Zero-sum games

November 28, 2018

1 Conflict and cooperation

Definition 1. *Game theory is a decision theory in conflict situations.*

Definition 2. *The game is a mathematical model in conflict situations.*

A theory of games, introduced in 1921 by Emile Borel, was established in 1928 by John von Neumann and Oskar Morgenstern, to develop it as a means of decision making in complex economic systems. In their book "The Theory of Games and Economic Behaviour", published in 1944, they asserted that the classical mathematics developed for applications in mechanics and physics fail to describe the real processes in economics and social life.

2 Definition of a two-person zero-sum game in normal form

Definition 3. *The system*

$$\Gamma = (X, Y, K), \tag{2.1}$$

where X and Y are nonempty sets, and the function $K : X \times Y \to R^1$, is called a two-person zero-sum game in normal form.

The elements $x \in X$ and $y \in Y$ are called the *strategies* of players 1 and 2, respectively, in the game Γ, the elements of the Cartesian product $X \times Y$ (i.e. the pairs of strategies (x, y), where $x \in X$ and $y \in Y$) are called *situations*, and the function K is the payoff of Player 1. Player 2's payoff in situation (x, y) is equal to $[-K(x, y)]$; therefore the function K is also called the *payoff function* of the game Γ and the game Γ is called a *zero-sum game*. Thus, in order to specify the game Γ, it is necessary to define the sets of strategies X, Y for players 1 and 2, and the payoff function K given on the set of all situations $X \times Y$.

The game Γ is interpreted as follows. Players simultaneously and independently choose strategies $x \in X, y \in Y$. Thereafter Player 1 receives the payoff equal to $K(x, y)$ and Player 2 receives the payoff equal to $(-K(x, y))$.

Definition 4. *Definition. The game* $\Gamma' = (X', Y', K')$ *is called a subgame of the game* $\Gamma = (X, Y, K)$ *if* $X' \subset X, Y' \subset Y$, *and the function* $K' : X' \times Y' \to R^1$ *is a restriction of function* K *on* $X' \times Y'$.

This section focuses on two-person zero-sum games in which the strategy sets of the players' are finite.

Definition 5. *Definition. Two-person zero-sum games in which both players have finite sets of strategies are called matrix games.*

Suppose that Player 1 in matrix game (2.1) has a total of m strategies. Let us order the strategy set X of the first player, i.e. set up a one-to-one correspondence between the sets $M = \{1, 2, \dots, m\}$ and X. Similarly, if Player 2 has n strategies, it is possible to set up a one-to-one correspondence between the sets $N = \{1, 2, \dots, n\}$ and Y. The game Γ is then fully defined by specifying the matrix $A = \{a_{ij}\}$, where $a_{ij} = K(x_i, y_i)$, $(i, j) \in M \times N$, $(x_i, y_j) \in X \times Y, i \in M, j \in N$ (whence comes the name of the game — the matrix game). In this case the game Γ is realized as follows. Player 1 chooses row $i \in M$ and Player 2 (simultaneously and independently from Player 1) chooses column $j \in N$. Thereafter Player 1 receives the payoff (a_{ij}) and Player 2 receives the payoff $(-a_{ij})$. If the payoff is equal to a negative number, then we are dealing with the actual loss of Player 1.

Denote the game Γ with the payoff matrix A by Γ_A and call it the $(m \times n)$ game according to the dimension of matrix A.

Example 1. Defense of the city. This example is known in literature as Colonel Blotto game. Colonel Blotto has m regiments and his enemy has n regiments. The enemy is defending two posts. The post will be taken by Colonel Blotto if when attacking the post he is more powerful in strength on this post. The opposing parties are two separate regiments between the two posts.

Define the payoff to the Colonel Blotto (Player 1) at each post. If Blotto has more regiments than the enemy at the post (Player 2), then his payoff at this post is equal to the number of the enemy's regiments plus one (the occupation of the post is equivalent to capturing of one regiment). If Player 2 has more regiments than Player 1 at the post, Player 1 loses his regiments at the post plus one (for the lost of the post). If each side has the same number of regiments at the post, it is a draw and each side gets zero. The total payoff to Player 1 is the sum of the payoffs at the two posts.

The game is zero-sum. We shall describe strategies of the players. Suppose that $m > n$. Player 1 has the following strategies: $x_0 = (m, 0)$ – to place all of the regiments at the first post; $x_1 = (m-1, 1)$ – to place $(m-1)$ regiments at the first post and one at the second; $x_2 = (m-2, 2), \dots, x_{m-1} = (1, m-1), x_m = (0, m)$. The enemy (Player 2) has the following strategies: $y_0 = (n, 0)$, $y_1 = (n-1, 1), \dots, y_n = (0, n)$.

Suppose that the Player 1 chooses strategy x_0 and Player 2 chooses strategy y_0. Compute the payoff a_{00} of Player 1 in this situation. Since $m > n$, Player 1 wins at the first post. His payoff is $n+1$ (one for holding the post). At the second post it is draw. Therefore $a_{00} = n+1$. Compute a_{01}. Since $m > n-1$, then in the first post Player 1's payoff is $n-1+1=n$. Player 2 wins at the second post. Therefore the loss of Player 1 at this post is one. Thus, $a_{01}=n-1$. Similarly, we obtain $a_{0j}=n-j+1-1=n-j,\ 1\le j\le n$. Further, if $m-1>n$ then $a_{10}=n+1+1=n+2,\ a_{11}=n-1+1=n,\ a_{1j}=n-j+1-1-1=n-j-1,\ 2\le j\le n$. In a general case (for any m and n) the elements $a_{ij},\ i=\overline{0,m},\ j=\overline{0,n}$, of the payoff matrix are computed as follows:

$$a_{ij}=K(x_i,y_j)=\begin{cases} n+2 & \text{if } m-i>n-j, \quad i>j,\\ n-j+1 & \text{if } m-i>n-j, \quad i=j,\\ n-j-i & \text{if } m-i>n-j, \quad i<j,\\ -m+i+j & \text{if } m-i<n-j, \quad i>j,\\ j+1 & \text{if } m-i=n-j, \quad i>j,\\ -m-2 & \text{if } m-i<n-j, \quad i<j,\\ -i-1 & \text{if } m-i=n-j, \quad i<j,\\ -m+i-1 & \text{if } m-i<n-j, \quad i=j,\\ 0 & \text{if } m-i=n-j, \quad i=j. \end{cases}$$

Thus, with $m=4, n=3$, considering all possible situations, we obtain the payoff matrix A of this game:

$$A=\begin{array}{c} \\ x_0\\ x_1\\ x_2\\ x_3\\ x_4 \end{array}\begin{array}{c} \begin{array}{cccc} y_0 & y_1 & y_2 & y_3 \end{array}\\ \begin{bmatrix} 4 & 2 & 1 & 0\\ 1 & 3 & 0 & -1\\ -2 & 2 & 2 & -2\\ -1 & 0 & 3 & 1\\ 0 & 1 & 2 & 4 \end{bmatrix} \end{array}.$$

Example 2. Game of Evasion. Players 1 and 2 choose integers i and j from the set $\{1,\dots,n\}$. Player 1 wins the amount $|i-j|$. The game is zero-sum. The payoff matrix is square $(n\times n)$ matrix, where $a_{ij}=|i-j|$. For $n=4$, the payoff matrix A has the form

$$A=\begin{array}{c} \\ 1\\ 2\\ 3\\ 4 \end{array}\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array}\\ \begin{bmatrix} 0 & 1 & 2 & 3\\ 1 & 0 & 1 & 2\\ 2 & 1 & 0 & 1\\ 3 & 2 & 1 & 0 \end{bmatrix} \end{array}.$$

Example 3. Discrete Type Game.

Players approach one another by taking n steps. After each step a player may or may not fire a bullet, but during the game he may fire only once. The probability that the player will hit his opponent (if he shoots) on the k-th step is assumed to be k/n $(k\le n)$.

A strategy for Player 1 (2) consists in taking a decision on shooting at the i-th (j-th) step. Suppose that $i<j$ and Player 1 makes a decision to shoot at the i-th step and Player 2 makes a decision to shoot at the j-th step. The payoff a_{ij} to Player 1 is then determined by

$$a_{ij}=\frac{i}{n}-\left(1-\frac{i}{n}\right)\frac{j}{n}=\frac{n(i-j)+ij}{n^2}.$$

Thus the payoff a_{ij} is the difference in the probabilities of hitting the opponent and failing to survive. In the case $i > j$, Player 2 is the first to fire and $a_{ij} = -a_{ji}$. If however, $i = j$, then we set $a_{ij} = 0$. Accordingly, if we set $n = 5$, the game matrix multiplied by 25 has the form

$$A = \begin{bmatrix} 0 & -3 & -7 & -11 & -15 \\ 3 & 0 & 1 & -2 & -5 \\ 7 & -1 & 0 & 7 & 5 \\ 11 & 2 & -7 & 0 & 15 \\ 15 & 5 & -5 & -15 & 0 \end{bmatrix}.$$

Example 4. Attack-Defense Game. Suppose that Player 1 wants to attack one of the targets $c_1, \ldots, c_n$ having positive values $\tau_1 > 0, \ldots, \tau_n > 0$. Player 2 defends one of these targets. We assume that if the undefended target c_i is attacked, it is necessarily destroyed (Player 1 wins τ_i) and the defended target is hit with probability $1 > \beta_i > 0$ (the target c_i withstands the attack with probability $1 - \beta_i > 0$), i.e. Player 1 wins (on the average) $\beta_i \tau_i$, $i = 1, 2, \ldots, n$.

The problem of choosing the target for attack (for Player 1) and the target for defense (for Player 2) reduces to the game with the payoff matrix

$$A = \begin{bmatrix} \beta_1\tau_1 & \tau_1 & \ldots & \tau_1 \\ \tau_2 & \beta_2\tau_2 & \ldots & \tau_2 \\ \ldots & \ldots & \ldots & \ldots \\ \tau_n & \tau_n & \ldots & \beta_n\tau_n \end{bmatrix}.$$

Example 5. Discrete Search Game.

