamma(\bar{x}(t), T - t; S)$ and also imputation set $E(\bar{x}(t), T - t)$ and the solution of subgames along this cooperative trajectory, $C(\bar{x}(t), T - t)$.

For each $\xi \in C(x_0, T - t_0)$ define the imputation distribution procedure IDP [3] $\beta(t) = (\beta_1(t), \ldots, \beta_i(t), \ldots, \beta_n(t))$

$$\xi = \int_{t_0}^{T} \beta(\tau) d\tau, \ \xi \in C(x_0, T - t_0).$$

The imputation $\xi \in C(x_0, T - t_0)$ is called dynamic stable [3–5] (time-consistent) if

$$\xi - \int_{t_0}^{t} \beta(\tau) d\tau \in C(\bar{x}(t), T - t), \ t \in [t_0, t].$$

Definition 1 The solution $C(x_0, T - t_0)$ is called time-consistent if all imputations $\xi \in C(x_0, T - t_0)$ are time-consistent.

Definition 2 Optimality principle $C(x_0, T - t_0)$ is called strongly dynamic stable [11] (strongly time-consistent) if for each $\xi \in C(x_0, T - t_0)$ there exist IDP $\beta(\tau)$ such that

$$\int_{t_0}^{t} \beta(\tau) d\tau \oplus C(\bar{x}(t), T - t) \subset C(x_0, T - t_0),$$

here $a \oplus B (a \in R^n, B \subset R^n)$ is defined as $\{a + b : b \in B\}$.

Since as it is well known time-consistency of cooperative solutions taken from the classical one-shot game theory takes place only in special cases it is clear that strongly time-consistency is a very special event. Note that strongly time-consistency has sense only for multivalued (set-valued) optimality principles (core, NM-solution).

1.2 Transformation of Characteristic Function

Let $v(y, T - t; S)$ be characteristic function in $\Gamma(y, T - t)$. Define the following integral transformation

$$\bar{v}(x_0, T - t_0; S) = \int_{t_0}^{T} \frac{v(\bar{x}(t), T - t; S) \sum_{i \in N} h_i(\bar{x}(t))}{v(\bar{x}(t), T - t; N)} dt,$$

here $v(\bar{x}(t), T - t; S)$ is characteristic function computed for subgame $\Gamma(\bar{x}(t), T - t)$ along cooperative trajectory. It can be seen that

$$\overline{v}(x_0, T - t; N) = v(x_0, T - t; N).$$

Define the imputation set $\overline{E}(x_0, T - t_0)$ and the core under the new characteristic function $\overline{v}(x_0, T - t_0; S)$, $\overline{C}(x_0, T - t_0) \subset \overline{E}(x_0, T - t_0)$ and define the integral transformation of the imputation $\xi \in E(x_0, T - t_0)$ to $\overline{\xi} \in \overline{E}(x_0, T - t_0)$ as

$$\overline{\xi}_i = \int_{t_0}^{T} \frac{\xi_i(t) \sum\limits_{i \in N} h_i(\bar{x}(t))}{V(\bar{x}(t), T - t; N)} dt, \; i \in N,$$

where $\xi(t) \in E(\bar{x}(t), T - t)$. Similarly let $\overline{E}(\bar{x}(t), T - t)$ $\overline{C}(\bar{x}(t), T - t)$ be the set of imputations and the core in subgame $\Gamma(\bar{x}(t), T - t)$ along cooperative trajectory under characteristic function

$$\overline{v}(\bar{x}(t), T - t; S) = \int_{t}^{T} \frac{v(\overline{x}(\tau), T - \tau; S) \sum\limits_{i \in N} h_i(\overline{x}(\tau))}{v(\overline{x}(\tau), T - \tau; N)} d\tau.$$

Theorem 1 $\overline{C}(x_0, T - t_0)$ *is strongly time-consistent.*

To prove it is sufficient to take for each $\overline{\xi} \in \overline{E}(x_0, T - t_0)$ as $\beta_i(t)$

$$\beta_i(t) = \frac{\xi_i(t) \sum\limits_{i \in N} h_i(\overline{x}(t))}{v(\overline{x}(t), T - t; N)},$$

where $\xi(t) \in C(\overline{x}(t), T - t)$ is an integrable selector from $C(\overline{x}(t), T - t)$.