There are n cells. Player 2 hide an object in one of n cells and Player 1 wishes to find it. In examining the i-th cell, Player 1 exerts $\tau_i > 0$ efforts, and the probability of finding the object in the i-th cell (if it is concealed there) is $0 < \beta_i \leq 1$, $i = 1, 2, \ldots, n$. If the object is found, Player 1 receives the amount α. The players' strategies are the numbers of cells wherein the players respectively hide and search for the object. Player 1's payoff is equal to the difference in the expected receipts and the efforts made in searching for the object. Thus, the problem of hiding and searching for the object reduces to the game with the payoff matrix

$$A = \begin{bmatrix} \alpha\beta_1 - \tau_1 & -\tau_1 & -\tau_1 & \ldots & -\tau_1 \\ -\tau_2 & \alpha\beta_2 - \tau_2 & -\tau_2 & \ldots & -\tau_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ -\tau_n & -\tau_n & -\tau_n & \ldots & \alpha\beta_n - \tau_n \end{bmatrix}.$$

Example 6. Noisy Search Suppose that Player 1 is searching for a mobile object (Player 2) for the purpose of detecting it. Player 2's objective is the opposite one (i.e. he seeks to avoid being detected). Player 1 can move with velocities $\alpha_1 = 1$, $\alpha_2 = 2$, $\alpha_3 = 3$ and Player 2 with velocities $\beta_1 = 1$, $\beta_2 = 2$, $\beta_3 = 3$, respectively. The range of the detecting device used by Player 1, depending on the velocities of the players is determined by the matrix

$$D = \begin{array}{c} \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{array} \begin{array}{c} \begin{array}{ccc} \beta_1 & \beta_2 & \beta_3 \end{array} \\ \begin{bmatrix} 4 & 5 & 6 \\ 3 & 4 & 5 \\ 1 & 2 & 3 \end{bmatrix} \end{array}.$$

Strategies of the players are the velocities, and Player 1's payoff in the situation (α_i, β_j) is assumed to be the search efficiency $a_{ij} = \alpha_i \delta_{ij}$, $i = \overline{1,3}, j = \overline{1,3}$, where δ_{ij} is an element of the matrix D. Then the problem of selecting velocities in a noisy search can be represented by the game with matrix

$$A = \begin{array}{c} \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{array} \begin{array}{c} \begin{array}{ccc} \beta_1 & \beta_2 & \beta_3 \end{array} \\ \begin{bmatrix} 4 & 5 & 6 \\ 6 & 8 & 10 \\ 3 & 6 & 9 \end{bmatrix} \end{array}.$$

3 Solution of matrix games

Consider a two-person zero-sum game $\Gamma = (X, Y, K)$. In this game each of the players seeks to maximize his payoff by choosing a proper strategy. But for Player 1 the payoff is determined by the function $K(x, y)$, and for Player 2 it is determined by $(-K(x, y))$, i.e. the players' objectives are directly opposite. Note that the payoff of Player 1 (2) (the payoff function) is determined on the set of situations $(x, y) \in X \times Y$. Each situation, and hence the player's payoff do not depend only on his own choice, but also on what strategy will be chosen by his opponent whose objective is directly opposite. Therefore, seeking to obtain the maximum possible payoff, each player must take into account the opponent's behavior.

In the theory of games it is supposed that the behavior of both players is rational, i.e. they wish to obtain the maximum payoff, assuming that the opponent is acting in the best (for himself) possible way. What maximal payoff can Player 1 guarantee himself? Suppose player 1 chooses strategy x. Then, at worst case he will win $\min_y K(x, y)$. Therefore, Player 1 can always guarantee himself the payoff $\max_x \min_y K(x, y)$. If the max and min are not reached, Player 1 can guarantee himself obtaining the payoff arbitrarily close to the quantity

$$\underline{v} = \sup_{x \in X} \inf_{y \in Y} K(x, y), \tag{3.1}$$

which is called the *lower value* of the game.

The principle of constructing strategy x based on the maximization of the minimal payoff is called the *maximin principle*, and the strategy x selected by this principle is called the *maximin strategy* of Player 1.

For Player 2 it is possible to provide similar reasonings. Suppose he chooses strategy y. Then, at worst, he will lose $\max_x K(x, y)$. Therefore, the second player can always guarantee himself the payoff $-\min_y \max_x K(x,y)$. The number

$$\overline{v} = \inf_{y \in Y} \sup_{x \in X} K(x,y) \tag{3.2}$$

is called the *upper value of the game* Γ.

The principle of constructing a strategy y, based on the minimization of maximum losses, is called the *minimax principle*, and the strategy y selected for this principle is called the *minimax strategy* of Player 2. It should be stressed that the existence of the minimax (maximin) strategy is determined by the reachability of the extremum in (3.1), (3.2).

Consider the $(m \times n)$ matrix game Γ_A. Then the extrema in (3.1) and (3.2) are reached and the lower and upper values of the game are, respectively equal to

$$\underline{v} = \max_{1 \le i \le m} \min_{1 \le j \le n} a_{ij}, \tag{3.3}$$

$$\overline{v} = \min_{1 \le j \le n} \max_{1 \le i \le m} a_{ij}. \tag{3.4}$$

The minimax and maximin for the game Γ_A can be found by the following scheme

$$\begin{array}{c} \left[\begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{array}\right] \\ \underbrace{\begin{array}{cccc} \max_i a_{i1} & \max_i a_{i2} & \dots & \max_i a_{in} \end{array}}_{\min\limits_j \max\limits_i a_{ij}} \end{array} \left.\begin{array}{c} \min_j a_{1j} \\ \min_j a_{2j} \\ \dots \\ \min_j a_{mj} \end{array}\right\} \max_i \min_j a_{ij}$$

Thus, in the game Γ_A with the matrix

$$A = \begin{bmatrix} 1 & 0 & 4 \\ 5 & 3 & 8 \\ 6 & 0 & 1 \end{bmatrix}$$

the lower value (maximin) $\underline{v}$ and the maximin strategy i_0 of the first player are $\underline{v} = 3$, $i_0 = 2$, respectively, and the upper value (minimax) $\overline{v}$ and the minimax strategy j_0 of the second player are $\overline{v} = 3$, $j_0 = 2$, respectively.

Definition 6. *In the two-person zero-sum game $\Gamma = (X, Y, K)$ the point (x^*, y^*) is called an equilibrium point, or a saddle point, if*

$$K(x, y^*) \le K(x^*, y^*), \tag{3.5}$$

$$K(x^*, y) \ge K(x^*, y^*) \tag{3.6}$$

for all $x \in X$ and $y \in Y$.

The set of all equilibrium points in the game Γ will be denoted as

$$Z(\Gamma),\ Z(\Gamma) \subset X \times Y.$$

Definition 7. *The value of the payoff function of the I-player in an equilibrium point is called the game value and is denoted by* v.

Equilibrium situations have the following properties.

Theorem 1. *Let* (x_1^*, y_1^*), (x_2^*, y_2^*) *be two arbitrary saddle points in the two-person zero-sum game* Γ. *Then:*

1. $K(x_1^*, y_1^*) = K(x_2^*, y_2^*)$;
2. $(x_1^*, y_2^*) \in Z(\Gamma)$, $(x_2^*, y_1^*) \in Z(\Gamma)$.

Proof. From the definition of a saddle point for all $x \in X$ and $y \in Y$ we have

$$K(x, y_1^*) \leq K(x_1^*, y_1^*) \leq K(x_1^*, y); \tag{3.7}$$

$$K(x, y_2^*) \leq K(x_2^*, y_2^*) \leq K(x_2^*, y). \tag{3.8}$$

We substitute x_2^* into the left-hand side of the inequality (3.7), y_2^* into the right-hand side, x_1^* into the left-hand side of the inequality (3.8) and y_1^* into the right-hand side. Then we get

$$K(x_2^*, y_1^*) \leq K(x_1^*, y_1^*) \leq K(x_1^*, y_2^*) \leq K(x_2^*, y_2^*) \leq K(x_2^*, y_1^*).$$

From this it follows that:

$$K(x_1^*, y_1^*) = K(x_2^*, y_2^*) = K(x_2^*, y_1^*) = K(x_1^*, y_2^*). \tag{3.9}$$

Show the validity of the second statement. Consider the point (x_2^*, y_1^*). From (3.7) -(3.9), we then have

$$K(x, y_1^*) \leq K(x_1^*, y_1^*) = K(x_2^*, y_1^*) = K(x_2^*, y_2^*) \leq K(x_2^*, y) \tag{3.10}$$

for all $x \in X, y \in Y$. The inclusion $(x_1^*, y_2^*) \in Z(\Gamma)$ can be proved in much the same way.

From the theorem it follows that the payoff function takes the same values at all saddle points. Therefore, it is meaningful to introduce the following definition.

For any game $\Gamma = (X, Y, K)$ the following proposition takes place.

Lemma 1. *In matrix game* Γ

$$\underline{v} \ \leq \ \overline{v} \tag{3.11}$$

or

$$\sup_{x \in X} \inf_{y \in Y} K(x, y) \ \leq \ \inf_{y \in Y} \sup_{x \in X} K(x, y). \tag{3.12}$$

Proof. Let $x \in X$ random player strategy 1. Then we have

$$K(x, y) \leq \sup_{x \in X} K(x, y).$$

Hence we get

$$\inf_{y\in Y} K(x,y) \le \inf_{y\in Y} \sup_{x\in X} K(x,y).$$

Now we note that on the right-hand side of the last inequality constant, and the value $x \in X$ was chosen arbitrarily. Therefore, inequality

$$\sup_{x\in X} \inf_{y\in Y} K(x,y) \le \inf_{y\in Y} \sup_{x\in X} K(x,y).$$

Theorem 2 (Necessary and sufficient conditions for the existence of the saddle point in the game $\Gamma = (X, Y, K)$)**.** *For the existence of the saddle point in the game $\Gamma = (X, Y, K)$, it is necessary and sufficient that the quantities*

$$\min_y \sup_x K(x,y), \;\; \max_x \inf_y K(x,y) \tag{3.13}$$

exist and the following equality holds:

$$\underline{v} = \max_x \inf_y K(x,y) = \min_y \sup_x K(x,y) = \overline{v}. \tag{3.13}$$

Proof.