What is the connection between $\overline{C}$ and C? If there is a nonvoid intersection of $\overline{C}$ and C, then this imputation set could be a good preferable optimality principle in $\Gamma(x_0, T - t)$. Introduce

$$\lambda(S) = \max_{t_0 \le t \le T} \frac{v(\overline{x}(t), T - t; S)}{v(\overline{x}(t), T - t; N)},$$

$$\lambda(N) = 1.$$

We have

$$\overline{v}(x_0, T - t_0; S) \le \lambda(S) \int_{t_0}^{T} \sum_{i \in N} h_i(\overline{x}(t)) dt = \lambda(S) v(x_0, T - t_0; N),$$

$$\overline{v}(x_0, T - t_0; N) = \lambda(N) v(x_0, T - t_0; N) = v(x_0, T - t_0; N),$$

$$\lambda(S) \ge \frac{v(x_0, T - t_0; S)}{v(x_0, T - t_0; N)},$$

$$v(x_0, T - t_0; S) \le \lambda(S) v(x_0, T - t_0; N).$$

Denote by $\hat{C}(x_0, T - t_0)$ the set of all solutions $\xi = \{\xi_1, \ldots, \xi_n\}$

$$\sum_{i \in S} \xi_i \geq \lambda(S)v(x_0, T - t_0; N), \ \ S \subset N, \ \sum_{i \in N} \xi_i = v(x_0, T - t_0; N).$$

From previous considerations it follows

$$\sum_{i \in S} \xi_i \geq \lambda(S)v(x_0, T - t_0; N) \geq v(x_0, T - t_0; S).$$

We see that

$$\hat{C}(x_0, T - t_0) \subset C(x_0, T - t_0) \cap \overline{C}(x_0, T - t_0)$$

and

$$\hat{C}(\overline{x}(t), T - t) \subset C(\overline{x}(t), T - t) \cap \overline{C}(\overline{x}(t), T - t).$$

The following theorem holds.

Theorem 2

$$\overline{C}(x_0, T - t_0) \supset \int_{t_0}^{t} \frac{\xi(t) \sum_{i=1}^{n} h_i(\overline{x}(t))}{v(\overline{x}(t), T - t; N)} \oplus \hat{C}(\overline{x}(t), T - t) \quad (6)$$

for any integrable selector $\xi(t) \in C(x(t), T - t)$.

Proof Theorem 2 follows from the inclusion $\hat{C}(\overline{x}(t), T - t) \subset \overline{C}(\overline{x}(t), T - t)$ and strongly time-consistency of $\overline{C}(x_0, T - t_0)$.

From Theorem 2 it follows that for each imputation $\xi_0 \in C(x_0, T - t_0) \cap \hat{C}(x_0, T - t_0)$ there exist IDP

$$\beta(t) = \frac{\xi(t) \sum_{i=1}^{n} h_i(\overline{x}(t))}{v(\overline{x}(t), T - t; N)},$$

where $\xi(t_0) = \xi_0$ and $\xi(t)$ is an integrable selector from $C(\overline{x}(t), T - t)$, such that

$$\int_{t_0}^{t} \beta(\tau)d\tau \oplus \hat{C}(\overline{x}(t), T - t) \subset \overline{C}(x_0, T - t_0). \quad (7)$$

$\square$

Suppose that $\hat{C}(x_0, T - t_0) \neq \emptyset$. The interpretation of (7) is as follows. $\hat{C}(x_0, T - t_0)$ is the subset of the original core $C(x_0, T - t_0)$ and for any imputation $\xi \in \hat{C}(x_0, T - t_0) \cap C(x_0, T - t_0)$ from this subset of original core $C(x_0, T - t_0)$ one can construct the IDP (the imputation distribution procedure) such that if in an intermediate time instant t players for some reasons would like to switch to another optimal imputation $(\xi^t)' \in \hat{C}(\bar{x}(t), T - t) \subset C(\bar{x}(t), T - t)$ from the subset of original core, they will still get the payments according to the imputation from $\bar{C}(x_0, T - t_0)$, resulting from the integral transformation of $C(x_0, T - t_0)$.