Necessity. Let $(x^*, y^*) \in Z(\Gamma)$. Then for all $x \in X$ and $y \in Y$ the following inequality holds:

$$K(x, y^*) \le K(x^*, y^*) \le K(x^*, y) \tag{3.14}$$

and hence

$$\sup_x K(x, y^*) \le K(x^*, y^*). \tag{3.15}$$

At the same time, we have

$$\inf_y \sup_x K(x,y) \le \sup_x K(x, y^*). \tag{3.16}$$

Comparing (3.15) and (3.16) we get

$$\inf_y \sup_x K(x,y) \le \sup_x K(x, y^*) \le K(x^*, y^*). \tag{3.17}$$

In the similar way we get the inequality

$$K(x^*, y^*) \le \inf_y K(x^*, y) \le \sup_x \inf_y K(x,y). \tag{3.18}$$

On the other hand, the inverse inequality (3.12) holds. Thus, we get

$$\sup_x \inf_y K(x,y) = \inf_y \sup_x K(x,y), \tag{3.19}$$

and finally we get

$$\min_y \sup_x K(x,y) = \sup_x K(x, y^*) = K(x^*, y^*),$$

$$\max_x \inf_y K(x,y) = \inf_y K(x^*,y) = K(x^*,y^*),$$

i.e. the exterior extrema of the min sup and max inf are reached at the points y^* and x^* respectively.

Sufficiency. Suppose there exist the min sup and max inf

$$\max_x \inf_y K(x,y) = \inf_y K(x^*,y); \tag{3.20}$$

$$\min_y \sup_x K(x,y) = \sup_x K(x,y^*) \tag{3.21}$$

and the equality (3.13) holds. We shall show that (x^*, y^*) is a saddle point. Indeed,

$$K(x^*,y^*) \geq \inf_y K(x^*,y) = \max_x \inf_y K(x,y); \tag{3.22}$$

$$K(x^*,y^*) \leq \sup_x K(x,y^*) = \min_y \sup_x K(x,y). \tag{3.23}$$

By (3.13) the min sup is equal to the max inf, and from (3.22), (3.23) it follows that the min sup is also equal to the $K(x^*,y^*)$, i.e. the inequalities in (3.22), (3.23) are satisfied as equalities. Now we have

$$K(x^*,y^*) = \inf_y K(x^*,y) \leq K(x^*,y),$$

$$K(x^*,y^*) = \sup_x K(x,y^*) \geq K(x,y^*)$$

for all $x \in X$ and $y \in Y$, i.e. $(x^*,y^*) \in Z(\Gamma)$.

The proof shows that the common value of the min sup and max inf is equal to $K(x^*,y^*) = v$, the value of the game, and any min sup (max inf) strategy $y^*(x^*)$ is optimal in terms of the theorem, i.e. the point (x^*,y^*) is a saddle point.

The proof of the theorem yields the following statement.

Corollary 1. *If the* min sup *and* max inf *in (3.12) exist and are reached on* $\overline{y}$ *and* $\overline{x}$*, respectively, then*

$$\max_x \inf_y K(x,y) = K(\overline{x},\overline{y}) = \min_y \sup_x K(x,y). \tag{3.24}$$

The games, in which saddle points exist, are called *strictly determined*. Therefore, this theorem establishes the criterion for strict determination of the game and can be restated as follows. For the game to be strictly determined it is necessary and sufficient that the min sup and max inf in (3.12) exist and the equality (3.13) is satisfied.

Note that, in the game Γ_A, the extrema in (3.12) are always reached and the theorem may be reformulated in the following form.

Corollary 2. *For the* $(m \times n)$ *matrix game to be strictly determined it is necessary and sufficient that the following equalities hold*

$$\min_{j=1,2,\ldots,n} \max_{i=1,2,\ldots,m} \alpha_{ij} = \max_{i=1,2,\ldots,m} \min_{j=1,2,\ldots,n} \alpha_{ij}. \tag{3.25}$$

For example, in the game with the matrix $\begin{bmatrix} 1 & 4 & 1 \\ 2 & 3 & 4 \\ 0 & -2 & 7 \end{bmatrix}$ the situation (2,1) is a saddle point. In this case

$$\max_i \min_j a_{ij} = \min_j \max_i a_{ij} = 2.$$

On the other hand, the game with the matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ does not have a saddle point, since

$$\min_j \max_i a_{ij} = 1 > \max_i \min_j a_{ij} = 0.$$

4 Mixed extension of a game

Definition 8. *The random variable whose values are strategies of a player is called a mixed strategy of the player.*

Thus, for the matrix game Γ_A, a mixed strategy of Player 1 is a random variable whose values are the row numbers $i \in M$, $M = \{1, 2, \ldots, m\}$. A similar definition applies to Player 2's mixed strategy whose values are the column numbers $j \in N$ of the matrix A.

Considering the above definition of *mixed strategies*, the former strategies will be referred to as *pure strategies*. Since the random variable is characterized by its distribution, the mixed strategy will be identified in what follows with the probability distribution over the set of pure strategies. Thus, Player 1's mixed strategy x in the game is the m-dimensional vector

$$x = (\xi_1, \ldots, \xi_m), \ \sum_{i=1}^{m} \xi_i = 1, \xi_i \geq 0, i = 1, \ldots, m. \tag{4.1}$$

Similarly, Player 2's mixed strategy y is the n-dimensional vector

$$y = (\eta_1, \ldots, \eta_n), \ \sum_{j=1}^{n} \eta_j = 1, \eta_j \geq 0, j = 1, \ldots, n. \tag{4.2}$$

In this case, $\xi_i \geq 0$ and $\eta_j \geq 0$ are the probabilities of choosing the pure strategies $i \in M$ and $j \in N$, respectively, when the players use mixed strategies x and y.

For example, the net strategy 1 for the first player – $x = (1, 0, \ldots, 0)$.

Definition 9. *The pair (x, y) of mixed strategies in the matrix game Γ_A is called the situation in mixed strategies.*

We shall define the payoff of Player 1 at the point (x, y) in mixed strategies for the $(m \times n)$ matrix game Γ_A as the mathematical expectation of his payoff provided that the players use mixed strategies x and y, respectively. The players choose their strategies independently; therefore the mathematical expectation of payoff $K(x, y)$ in mixed strategies $x = (\xi_1, \dots, \xi_m)$, $y = (\eta_1, \dots, \eta_n)$ is equal to

$$K(x, y) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \xi_i \eta_j = (xA)y = x(Ay) \tag{4.3}$$

The function $K(x, y)$ is continuous in $x \in X$ and $y \in Y$. Notice that when one player uses a pure strategy (i or j, respectively) and the other uses a mixed strategy (y or x), the payoffs $K(i, y), K(x, j)$ are computed by formulas

$$K(i, y) = K(u_i, y) = \sum_{j=1}^{n} a_{ij} \eta_j = a_i y, i = 1, \dots, m,$$

$$K(x, j) = K(x, w_j) = \sum_{i=1}^{m} a_{ij} \xi_i = x a^j, j = 1, \dots, n,$$

where a_i, a^j are respectively the ith row and the jth column of the $(m \times n)$ matrix A.

Thus, from the matrix game $\Gamma_A = (M, N, A)$ we have arrived at a new game $\overline{\Gamma}_A = (X, Y, K)$, where X and Y are the sets of mixed strategies in the game Γ_A and K is the payoff function in mixed strategies (mathematical expectation of the payoff). The game $\overline{\Gamma}_A$ will be called a *mixed extension* of the game Γ_A. The game Γ_A is a subgame for $\overline{\Gamma}_A$, i.e. $\Gamma_A \subset \overline{\Gamma}_A$.

Definition 10. *The point (x^*, y^*) in the game $\overline{\Gamma}_A$ forms a saddle point and the number $v = K(x^*, y^*)$ is the value of the game $\overline{\Gamma}_A$ if for all $x \in X$ and $y \in Y$*

$$K(x, y^*) \le K(x^*, y^*) \le K(x^*, y). \tag{4.4}$$

The strategies (x^*, y^*) appearing in the saddle point are called optimal. Moreover, by Theorem 2, the strategies x^* and y^* are respectively the maximin and minimax strategies, since the exterior extrema in (3.12) are reachable (the function $K(x, y)$ is continuous on the compact sets X and Y).

Lemma 2. *Let Γ_A and $\Gamma_{A'}$ be two $(m \times n)$ matrix games, where*

$$A' = \alpha A + B, \;\; \alpha > 0, \alpha = const,$$

and B is the matrix with the same elements β, i.e. $\beta_{ij} = \beta$ for all i and j. Then $Z(\overline{\Gamma}_{A'}) = Z(\overline{\Gamma}_A)$, $\overline{v}_{A'} = \alpha \overline{v}_A + \beta$, where $\overline{\Gamma}_{A'}$ and $\overline{\Gamma}_A$ are the mixed extensions of the games $\Gamma_{A'}$ and Γ_A, respectively, and $\overline{v}_{A'}, \overline{v}_A$ are the values of the games $\overline{\Gamma}_{A'}$ and $\overline{\Gamma}_A$.

Proof. Both matrices A and A' are of dimension $m \times n$; therefore the sets of mixed strategies in the games $\Gamma_{A'}$ and Γ_A coincide. We shall show that for any situation in mixed strategies (x, y) the following equality holds

$$K'(x, y) = \alpha K(x, y) + \beta, \tag{4.5}$$

where K' and K are Player 1's payoffs in the games $\overline{\Gamma}_{A'}$ and $\overline{\Gamma}_A$, respectively.

Indeed, for all $x \in X$ and $y \in Y$ we have

$$K'(x,y) = xA'y = \alpha(xAy) + xBy = \alpha K(x,y) + \beta.$$

From Scale Lemma it then follows that $Z(\overline{\Gamma}_{A'}) = Z(\overline{\Gamma}_A)$, $\overline{v}_{A'} = \alpha \overline{v}_A + \beta$.

Example 7. Verify that the strategies $y^* = (\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$, $x^* = (\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$ are optimal and $v = 0$ is the value of the game $\overline{\Gamma}_A$ with matrix

$$A = \begin{bmatrix} 1 & -1 & -1 \\ -1 & -1 & 3 \\ -1 & 3 & -1 \end{bmatrix}.$$

We shall simplify the matrix A (to obtain the maximum number of zeros). Adding a unity to all elements of the matrix A, we get the matrix

$$A' = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 0 \end{bmatrix}.$$

Each element of the matrix A' can be divided by 2. The new matrix is of the form

$$A'' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 2 & 0 \end{bmatrix}.$$

By the lemma we have $v_{A''} = \frac{1}{2} v_{A'} = \frac{1}{2}(v_A + 1)$. Verify that the value of the game Γ_A is equal to 1/2. Indeed, $K(x^*, y^*) = Ay^* = 1/2$. On the other hand, for each strategy $y \in Y, y = (\eta_1, \eta_2, \eta_3)$ we have $K(x^*, y) = \frac{1}{2}\eta_1 + \frac{1}{2}\eta_2 + \frac{1}{2}\eta_3 = \frac{1}{2} \cdot 1 = \frac{1}{2}$, and for all $x = (\xi_1, \xi_2, \xi_3), x \in X$, $K(x, y^*) = \frac{1}{2}\xi_1 + \frac{1}{2}\xi_2 + \frac{1}{2}\xi_3 = \frac{1}{2}$. Consequently, the above-mentioned strategies x^*, y^* are optimal and $v_A = 0$.

In what follows, whenever the matrix game Γ_A is mentioned, we shall mean its mixed extension $\overline{\Gamma}_A$.

Theorem 3. *The main theorem of matrix games. Any matrix game has a saddle point in mixed strategies. [von Neumann, 1928].*

Proof. Let Γ_A be an arbitrary $(m \times n)$ game with a strictly positive matrix $A = \{a_{ij}\}$, i.e. $a_{ij} > 0$ for all $i = \overline{1,m}$ and $j = \overline{1,n}$. Show that in this case the theorem is true. To do this, we shall consider an auxiliary linear programming problem

$$\min xu, \ xA \geq w, \ x \geq 0 \tag{4.6}$$

and its dual problem

$$\max yw, \ Ay \leq u, \ y \geq 0, \tag{4.7}$$

where $u = (1, \ldots, 1) \in R^m$, $w = (1, \ldots, 1) \in R^n$. From the strict positivity of the matrix A it follows that there exists a vector $x > 0$ for which $xA > w$, i.e. problem (4.6) has a feasible solution. On the other hand, the vector $y = 0$ is a feasible solution of problem (4.7). And it can by easily seen that there exist a feasible solution of (4.7) y' with $|y'| > 0$. Therefore, by the duality theorem of linear programming, both problems (4.6) and (4.7) have optimal solutions $\overline{x}, \overline{y}$, respectively, and

$$\overline{x}u = \overline{y}w = \Theta > 0. \tag{4.8}$$

Consider vectors $x^* = \overline{x}/\Theta$ and $y^* = \overline{y}/\Theta$ and show that they are optimal strategies for the players 1 and 2 in the game $\overline{\Gamma}_A$, respectively and the value of the game is equal to $1/\Theta$.

Indeed, from (4.8) we have

$$x^*u = (\overline{x}u)/\Theta = (\overline{y}w)/\Theta = y^*w = 1,$$

and from feasibility of $\overline{x}$ and $\overline{y}$ for problems (4.6), (4.7), it follows that $x^* = \overline{x}/\Theta \geq 0$ and $y^* = \overline{y}/\Theta \geq 0$, i.e. x^* and y^* are the mixed strategies of players 1 and 2 in the game Γ_A.

Let us compute a payoff to Player 1 at (x^*, y^*):

$$K(x^*, y^*) = x^*Ay^* = (\overline{x}A\overline{y})/\Theta^2. \tag{4.9}$$

On the other hand, from the feasibility of vectors $\overline{x}$ and $\overline{y}$ for problems (4.6),(4.7) and equality (4.8), we have

$$\Theta = w\overline{y} \leq (\overline{x}A)\overline{y} = \overline{x}(A\overline{y}) \leq \overline{x}u = \Theta. \tag{4.10}$$

Thus, $\overline{x}A\overline{y} = \Theta$ and (4.9) implies that

$$K(x^*, y^*) = 1/\Theta. \tag{4.11}$$

Let $x \in X$ and $y \in Y$ be arbitrary mixed strategies for players 1 and 2. The following inequalities hold:

$$K(x^*, y) = (x^*A)y = (\overline{x}A)y/\Theta \geq (wy)/\Theta = 1/\Theta, \tag{4.12}$$

$$K(x, y^*) = x(Ay^*) = x(A\overline{y})/\Theta \leq \ (xu)/\Theta = 1/\Theta. \tag{4.13}$$

Comparing (4.11)–(4.13), we have that (x^*, y^*) is a saddle point and $1/\Theta$ is the value of the game Γ_A with a strictly positive matrix A.

Now consider the $(m \times n)$ game $\Gamma_{A'}$ with an arbitrary matrix $A' = \{a'_{ij}\}$. Then there exists such constant $\beta > 0$ that the matrix $A = A' + B$ is strictly positive, where $B = \{\beta_{ij}\}$ is an $(m \times n)$ matrix, $\beta_{ij} = \beta$, $i = \overline{1, m}$, $j = \overline{1, n}$. In the game Γ_A there exists a saddle point (x^*, y^*) in mixed strategies, and the value of the game equals $v_A = 1/\Theta$, where Θ is determined as in (4.8).

From Lemma 2, it follows that $(x^*, y^*) \in Z(\overline{\Gamma}_{A'})$ is a saddle point in the game $\Gamma_{A'}$ in mixed strategies and the value of the game is equal to $v_{A'} = v_A - \beta = 1/\Theta - \beta$. This completes the proof of Theorem.

Informally, the existence of a solution in the class of mixed strategies implies that, by randomizing the set of pure strategies, the players can always eliminate uncertainty in choosing their strategies they have encountered before the game starts. Note that the mixed strategy solution does not necessarily exist in zero-sum games. Examples of such games with an infinite number of strategies are given in Secs. 2.3, 2.4.

Notice that the proof of theorem is constructive, since the solution of the matrix game is reduced to a linear programming problem, and the solution algorithm for the game $\Gamma_{A'}$ is as follows.

1. By employing the matrix A', construct a strictly positive matrix $A = A' + B$, where $B = \{\beta_{ij}\}, \beta_{ij} = \beta > 0$.
2. Solve the linear programming problems (4.6), (4.7). Find vectors $\overline{x}, \overline{y}$ and a number Θ (see (4.8)).
3. Construct optimal strategies for the players 1 and 2, respectively,
$$x^* = \overline{x}/\Theta,\ y^* = \overline{y}/\Theta.$$
4. Compute the value of the game $\Gamma_{A'}$
$$v_{A'} = 1/\Theta - \beta.$$

Example 9. Consider the matrix game Γ_A determined by the matrix

$$A = \begin{bmatrix} 4 & 0 \\ 2 & 3 \end{bmatrix}.$$

Associated problems of linear programming are of the form

$$\begin{array}{ll} \min \xi_1 + \xi_2, & \max \eta_1 + \eta_2, \\ 4\xi_1 + 2\xi_2 \ge 1, & 4\eta_1 \le 1 \\ 3\xi_2 \ge 1, & 2\eta_1 + 3\eta_2 \le 1, \\ \xi_1 \ge 0, \xi_2 \ge 0, & \eta_1 \ge 0, \eta_2 \ge 0. \end{array}$$

Note that, these problems may be written in the equivalent form with constraints in the form of equalities

$$\begin{array}{ll} \min \xi_1 + \xi_2, & \max \eta_1 + \eta_2, \\ 4\xi_1 + 2\xi_2 - \xi_3 = 1, & 4\eta_1 + \eta_3 = 1 \\ 3\xi_2 - \xi_4 = 1, & 2\eta_1 + 3\eta_2 + \eta_4 = 1, \\ \xi_1 \ge 0, \xi_2 \ge 0, \xi_3 \ge 0, \xi_4 \ge 0, & \eta_1 \ge 0, \eta_2 \ge 0, \eta_3 \ge 0, \eta_4 \ge 0. \end{array}$$

Thus, any method of solving the linear programming problems can be used to solve the matrix games. The simplex method is most commonly used to solve such problems. Its systematic discussion may be found in [Ashmanov (1981)], [Gale (1960)], [Hu (1970)].

In a sense, the linear programming problem is equivalent to the matrix game Γ_A. Indeed, consider the following direct and dual problems of linear programming

$$\begin{array}{l} \min xu \\ xA \ge w, \\ x \ge 0, \end{array} \tag{4.14}$$

$$\begin{aligned} &\max yw \\ &Ay \le u, \\ &y \ge 0. \end{aligned} \tag{4.15}$$

Let $\overline{X}$ and $\overline{Y}$ be the sets of optimal solutions of the problems (4.14) and (4.15), respectively. Denote $(1/\Theta)\overline{X} = \{\overline{x}/\Theta\ |\overline{x} \in X\}$, $(1/\Theta)\overline{Y} = \{\overline{y}/\Theta\ |\overline{y} \in Y\}$, $\Theta > 0$.

Theorem 4. *Let Γ_A be the $(m\times n)$ game with the positive matrix A (all elements are positive) and let there be given two dual problems of linear programming (4.14) and (4.15). Then the following statements hold.*

1. *Both linear programming problems have a solution ($\overline{X} \neq \oslash$ and $\overline{Y} \neq \oslash$), in which case*
$$\Theta = \min_x xu = \max_y yw.$$

2. *The value v_A of the game Γ_A is*
$$v_A = 1/\Theta,$$
and the strategies
$$x^* = \overline{x}/\Theta,\ y^* = \overline{y}/\Theta$$
are optimal, where $\overline{x} \in \overline{X}$ is an optimal solution of the direct problem (4.14) and $\overline{y} \in \overline{Y}$ is the solution of the dual problem (4.15).

3. *Any optimal strategies $x^* \in X^*$ and $y^* \in Y^*$ of the players can be constructed as shown above, i.e.*

$$X^* = (1/\Theta)\overline{X},\ Y^* = (1/\Theta)\overline{Y}.$$

Proof. Statements 1, 2 and inclusions $(1/\Theta)\overline{X} \subset X^*$, $(1/\Theta)\overline{Y} \subset Y^*$, immediately follow from the proof of Theorem 3.