2 Repeated Games

Folk theorems are well known in game theory [1, 2, 6–9]. By using the so-called punishment strategies they show the possibility to attain in some sense preferable outcomes. These outcomes are stable against deviations of single players. But the natural question arises: is it possible to get "good" outcomes stable against deviations of coalitions (coalition-proofness). Now we try to construct a mechanism based on the introduction of an analog of characteristic function which makes it possible (under some conditions on this newly defined characteristic function) to get coalition-proofness for repeated and multistage games [9]. This will show us the way of constructing strongly time-consistent optimality principles in multistage games.

Denote by G the infinity repeated n-person game with the game Γ played on each stage. For simplicity suppose that the stage game Γ is finite (has finite sets of strategies).

$$\Gamma = < N; U_1, \dots, U_i, \dots, U_n; K_1, \dots, K_i, \dots, K_n > .$$

If on stage $k(1 \leq k \leq \infty)$ strategy profile $u^k = (u_1^k, \dots, u_i^k, \dots, u_n^k)$ is chosen, the payoff in G is defined as

$$H_i(u_1(\cdot), \dots, u_i(\cdot), \dots, u_n(\cdot)) = \sum_{k=1}^{\infty} \delta^{k-1} K_i(u_1^k, \dots, u_i^k, \dots, u_n^k) =$$

$$= \sum_{k=1}^{\infty} \delta^{k-1} K_i(u^k) = H_i(u(\cdot)), \ i \in N, \quad (8)$$

here $u_1(\cdot) = (u_1^1, \dots, u_1^k, \dots), \dots, u_i(\cdot) = (u_i^1, \dots, u_i^k, \dots), \dots, u_n(\cdot) = (u_n^1, \dots, u_n^k, \dots), \delta \in (0, 1)$.

Here in the expression $u_i(\cdot) = (u_i^1, \dots, u_i^k, \dots), i \in N$ u_i^k is the strategy chosen by player i in the game Γ on stage k. We suppose that on stage k when choosing u_i^k player i knows the choices of other players and remembers his choices on previous stages. Thus u_i^k is function of history

$$h^k = (u_1^1, \ldots, u_1^{k-1}; \ldots; u_i^1, \ldots, u_i^{k-1}; \ldots; u_n^1, \ldots, u_n^{k-1}).$$

Formally we have to write $u_i^k(h^k)$, i.e. u_i^k depends upon history h^k, $k = 1, \ldots$. However in this paper for convenience we shall write u_i^k instead $u_i^k(h^k)$.

Consider the strategy profile $\bar{u}(\cdot) = (\bar{u}_1(\cdot), \ldots, \bar{u}_i(\cdot), \ldots, \bar{u}_n(\cdot))$ such that

$$\sum_{i \in N} H_i(\bar{u}) = \max_{u(\cdot)} \sum_{i \in N} H_i(u). \tag{9}$$

It is evident that such strategy profile always exists.