Show the inverse inclusion. To do this, consider the vectors $x^* = (\xi_1^*, \ldots, \xi_m^*) \in X^*$ and $\overline{x} = (\overline{\xi}_1, \ldots, \overline{\xi}_m)$, where $\overline{x} = \Theta x^*$. Then for all $j \in N$ we have

$$\overline{x}a^j = \Theta x^* a^j \ge \Theta(1/\Theta) = 1,$$

in which case $\overline{x} \ge 0$, since $\Theta > 0$ and $x^* \ge 0$. Therefore $\overline{x}$ is a feasible solution to problem (4.14).

Let us compute the value of the objective function

$$\overline{x}u = \Theta x^* u = \Theta = \min_x xu,$$

i.e. $\overline{x} \in \overline{X}$ is an optimal solution to problem (4.14).

The inclusion $Y^* \subset (1/\Theta)\overline{Y}$ can be proved in a similar manner. This completes the proof of the theorem.

5 Dominance of strategies

The complexity of solving a matrix game increases as the dimensions of the matrix A increase. In some cases, however, the analysis of payoff matrices permits a conclusion that some pure strategies do not appear in the spectrum of optimal strategy. This can result in replacement of the original matrix by the payoff matrix of a smaller dimension.

Definition 11. *Strategy x' of Player 1 is said to dominate strategy x'' in the $(m \times n)$ game Γ_A if the following inequalities hold for all pure strategies $j \in \{1, \ldots, n\}$ of Player 2*

$$x'a^j \geq x''a^j. \tag{5.1}$$

Similarly, strategy y' of Player 2 dominates his strategy y'' if for all pure strategies $i \in \{1, \ldots, m\}$ of Player 1

$$a_i y' \leq a_i y''. \tag{5.2}$$

If inequalities (5.1), (5.2) are satisfied as strict inequalities, then we are dealing with a *strict dominance.* A special case of the dominance of strategies is their equivalence.

Definition 12. *Strategies x' and x'' of Player 1 are equivalent in the game Γ_A if for all $j \in \{1, \ldots, n\}$*

$$x'a^j = x''a^j.$$

We shall denote this fact by $x' \sim x''$.

For two equivalent strategies x' and x'' the following equality holds (for every $y \in Y$)

$$K(x', y) = K(x'', y).$$

Similarly, strategies y' and y'' of Player 2 are equivalent ($y' \sim y''$) in the game Γ_A if for all $i \in \{1, \ldots, m\}$

$$y'a_i = y''a_i.$$

Hence we have that for any mixed strategy $x \in X$ of Player 1 the following equality holds

$$K(x, y') = K(x, y'').$$

For pure strategies the above definitions are transformed as follows. If Player 1's pure strategy i' dominates strategy i'' and Player 2's pure strategy j' dominates strategy j'' of the same player, then for all $i = 1, \ldots, m$; $j = 1, \ldots, n$ the following inequalities hold

$$a_{i'j} \geq a_{i''j}, \; a_{ij'} \leq a_{ij''}.$$

Definition 13. *The strategy $x''(y'')$ of Player 1(2) is dominated if there exists a strategy $x' \neq x''$ $(y' \neq y'')$ of this player which dominates $x''(y'')$; otherwise strategy $x''(y'')$ is an undominated strategy.*

Similarly, strategy $x''(y'')$ of Player 1(2) is *strictly dominated* if there exists a strategy $x'(y')$ of this player which strictly dominates $x''(y'')$, i.e. for all $j = \overline{1, n}(i = \overline{1, m})$ the following inequalities hold

$$x'a^j > x''a^j, \ a_i y' < a_i y'';$$

otherwise strategy $x''(y'')$ of Player 1(2) is not strictly dominated.

Show that players playing optimally do not use dominated strategies. This establishes the following assertion.

Theorem 5. *If, in the game $\overline{\Gamma}_A$, strategy x' of one of the players dominates an optimal strategy x^*, then strategy x' is also optimal.*

Proof. Let x' and x^* be strategies of Player 1. Then, by dominance,

$$x'a^j \geq x^* a^j$$

for all $j = \overline{1, n}$. Hence, using the optimality of strategy x^*, we get

$$v_A = \min_j x^* a^j \geq \min_j x' a^j \geq \min_j x^* a^j = v_A$$

for all $j = \overline{1, n}$. Therefore, by Theorem 11, strategy x' is also optimal.

Thus, an optimal strategy can be dominated only by another optimal strategy. On the other hand, no optimal strategy is strictly dominated; hence the players when playing optimally must not use strictly dominated strategies.

Theorem 6. *If, in the game $\overline{\Gamma}_A$, strategy x^* of one of the players is optimal, then strategy x^* is not strictly dominated.*

Proof. For definiteness, let x^* be an optimal strategy of Player 1. Assume that x^* is strictly dominated, i.e. there exist such strategy $x' \in X$ that

$$x'a^j > x^* a^j, \ j = 1, 2, \ldots, n.$$

Hence

$$\min_j x' a^j > \min_j x^* a^j.$$

However, by the optimality of $x^* \in X$, the equality $\min\limits_j x^* a^j = v_A$ is satisfied. Therefore, the strict inequality

$$\max_x \min_j x a^j > v_A$$

holds and this contradicts to the fact that v_A is the value of the game. The contradiction proves the theorem.

It is clear that the reverse assertion is generally not true. Thus, in the game with the matrix $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ the first and second strategies of Player 1 are not strictly dominated, although they are not optimal.

On the other hand, it is intuitively clear that if the ith row of the matrix A (the jth column) is dominated, then there is no need to assign positive probability to it. Thus, in order to find optimal strategies instead of the game Γ_A, it suffices to solve a subgame $\Gamma_{A'}$, where A' is the matrix obtained from the matrix A by deleting the dominated rows and columns.

Before proceeding to a precise formulation and proof of this result, we will introduce the notion of an *extension of mixed strategy x at the ith place.* If $x = (\xi_1, \dots, \xi_m) \in X$ and $1 \le i \le m+1$, then the extension of strategy x at the ith place is called the vector $\overline{x}_i = (\xi_1, \dots, \xi_{i-1}, 0, \xi_i, \dots, \xi_m) \in R^{m+1}$. Thus the extension of vector $(1/3, 2/3, 1/3)$ at the 2nd place is the vector $(1/3, 0, 2/3, 1/3)$; the extension at the 4th place is the vector $(1/3, 2/3, 1/3, 0)$; the extension at the 1st place is the vector $(0, 1/3, 2/3, 1/3)$.

Theorem 7. *Let Γ_A be an $(m \times n)$ game. We assume that the ith row of matrix A is dominated (i.e. Player 1's pure strategy i is dominated) and let $\Gamma_{A'}$ be the game with the matrix A' obtained from A by deleting the ith row. Then the following assertions hold.*

1. $v_A = v_{A'}$.
2. *Any optimal strategy y^* of Player 2 in the game $\Gamma_{A'}$ is also optimal in the game Γ_A.*
3. *If x^* is an arbitrary optimal strategy of Player 1 in the game $\Gamma_{A'}$ and $\overline{x}^*_i$ is the extension of strategy x^* at the ith place, then $\overline{x}^*_i$ is an optimal strategy of that player in the game Γ_A.*
4. *If the ith row of the matrix A is strictly dominated, then an arbitrary optimal strategy $\overline{x}^*$ of Player 1 in the game Γ_A can be obtained from an optimal strategy x^* in the game $\Gamma_{A'}$ by the extension at the ith place.*

Proof. Without loss of generality, we may assume, that the last mth row is dominated. Let $x = (\xi_1, \dots, \xi_m)$ be a mixed strategy which dominates the row m. If $\xi_m = 0$, then from the dominance condition for all $j = 1, 2, \dots, n$ we get

$$\sum_{i=1}^{m} \xi_i \alpha_{ij} = \sum_{i=1}^{m-1} \xi_i \alpha_{ij} \ge \alpha_{mj},$$

$$\sum_{i=1}^{m-1} \xi_i = 1, \ \xi_i \ge 0, \ i = 1, \dots, m-1. \tag{5.3}$$

Otherwise ($\xi_m > 0$), consider the vector $x' = (\xi'_1, \dots, \xi'_m)$, where

$$\xi'_i = \begin{cases} \xi_i/(1-\xi_m), & i \ne m, \\ 0, & i = m. \end{cases} \tag{5.4}$$

Components of the vector x are non-negative, ($\xi_i' \geq 0, i = 1, \ldots, m$) and $\sum_{i=1}^m \xi_i' = 1$. On the other hand, for all $i = 1, \ldots, n$ we have

$$\frac{1}{1-\xi_m} \sum_{i=1}^{m} \xi_i \alpha_{ij} \geq \alpha_{mj} \frac{1}{1-\xi_m} \sum_{i=1}^{m} \xi_i$$

or

$$\frac{1}{1-\xi_m} \sum_{i=1}^{m-1} \xi_i \alpha_{ij} \geq \alpha_{mj} \frac{1}{1-\xi_m} \sum_{i=1}^{m-1} \xi_i.$$

Considering (5.4) we get

$$\sum_{i=1}^{m-1} \xi_i' \alpha_{ij} \geq \alpha_{mj} \sum_{i=1}^{m-1} \xi_i' = \alpha_{mj},\ j = 1, \ldots, n,$$

$$\sum_{i=1}^{m-1} \xi_i' = 1,\ \xi_i' \geq 0,\ i = 1, \ldots, m-1. \tag{5.5}$$

Thus, from the dominance of the mth row it always follows that it does not exceed a convex linear combination of the remaining $m-1$ rows (see 5.5).

Let $(x^*, y^*) \in Z(\Gamma_{A'})$ be a saddle point in the game $\Gamma_{A'}$, $x^* = (\xi_1^*, \ldots, \xi_{m-1}^*)$, $y^* = (\eta_1^*, \ldots, \eta_n^*)$. To prove assertions 1, 2 and 3 of the theorem, it suffices to show that $K(x_m^*, y^*) = v_{A'}$ and

$$\sum_{j=1}^{n} \alpha_{ij} \eta_j^* \leq v_{A'} \leq \sum_{i=1}^{m-1} \alpha_{ij} \xi_i^* + 0 \cdot \alpha_{mj} \tag{5.6}$$

for all $i = 1, \ldots, m,\ j = 1, \ldots, n$.

The following theorem is presented without the proof.

Theorem 8. *Let Γ_A be an $(m \times n)$ game. Assume that the jth column of the matrix A is dominated and $\Gamma_{A'}$ is the game having the matrix A' obtained from A by deleting the jth column. Then the following assertions are true.*

1. *$v_A = v_{A'}$.*

2. *Any optimal strategy x^* of Player 1 in the game $\Gamma_{A'}$ is also optimal in the game Γ_A.*

3. *If y^* is an arbitrary optimal strategy of Player 2 in the game $\Gamma_{A''}$ and $\overline{y}_j^*$ is the extension of strategy y at the jth place, then $\overline{y}_j^*$ is an optimal strategy of Player 2 in the game Γ_A.*

4. *Further, if the jth column of the matrix A is strictly dominated, then an arbitrary optimal strategy $\overline{y}^*$ of Player 2 in the game Γ_A can be obtained from an optimal strategy y^* in the game $\Gamma_{A'}$ by extension at the jth place.*

To summarize: The theorems yield an algorithm for reducing the dimension of a matrix game. Thus, if the matrix row (column) is not greater (not smaller) than a convex linear combination of the remaining rows (columns) of the matrix, then to find a solution of the game, this row (column) can be deleted. In this case, an extension of optimal strategy in the truncated matrix game yields an optimal solution of the original game. If the inequalities are satisfied as strict inequalities, the set of optimal strategies in the original game can be obtained by extending the set of optimal strategies in the truncated game; otherwise this procedure may cause a loss of optimal strategies. An application of these theorems is illustrated by the following example.

Example 10. Consider the game with the matrix

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 5 & 3 & 1 & 3 \\ 0 & 1 & 2 & 0 \\ 7 & 3 & 0 & 6 \end{bmatrix}.$$

Since the 3rd row a_3 dominates the 1st row ($a_3 \geq a_1$), then, by deleting the 1st row, we obtain

$$A_1 = \begin{bmatrix} 5 & 3 & 1 & 3 \\ 0 & 1 & 2 & 0 \\ 7 & 3 & 0 & 6 \end{bmatrix}.$$

In this matrix the 1st column a^3 dominates the 3rd column a^1. Hence we get

$$A_2 = \begin{bmatrix} 3 & 1 & 3 \\ 1 & 2 & 0 \\ 3 & 0 & 6 \end{bmatrix}.$$

In the latter matrix no row (column) is dominated by the other row (column). At the same time, the 1st column a^1 is dominated by the convex linear combination of columns a^2 and a^3, i.e. $a^1 \geq 1/2a^2 + 1/2a^3$, since $3 > 1/2 + 1/2 \cdot 3$, $1 = 1/2 \cdot 2 + 1/2 \cdot 0$, $3 = 0 \cdot 1/2 + 1/2 \cdot 6$. By eliminating the 1st column, we obtain

$$A_3 = \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 0 & 6 \end{bmatrix}.$$

In this matrix the 1st row is equal to the linear convex combination of the second and third rows with a mixed strategy $x = (0, 1/2, 1/2)$, since $1 = 1/2 \cdot 2 + 0 \cdot 1/2$, $3 = 0 \cdot 1/2 + 6 \cdot 1/2$. Thus, by eliminating the 1st row, we obtain the matrix

$$A_4 = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}.$$

The players' optimal strategies x^* and y^* in the game with this matrix are $x^* = y^* = (3/4, 1/4)$, in which case the game value v is $3/2$.

The latter matrix is obtained by deleting the first two rows and columns; hence the players' optimal strategies in the original game are extensions of these strategies at the 1st and 2nd places, i.e. $\overline{x}^*_{12} = \overline{y}^*_{12} = (0, 0, 3/4, 1/4)$.

6 Properties of optimal strategies and value of the game

Consider the properties of optimal strategies which, in some cases, assist in finding the value of the game and a saddle point.

Let $(x^*, y^*) \in X \times Y$ be a saddle point in mixed strategies for the game Γ_A. It turns out that, to test the point (x^*, y^*) for a saddle, it will suffice to test the inequalities (4.4) only for $i \in M$ and $j \in N$, not for all $x \in X$ and $y \in Y$, since the following statement is true.

Theorem 9. *For the situation (x^*, y^*) to be an equilibrium (saddle point) in the game Γ_A, and the number $v = K(x^*, y^*)$ be the value, it is necessary and sufficient that the following inequalities hold for all $i \in M$ and $j \in N$:*

$$K(i, y^*) \le K(x^*, y^*) \le K(x^*, j). \tag{6.1}$$

Proof. Necessity. Let (x^*, y^*) be a saddle point in the game Γ_A. Then

$$K(x, y^*) \le K(x^*, y^*) \le K(x^*, y)$$

for all $x \in X$, $y \in Y$. Hence, in particular, for $u_i \in X$ and $w_j \in Y$ we have

$$K(i, y^*) = K(u_i, y^*) \le K(x^*, y^*) \le K(x^*, w_j) = K(x^*, j)$$

for all $i \in M$ and $j \in N$.

Sufficiency. Let (x^*, y^*) be a pair of mixed strategies for which the inequalities (6.1) hold. Also, let $x = (\xi_1, \ldots, \xi_m) \in X$ and $y = (\eta_1, \ldots, \eta_n) \in Y$ be arbitrary mixed strategies for the players 1 and 2, respectively. Multiplying the first and second inequalities (6.1) by ξ_i and η_j, respectively, and summing, we get

$$\sum_{i=1}^{m} \xi_i K(i, y^*) \le K(x^*, y^*) \sum_{i=1}^{m} \xi_i = K(x^*, y^*), \tag{6.2}$$

$$\sum_{j=1}^{n} \eta_j K(x^*, j) \ge K(x^*, y^*) \sum_{j=1}^{n} \eta_j = K(x^*, y^*). \tag{6.3}$$

In this case we have

$$\sum_{i=1}^{m} \xi_i K(i, y^*) = K(x, y^*), \tag{6.4}$$

$$\sum_{j=1}^{n} \eta_j K(x^*, j) = K(x^*, y). \tag{6.5}$$

Substituting (6.4), (6.5) into (6.2) and (6.3), respectively, and taking into account the arbitrariness of strategies $x \in X$ and $y \in Y$, we obtain saddle point conditions for the pair of mixed strategies (x^*, y^*).

Corollary 3. *Let (i^*, j^*) be a saddle point in the game Γ_A. Then the situation (i^*, j^*) is also a saddle point in the game $\overline{\Gamma}_A$.*

Example 11. Solution of the Evasion-type Game. Suppose the players select integers i and j between 1 and n, and Player 1 wins the amount $a_{ij} = |i - j|$, i.e. the distance between the numbers i and j.

Suppose the first player uses strategy $x^* = (1/2, 0, \ldots, 0, 1/2)$. Then

$$K(x^*, j) = 1/2|1 - j| + 1/2|n - j| = 1/2(j - 1) + 1/2(n - j) = (n - 1)/2$$

for all $1 \leq j \leq n$.

a) Let $n = 2k+1$ be odd. Then Player 2 has a pure strategy $j^* = (n+1)/2 = k + 1$ such that

$$a_{ij^*} = |i - (n + 1)/2| = |i - k - 1| \leq k = (n - 1)/2$$

for all $i = 1, 2, \ldots, n$.

b) Let $n = 2k$ be even. Then Player 2 has a strategy
$y^* = (0, 0, \ldots, 1/2, 1/2, 0, \ldots, 0)$, where $\eta_k^* = 1/2$, $\eta_{k+1}^* = 1/2$, $\eta_j^* = 0$, $j \neq k+1$, $j \neq k$, and

$$K(j, y^*) = 1/2|i - k| + 1/2|i - k - 1| \leq 1/2k + 1/2(k - 1) = (n - 1)/2$$

for all $1 \leq i \leq n$.

It can be easily seen that the value of the game is $v = (n - 1)/2$, Player 1 has optimal strategy x^*, and Player 2's optimal strategy is j^* if $n = 2k + 1$, and y^* if $n = 2k$.

Theorem 10. *Let Γ_A be an $(m \times n)$ game. For the situation in mixed strategies, let (x^*, y^*) be an equilibrium (saddle point) in the game $\overline{\Gamma}_A$, it is necessary and sufficient that the following equality holds*

$$\max_{1 \leq i \leq m} K(i, y^*) = \min_{1 \leq j \leq n} K(x^*, j). \tag{6.6}$$

Proof.

Necessity. If (x^*, y^*) is a saddle point, then, by Theorem 9, we have

$$K(i, y^*) \leq K(x^*, y^*) \leq K(x^*, j)$$

for all $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$. Therefore

$$K(i, y^*) \leq K(x^*, j)$$

for each i and j. Suppose the opposite is true, i.e. (6.6) is not satisfied. Then

$$\max_{1 \leq i \leq m} K(i, y^*) < \min_{1 \leq j \leq n} K(x^*, j).$$

Consequently, the following inequalities hold

$$K(x^*, y^*) = \sum_{i=1}^{m} \xi_i^* K(i, y^*) \le \max_{1\le i\le m} K(i, y^*) < \min_{1\le j\le n} K(x^*, j)$$

$$\le \sum_{j=1}^{n} \eta_j^* K(x^*, j) = K(x^*, y^*).$$

The obtained contradiction proves the necessity of the Theorem assertion.

Sufficiency. Let a pair of mixed strategies $(\tilde{x}, \tilde{y})$ be such that $\max_i K(i, \tilde{y}) = \min_j K(\tilde{x}, j)$. Show that in this case $(\tilde{x}, \tilde{y})$ is a saddle point in the game $\overline{\Gamma}_A$.

The following relations hold

$$\min_{1\le j\le n} K(\tilde{x}, j) \le \sum_{j=1}^{n} \tilde{\eta}_j K(\tilde{x}, j) = K(\tilde{x}, \tilde{y})$$

$$= \sum_{i=1}^{m} \tilde{\xi}_i K(i, \tilde{y}) \le \max_{1\le i\le m} K(i, \tilde{y}).$$

Hence we have

$$K(i, \tilde{y}) \le \max_{1\le i\le m} K(i, \tilde{y}) = K(\tilde{x}, \tilde{y}) = \min_{1\le j\le n} K(\tilde{x}, j) \le K(\tilde{x}, j)$$

for all $1 \le i \le m$ and $1 \le j \le n$, then, by the Theorem 9, $(\tilde{x}, \tilde{y})$ is the saddle point in the game $\overline{\Gamma}_A$.