One can take $\bar{u}_i(\cdot) = (\bar{u}_i^1, \ldots, \bar{u}_i^k, \ldots,) \; i \in N$ such that

$$\sum_{i \in N} K_i(\bar{u}_1, \ldots, \bar{u}_i, \ldots, \bar{u}_n) = \max_{u_1, \ldots, u_i, \ldots, u_n} \sum_{i \in N} K_i(u_1, \ldots, u_i, \ldots, u_n) \tag{10}$$

and since the stage games are the same (G is repeated game) we can take $\bar{u}_i^k = \bar{u}_i$ for all $k = 1, \ldots, n$. Then from (8)–(10) we get that

$$\sum_{i \in N} H_i(\bar{u}) = \sum_{i \in N} \left(\sum_{k=1}^{\infty} \delta^{k-1} K_i(\bar{u}_1^k, \ldots, \bar{u}_n^k) \right) =$$
$$= \sum_{i \in N} \left(\sum_{k=1}^{\infty} \delta^{k-1} K_i(\bar{u}_1, \ldots, \bar{u}_n) \right) = \frac{1}{1-\delta} \sum_{i \in N} K_i(\bar{u}_1, \ldots, \bar{u}_n). \tag{11}$$

Introduce characteristic function $V(S), \; S \subset N$ in Γ in classical sense. Then we shall have

$$V(N) = \sum_{i \in N} K_i(\bar{u}_1, \ldots, \bar{u}_n) \tag{12}$$

and it can be easily shown that the characteristic function $W(S), \; S \subset N$ in G will have the form

$$W(S) = \frac{1}{1-\delta} V(S), \; S \subset N. \tag{13}$$

Remind now the definition of strong (or coalition proof) Nash equilibrium.

Definition 3 The n-tuple of strategies $(\hat{u}_1, \ldots \hat{u}_2, \ldots \hat{u}_n) = \hat{u}$ is called strong (or coalition proof) Nash equilibrium (SNE) if for all $S \subset N$, and all $u_S = \{u_i, i \in S\}$ the following inequality holds

$$\sum_{i \in S} K_i(\hat{u}) \geq \sum_{i \in S} K_i(\hat{u}||u_S). \tag{14}$$

Consider now the core C in Γ, and suppose that $C \neq \emptyset$, and suppose also that there exist an imputation $\alpha \in C$ such that

$$\sum_{i \in S} \alpha_i > V(S), \ S \subset N, \ S \neq N. \tag{15}$$

2.1 Associated Zero-Sum Games

Consider a family of zero-sum games $\Gamma_{N\setminus i, i}$ with coalition $N\setminus\{i\}$ as first player and coalition $\{i\}$ as second. The payoff of $N\setminus\{i\}$ is equal to the sum of payoffs of players from $N\setminus\{i\}$. Denote by $V(N\setminus i)$ the value of $\Gamma_{N\setminus i, i}$. Let $(\bar{\mu}_{N\setminus i}, \bar{\mu}_i)$ be the saddle point (in mixed strategies) in $\Gamma_{N\setminus i, i}$.

Consider the n-tuple of strategies $\bar{\mu} = (\bar{\mu}_1, \ldots, \bar{\mu}_n)$, and define

$$\overline{W}(S) = \max_{\mu_S} \sum_{i \in S} K_i(\mu_S; \bar{\mu}_{N\setminus S}),$$

here $\mu_S = \{\mu_i, \ i \in S\}$, $\bar{\mu}_{N\setminus S} = \{\bar{\mu}_i, \ i \in N\setminus S\}$. It is clear that

$$\overline{W}(S) \geq V(S), \ \overline{W}(N) = V(N), \ S \subset N.$$

Suppose, that there exist the solution of the system

$$\sum_{i \in S} \alpha_i > \overline{W}(S), \ \sum_{i \in N} \alpha_i = \overline{W}(N) = V(N). \tag{16}$$

Construct now the modification G^α of the game G. The difference between G^α and G is in payoffs defined in stage games Γ when the cooperative strategies $\bar{u} = (\bar{u}_1, \ldots, \bar{u}_n)$ are used and the payoff in this case is equal to $\alpha = (\alpha_1, \ldots, \alpha_n)$, where α satisfies (16). For all other strategy combinations the payoffs remain as in Γ.

The following theorem holds [10].