From the proof it follows that any one of the numbers in (6.6) is the value of the game.

Theorem 11. *The following relation holds for the matrix game Γ_A*

$$\max_x \min_j K(x, j) = v_A = \min_y \max_i K(i, y), \tag{6.7}$$

in which case the extrema are achieved on the players' optimal strategies.

This theorem follows from the Theorems 2 and 10, and its proof is left to the reader.

Theorem 12. *In the matrix game Γ_A the players' sets of optimal mixed strategies X^* and Y^* are convex polyhedra.*

Proof. By Theorem 9, the set X^* is the set of all solutions of the system of inequalities

$$xa^j \ge v_A, \ j \in N,$$

$$xu = 1,$$

$$x \ge 0,$$

where $u = (1, \dots, 1) \in R^m$, v_A is the value of the game. Thus, X^* is a convex polyhedral set. On the other hand, $X^* \subset X$, where X is a convex polyhedron. Therefore X^* is bounded and a convex polyhedron.

In a similar manner, it may be proved that Y^* is a convex polyhedron.

As an application of Theorem 11, we shall provide a geometric solution to the games with two strategies for one of the players $(2 \times n)$ and $(m \times 2)$ games. This method is based on the property that the optimal strategies x^* and y^* deliver exterior extrema in the equality

$$v_A = \max_x \min_j K(x, j) = \min_y \max_i K(i, y).$$

Example 12. $(2 \times n)$ *game.* We shall examine the game in which Player 1 has two strategies and Player 2 has n strategies. The matrix is of the form

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \end{bmatrix}.$$

Suppose Player 1 chooses mixed strategy $x = (\xi, 1-\xi)$ and Player 2 chooses pure strategy $j \in N$. Then a payoff to Player 1 at (x, j) is

$$K(x, j) = \xi\alpha_{1j} + (1 - \xi)\alpha_{2j}. \tag{6.8}$$

Geometrically, the payoff is a straight line segment with coordinates (ξ, K). Accordingly, to each pure strategy j corresponds a straight line. The graph of the function

$$H(\xi) = \min_j K(x, j)$$

is the lower envelope of the family of straight lines (6.8). This function is concave as the lower envelope of the family of concave (linear in the case) function. The point ξ^*, at which the maximum of the function $H(\xi)$ is achieved with respect to $\xi \in [0, 1]$, yields the required optimal solution $x^* = (\xi^*, 1-\xi^*)$ and the value of the game $v_A = H(\xi^*)$.

For definiteness, we shall consider the game with the matrix

$$A = \begin{bmatrix} 1 & 3 & 1 & 4 \\ 2 & 1 & 4 & 0 \end{bmatrix}.$$

For each $j = 1, 2, 3, 4$ we have: $K(x, 1) = -\xi+2$, $K(x, 2) = 2\xi+1$, $K(x, 3) = -3\xi + 4$, $K(x, 4) = 4\xi$. The lower envelope $N(\xi)$ of the family of straight lines $\{K(x, j)\}$ and the lines themselves, $K(x, j)$, $j = 1, 2, 3, 4$ are shown in Fig. 1.1. The maximum $H(\xi^*)$ of the function $H(\xi)$ is found as the intersection of the first and the fourth lines. Thus, ξ^* is a solution of the equation.

$$4\xi^* = -\xi^* + 2 = v_A.$$

Hence we get the optimal strategy $x^* = (2/5, 3/5)$ of Player 1 and the value of the game is $v_A = 8/5$. Player 2's optimal strategy is found from the following reasonings. Note that in the case studied $K(x^*, 1) = K(x^*, 4) = v_A = 8/5$.

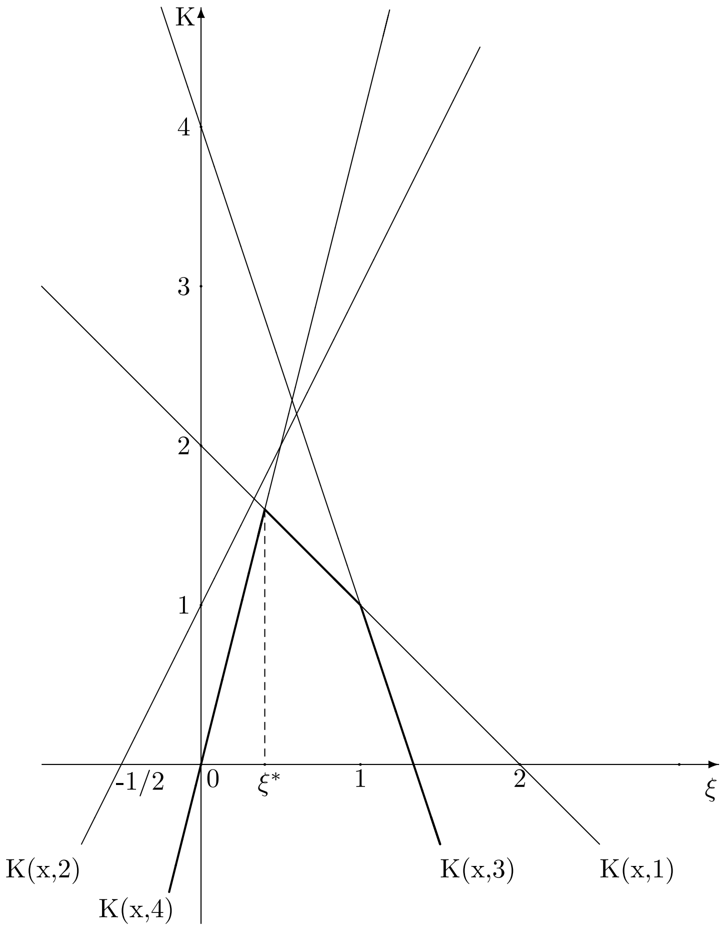


Figure 1:

For the optimal strategy $y^* = (\eta_1^*, \eta_2^*, \eta_3^*, \eta_4^*)$ the following equality must hold

$$v_A = K(x^*, y^*) = \eta_1^* K(x^*, 1) + \eta_2^* K(x^*, 2) + \eta_3^* K(x^*, 3) + \eta_4^* K(x^*, 4).$$

In this case $K(x^*, 2) > 8/5$, $K(x^*, 3) > 8/5$; therefore $\eta_2^* = \eta_3^* = 0$, and η_1^*, η_4^* can be found from the conditions

$$\eta_1^* + 4\eta_4^* = 8/5,$$

$$2\eta_1^* = 8/5.$$

Thus, $\eta_1^* = 4/5$, $\eta_4^* = 1/5$ and the optimal strategy of Player 2 is $y^* = (4/5, 0, 0, 1/5)$.

Example 13. *$(m \times 2)$ game.* In this example, Player 2 has two strategies and Player 1 has m strategies. The matrix A is of the form

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \dots & \dots \\ \alpha_{m1} & \alpha_{m2} \end{bmatrix}.$$

This game can be analyzed in a similar manner. Indeed, let $y = (\eta, 1-\eta)$ be an arbitrary mixed strategy of Player 2. Then Player 1's payoff in situation (i, y) is

$$K(i, y) = \alpha_{i1}\eta + \alpha_{i2}(1-\eta) = (\alpha_{i1} - \alpha_{i2})\eta + \alpha_{i2}.$$

The graph of the function $K(i, y)$ is a straight line. Consider the upper envelope of these straight lines, i.e. the function

$$H(\eta) = \max_i [(\alpha_{i1} - \alpha_{i2})\eta + \alpha_{i2}].$$

The function $H(\eta)$ is convex (as the upper envelope of the family of convex functions).

The point of minimum η^* of the function $H(\eta)$ yields the optimal strategy $y^* = (\eta^*, 1-\eta^*)$ and the value of the game is $v_A = H(\eta^*) = \min\limits_{\eta \in [0,1]} H(\eta)$.

We shall provide a theorem that is useful in finding a solution of the game.

Theorem 13. *Let $x^* = (\xi_1^*, \dots, \xi_m^*)$ and $y^* = (\eta_1^*, \dots, \eta_n^*)$ be optimal strategies in the game $\overline{\Gamma}_A$ and v_A be the value of the game. Then for any i, for which $K(i, y^*) < v_A$, there must be $\xi_i^* = 0$, and for any j such that $v_A < K(x^*, j)$ there must be $\eta_j^* = 0$.*

Conversely, if $\xi_i^* > 0$, then $K(i, y^*) = v_A$, and if $\eta_j^* > 0$, then $K(x^*, j) = v_A$.

Proof. Suppose that for some $i_0 \in M$, $K(i_0, y^*) < v_A$ and $\xi_{i_0}^* \neq 0$. Then we have

$$K(i_0, y^*)\xi_{i_0}^* < v_A \xi_{i_0}^*.$$

For all $i \in M$, $K(i, y^*) \le v_A$, therefore

$$K(i, y^*)\xi_i^* \le v_A \xi_i^*.$$

Consequently, $K(x^*, y^*) < v_A$, which contradicts to the fact that v_A is the value of the game. The second part of the Theorem can be proved in a similar manner.

This result is a counterpart of the complementary stackness theorem [Hu (1970)] or, as it is sometimes called the canonical equilibrium theorem for the linear programming problem [Gale (1960)].

Definition 14. *Player 1's (2's) pure strategy $i \in M (j \in N)$ is called an essential or active strategy if there exists the player's optimal strategy $x^* = (\xi_1^*, \dots, \xi_m^*)$ $(y^* = (\eta_1^*, \dots, \eta_n^*))$ for which $\xi_i^* > 0$ $(\eta_j^* > 0)$.*

From the definition, and from the latter theorem, it follows that for each essential strategy i of Player 1 and any optimal strategy $y^* \in Y^*$ of Player 2 in the game Γ_A the following equality holds:

$$K(i, y^*) = a_i y^* = v_A.$$

A similar equality holds for any essential strategy $j \in N$ of Player 2 and for the optimal strategy $x^* \in X^*$ of Player 1

$$K(x^*, j) = a^j x^* = v_A.$$

If the equality $a_i y = v_A$ holds for the pure strategy $i \in M$ and mixed strategy $y \in Y$, then the strategy i is the best reply to the mixed strategy y in the game Γ_A.