Theorem 3 *In game G^α there exist $\delta \in (0, 1)$ and SNE such that payoffs in this SNE are equal to $\alpha_i \dfrac{1}{1 - \delta}$, which are payoffs in G^α under cooperation.*

2.2 Multistage Games

Multistage game G starts from a fixed stage game $\Gamma(z_1)$ which can be considered as situated in the position (root) z_1 of the game tree G.

$$\Gamma(z_1) = \langle N; U_1^{z_1}, \ldots, U_i^{z_1}, \ldots, U_n^{z_1}; K_1^{z_1}, \ldots, K_i^{z_1}, \ldots, K_n^{z_1} \rangle. \tag{17}$$

For simplicity we suppose that the set of players N is the same in all stage games. When the game G develops the infinite sequence of stage games is realized but only a finite number of them are different since we suppose that the total number of different stage game $\Gamma(z)$ is finite. As usual in multistage games we consider the general case when the next stage game depends upon controls chosen by players only in previous stage game. Like in previous section denote by $u_i(\cdot)$ the strategy of player i in G (defined as function of histories). The strategy profile which maximizes the sum of players payoffs in G is called "cooperative" strategy profile and the corresponding sequence of stage games (or equivalently sequence of positions on the tree G) "cooperative trajectory." Suppose that for each stage game $\Gamma(z)$ the characteristic function $V(z, S)$ (in classical sense) is defined.

For each stage game $\Gamma(z)$ consider the family of zero-sum games $\Gamma_{N\setminus i, i}(z)$ and corresponding saddle points $\bar{\mu}^z_{N\setminus i}$, μ^z_i, and $\bar{\mu}^z = (\bar{\mu}^z_1, \ldots, \bar{\mu}^z_n)$, define

$$\overline{W}(z, S) = \max_{\mu^z_S} \sum_{i \in S} K^z_i(\mu^z_S, \bar{\mu}^z_{N\setminus S}).$$

Let

$$\overline{W}(S) = \sup_z \overline{W}(z, S).$$

Suppose that

$$\overline{W}(S) < \inf_z W(z, N) = \inf_z V(z, N).$$

Suppose the core $C(z)$ is not empty in each stage game $\Gamma(z)$, denote by $D(z)$ the subcore of $C(z)$ as set of all imputations $\alpha^z = (\alpha^z_1, \ldots, \alpha^z_n)$, $\sum_{i \in S} \alpha^z_i \geq \overline{W}(S)$, for all S.

Suppose that for all $z \in G$, $D(z) \neq \emptyset$ and suppose also that there exist imputation $\alpha^z = (\alpha^z_1, \ldots, \alpha^z_n)$ such that

$$\sum_{i \in S} \alpha^z_i > \overline{W}(S) \text{ for all } S, \tag{18}$$

$$\inf_{S,z} \left[\sum_{i \in S} \alpha^z_i - \overline{W}(S) \right] = A > 0. \tag{19}$$

For simplicity we shall consider the special case when $V(z, N) = \overline{W}(N)$ for all z the previous conditions (18) and (19) can be written as

$$\sum_{i \in S} \alpha_i > \overline{W}(S) \text{ for all } S, \tag{20}$$

$$\inf_{S} \left[\sum_{i \in S} \alpha_i - \overline{W}(S) \right] = A > 0, \tag{21}$$

since the number of different stage games is finite and we can select α the same in all stage games.

Construct now the modification G^{α} of the game in the same way as it was done in Sect. 1. Theorem 1 from Sect. 1 holds also for the game G^{α}.

Theorem 4 *In the game G^{α} there exist $\delta \in (0, 1)$ and SNE such that payoffs in this SNE are equal to $\alpha_i \frac{1}{1-\delta}$, which are payoffs in G^{α} under cooperation.*

2.3 Time-Consistency and Strongly Time-Consistency

Consider cooperative version of game G and subgame $G(z)$. Introduce the following characteristic function in G and in $G(z)$, respectively,

$$\hat{W}(S) = \frac{1}{1-\delta} W(S).$$

Denote the analog of the core $\hat{C}$ and $\hat{C}(z)$ in G under the defined above c.f.

Strongly time-consistency in this case means that for each imputation $\bar{\alpha} \in \hat{C}(\bar{z}_0)$ there exist corresponding IDP $\bar{\beta}(1), \ldots, \bar{\beta}(l), \ldots$ such that

$$\sum_{k=0}^{l} \delta^k \bar{\beta}(k) \oplus \delta^{l+1} \hat{C}(\bar{z}_{l+1}) \subset \hat{C}(\bar{z}_0). \tag{22}$$

It can be easily seen that if $D(z) = D \neq \emptyset$, by selecting $\bar{\beta}(k) = \beta \in D(\bar{z}_k)$ we can guarantee the strongly time-consistency of $\hat{C}(\bar{z}_0)$.

Suppose $\alpha \in \hat{C}(\bar{z}_0)$, then by definition we have

$$\sum_{i \in S} \bar{\alpha}_i \geq \hat{W}(S) = \frac{1}{1-\delta} \bar{W}(S); \quad \sum_{i \in N} \bar{\alpha}_i = \hat{W}(N) = \frac{1}{1-\delta} \bar{W}(N).$$

Represent $\bar{\alpha}$ in the form

$$\bar{\alpha} = \sum_{k=0}^{\infty} \delta^k \bar{\beta},$$

since $\bar{\alpha} \in \hat{C}(\bar{z}_0)$

$$\sum_{i \in S} \bar{\alpha}_i = \sum_{i \in S} \frac{1}{1-\delta} \bar{\beta}_i \geq \hat{W}(S) = \sum_{i \in S} \frac{1}{1-\delta} \bar{W}(S),$$

and

$$\sum_{i \in S} \bar{\beta}_i \geq \bar{W}(S), \quad \sum_{i \in N} \bar{\beta}_i = \bar{W}(N).$$

Thus $\bar{\beta} \in D(\bar{z}_k) = D$, $k = 0, 1, \ldots, l, \ldots$. And we get that each imputation $\bar{\alpha} \in \hat{C}(\bar{z}_0)$ can be represented in the form $\bar{\alpha} = \sum_{k=0}^{\infty} \delta^k \bar{\beta}(k)$, when $\bar{\beta}(k) = \bar{\beta} \in D(\bar{z}_k) = D$.

This will give us also strongly time-consistency of $\hat{C}(\bar{z}_0)$.

We have seen that for arbitrary $\bar{\alpha} \in \hat{C}(\bar{z}_0)$ there exist such IDP $\bar{\beta}(0), \bar{\beta}(1), \ldots, \bar{\beta}(k), \ldots$ (in our case $\bar{\beta}(k) = \bar{\beta} \in D$), that

$$\bar{\alpha} = \sum_{k=0}^{\infty} \delta^k \bar{\beta}(k).$$

Suppose that $\alpha' \in \sum_{k=0}^{l} \delta^k \bar{\beta}(k) \oplus \delta^{l+1} \hat{C}(\bar{z}_{l+1})$. To prove (22) we have to prove that in this case $\alpha' \in \hat{C}(\bar{z}_0)$. Consider the stage l then we can write the imputation α' in the form

$$\alpha' = \sum_{k=0}^{l} \delta^k \bar{\beta}(k) + \delta^{l+1} \alpha'',$$

here $\bar{\beta}(k) = \bar{\beta} \in D$), where $\alpha'' \in \hat{C}(\bar{z}_{l+1})$.

Since $\alpha'' \in \hat{C}(\bar{z}_{l+1})$ we have

$$\sum_{i \in S} \alpha_i'' \geq \hat{W}(S) = \frac{1}{1-\delta} \bar{W}(S), \quad \sum_{i \in N} \alpha_i'' = \hat{W}(N) = \frac{1}{1-\delta} \bar{W}(N),$$

and we can show that similar to previous case when $\alpha \in \hat{C}(\bar{z}_0)$, α'' can be represented in the form

$$\alpha'' = \sum_{k=l+1}^{\infty} \delta^{k-(l+1)} \beta''(k),$$

where $\beta''(k) = \beta'' \in D$, $k = l+1, \ldots$.

Then we get

$$\alpha' = \sum_{k=0}^{l} \delta^k \bar{\beta}(k) + \delta^{l+1} \sum_{k=l+1}^{\infty} \delta^{k-(l+1)} \bar{\bar{\beta}}(k) = \sum_{k=0}^{\infty} \delta^k \tilde{\beta}(k),$$

where $\tilde{\beta}(k) \in D, \tilde{\beta}(k) = \bar{\beta}(k) = \bar{\beta}, k = 1, \ldots, l, \tilde{\beta}(k) = \bar{\bar{\beta}}(k) = \beta'', k = l+1, \ldots$.
And we have

$$\sum_{i \in S} \alpha' = \sum_{k=0}^{l} \delta^k \sum_{i \in S} \tilde{\beta}_i(k) + \sum_{k=l+1}^{\infty} \delta^k \sum_{i \in S} \tilde{\beta}_i(k) = \sum_{k=1}^{l} \delta^k \sum_{i \in S} \bar{\beta}_i(k) + \sum_{k=l+1}^{\infty} \delta^k \sum_{i \in S} \bar{\bar{\beta}}_i(k) \geq,$$

$$\geq \sum_{k=0}^{l} \delta^k \bar{W}(S) + \sum_{k=l+1}^{\infty} \delta^k \bar{W}(S) = \sum_{k=0}^{\infty} \delta^k \bar{W}(S) = \frac{1}{1-\delta} \bar{W}(S) = \hat{W}(S).$$

In the similar way we can prove that $\sum_{i \in N} \alpha'_i = \hat{W}(S)$. This proves that $\alpha' \in \hat{C}(\bar{z}_0)$.

References

1. Aumann, R.J., Maschler, M.: Repeated Games with Incomplete Information. MIT Press, Cambridge (1995)
2. Myerson, R.B.: Multistage Games with Communication. Econometrica. **54**, 323–358 (1986)
3. Petrosyan, L.A., Danilov, N.N.: Stability of the solutions in nonantagonistic differential games with transferable payoffs. Vestnik Leningrad. Univ. Mat. Mekh. Astronom. **1**, 52–59 (1979)
4. Petrosyan, L.A., Zaccour, G.: Time-consistent Shapley value allocation of pollution cost reduction. Journal of Economics and Control. **27**, 3, 381–398 (2003)
5. Yeung, D.W.K., Petrosyan, L.A.: Subgame Consistent Economic Optimization. Birkhauser (2012)
6. M. Maschler, M., Solan, E., Zamir, S.: Game Theory. Cambridge University Press (2013)
7. Aumann, R., Shapley, L.: Long-Term Competition – A Game-Theoretic Analysis. Essays in Game Theory. (1994). https://doi.org/10.1007/978-1-4612-2648-21
8. Rubinstein, A.: Equilibrium in Supergames. Essays in Game Theory. (1994). https://doi.org/10.1007/978-1-4612-2648-22
9. Fudenberg, D., Maskin, E.: The Folk Theorem in Repeated Games with Discounting or with Incomplete Information. Econometrica. **54**, 3, 533–554 (1986). https://doi.org/10.2307/1911307.JSTOR1911307
10. Petrosjan L.A., Pankratova, Y.B.: Construction of Strong Nash Equilibria in a class of infinite nonzero-sum games. Trudy Inst. Mat. Mekh. UrO RAN. **24** (2018)
11. Petrosjan L.A.: Strongly time-consistent differential optimality principles. Vestnik St. Petersburg Univ. Math. **26**, 4, 40–46 (1993)