Thus, using this terminology, the theorem can be restated as follows. If the pure strategy of the player is essential, then it is the best reply to any optimal strategy of the opponent.

A knowledge of the optimal strategy spectrum simplifies to finding a solution of the game. Indeed, let M_{X^*} be the spectrum of Player 1's optimal strategy x^*. Then each optimal strategy $y^* = (\eta_1^*, \dots, \eta_n^*)$ of Player 2 and the value of the game v satisfy the system of inequalities

$$a_i y^* = v, \ i \in M_{x^*},$$

$$a_i y^* \leq v, \ i \in M \setminus M_{x^*},$$

$$\sum_{j=1}^{n} \eta_j^* = 1, \ \eta_j^* \geq 0, \ j \in N.$$

Thus, only essential strategies may appear in the spectrum M_{x^*} of any optimal strategy x^*.

7 Methods for solving matrix games

Analytical solution method

Example 14. Consider the analytical solution of Attack and Defence game Let us consider the game with the $(n \times n)$ matrix A.

$$A = \begin{bmatrix} \beta_1\tau_1 & \tau_1 & \dots & \tau_1 \\ \tau_2 & \beta_2\tau_2 & \dots & \tau_2 \\ \dots & \dots & \dots & \dots \\ \tau_n & \tau_n & \dots & \beta_n\tau_n \end{bmatrix}.$$

Here $\tau_i > 0$ is the value and $0 < \beta_i < 1$ is the probability of hitting the target $C_i, i = 1, 2, \dots, n$ provided that it is defended.

Let $\tau_1 \leq \tau_2 \leq \ldots \leq \tau_n$. We shall define the function φ, of integers $1, 2, \ldots, n$ as follows:

$$\varphi(k) = \{\sum_{i=k}^{n}(1-\beta_i)^{-1} - 1\}/\sum_{i=k}^{n}(\tau_i(1-\beta_i))^{-1} \tag{7.1}$$

and let $l \in \{1, 2, \ldots, n\}$ be an integer which maximize the function $\varphi(k)$, i.e.

$$\varphi(l) = \max_{k=1,2,\ldots,n} \varphi(k). \tag{7.2}$$

We shall establish properties of the function $\varphi(k)$. Denote by R one of the signs of the order relation $\{>, =, <\}$. In this case

$$\varphi(k) R \varphi(k+1) \tag{7.3}$$

if and only if

$$\tau_k R \varphi(k),\ k = 1, 2, \ldots, n-1,\ \tau_0 \equiv 0. \tag{7.4}$$

Indeed, from (7.1) we obtain

$$\frac{\varphi(k)}{\tau_k}\frac{(1-\beta_k)^{-1}}{\sum_{i=k+1}^{n}(\tau_i(1-\beta_i))^{-1}} + \varphi(k) = \varphi(k+1)$$

$$+\frac{(1-\beta_k)^{-1}}{\sum_{i=k+1}^{n}(\tau_i(1-\beta_i))^{-1}}.$$

Then we have

$$\left(\frac{\varphi(k)}{\tau_k} - 1\right)\frac{(1-\beta_k)^{-1}}{\sum_{i=k+1}^{n}(\tau_i(1-\beta_i))^{-1}} + \varphi(k) = \varphi(k+1). \tag{7.5}$$

Note that the coefficient in (7.5) placed after brackets, is positive. Therefore, from (7.5) we obtained equivalence of relations (7.3) and (7.4).

Now, since $\varphi(l) \geq \varphi(l-1)$ or $\varphi(l) \geq \varphi(l+1)$, (in this case $\tau_{l-1} \leq \varphi(l-1)$ or $\tau_l \geq \varphi(l)$), then from relations (7.2), (7.3) we have

$$\tau_{l-1} \leq \varphi(l) \leq \tau_l. \tag{7.6}$$

Find optimal strategies in the game Γ_A. Recall that we have inequalities $\tau_1 \leq \tau_2 \leq \ldots \tau_n$. Then the optimal strategies $x^* = (\xi_1^*, \ldots, \xi_n^*)$ and $y^* = (\eta_1^*, \ldots, \eta_n^*)$ for players 1 and 2 respectively, are as follows:

$$\xi_i^* = \begin{cases} 0,\ i = 1, \ldots, l-1, \\ (\tau_i(1-\beta_i))^{-1}/\sum_{j=l}^{n}(\tau_j(1-\beta_j))^{-1},\ i = l, \ldots, n, \end{cases} \tag{7.7}$$

$$\eta_j^* = \begin{cases} 0,\ j = 1, \ldots, l-1, \\ (\tau_j - \varphi(l))/(\tau_j(1-\beta_j)),\ j = l, \ldots, n, \end{cases} \tag{7.8}$$

and the value of the game is

$$v_A = \varphi(l).$$

We have that $\xi_i^* \geq 0, i = 1, 2, \ldots, n$ and $\sum_{i=1}^n \xi_i^* = 1$. From the definition of $\varphi(l)$ and (7.6) we have that $\eta_j^* \geq 0$, $j = 1, 2, \ldots, n$ and $\sum_{j=1}^n \eta_j^* = 1$.

Let $K(x^*, j)$ be a payoff of Player 1 at (x^*, j). Similarly, let $K(i, y^*)$ be a payoff at (i, y^*). Substituting (7.7), (7.8) into the payoff function and using the assumption that the values of targets do not decrease, and using (7.6), we obtain

$$K(x^*, j) = \begin{cases} \sum\limits_{i=l}^{n} \tau_i \xi_i^* = \varphi(l) + \sum\limits_{j=l}^{n} (\tau_j(1-\beta_j))^{-1} > \varphi(l),\ j = \overline{1, l-1}, \\ \sum\limits_{i=l}^{n} \tau_i \xi_i^* - (1-\beta_j)\tau_j \xi_j^* = \varphi(l),\ j = \overline{l, n}, \end{cases}$$

$$K(i, y^*) = \begin{cases} \tau_i \leq \varphi(l),\ i = \overline{1, l-1}, \\ \tau_i - \tau_i(1-\beta_i)\eta_i^* = \varphi(l),\ i = \overline{l, n}. \end{cases}$$

Thus, for all $i, j = 1, \ldots, n$ the following inequalities hold

$$K(i, y^*) \leq \varphi(l) \leq K(x^*, j).$$

Then, by Theorem 9, x^* and y^* are optimal and $v_A = \varphi(l)$. This completes the solution of the game.

Iterative method

Consider Brown-Robinson iterative method (fictitious play method). This method employs a repeated fictitious play of game having a given payoff matrix. One repetition of the game is called a play. Suppose the game is played with an $(m \times n)$ matrix $A = \{a_{ij}\}$. In the 1st play both players choose arbitrary pure strategies. In the kth play each player chooses the pure strategy which maximizes his expected payoff against the observed empirical probability distribution of the opponents pure strategies for $(k-1)$ plays.

Thus, we assume that in the first k plays Player 1 uses the ith strategy ξ_i^k times $(i = 1, \ldots, m)$ and Player 2 uses the jth strategy η_j^k times $(j = 1, \ldots, n)$. In the $(k+1)$ play, Player 1 will then use i_{k+1} strategy and Player 2 will use his j_{k+1} strategy, where

$$\overline{v}^k = \max_i \sum_j a_{ij} \eta_j^k = \sum_j a_{i_{k+1} j} \eta_j^k$$

and

$$\underline{v}^k = \min_j \sum_i a_{ij} \xi_i^k = \sum_i a_{i j_{k+1}} \xi_i^k.$$

Let v be the value of the matrix game Γ_A. Consider the expressions

$$\overline{v}^k / k = \max_i \sum_j \alpha_{ij} \eta_j^k / k = \sum_j \alpha_{i_{k+1} j} \eta_j^k / k,$$

$$\underline{v}^k/k = \min_j \sum_i \alpha_{ij}\xi_i^k/k = \sum_i \alpha_{ij_{k+1}}\xi_i^k/k.$$

The vectors $x^k = (\xi_1^k/k, \ldots, \xi_m^k/k)$ and $y^k = (\eta_1^k/k, \ldots, \eta_n^k/k)$ are mixed strategies for the players 1 and 2, respectively; hence, by the definition of the value of the game we have

$$\max_k \underline{v}^k/k \le v \le \min_k \overline{v}^k/k.$$

We have thus obtained an iterative process which enables us to find an approximate solution of the matrix game, the degree of approximation to the true value of the game being determined by the length of the interval $[\max_k \underline{v}^k/k, \min_k \overline{v}^k/k]$. Convergence of the algorithm is guaranteed by the Theorem [Robinson (1950)].

Theorem 14.

$$\lim_{k\to\infty}(\min_k \overline{v}^k/k) = \lim_{k\to\infty}(\max_k \underline{v}^k/k) = v.$$

Example 15. Find an approximate solution to the game having the matrix

$$A = \begin{array}{c} \\ \alpha \\ \beta \\ \gamma \end{array} \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{bmatrix} 2 & 1 & 3 \\ 3 & 0 & 1 \\ 1 & 2 & 1 \end{bmatrix} \end{array}.$$

Denote Player 1's strategies by α, β, γ, and Player 2's strategies by a, b, c. Suppose the players first choose strategies α and a, respectively. If Player 1 chooses strategy α, then Player 1 can receive one of the payoffs (2,1,3). If Player 2 chooses strategy a, then Player 1 can receive one of the payoffs (2,3,1). In the 2nd and 3rd plays, Player 1 chooses strategy β and Player 2 chooses strategy b, since these strategies ensure the best result, etc.

Table 1 shows the results of plays, the players' strategies, the accumulated payoff, and the average payoff.

Thus, for 12 plays, we obtain an approximate solution

$$x^{12} = (1/4, 1/6, 7/12), \quad y^{12} = (1/12, 7/12, 1/3)$$

and the accuracy can be estimated by the number 5/12. The principal disadvantage of this method is its low speed of convergence which decreases as the matrix dimension increases. This also results from the nonmonotonicity of sequences $\overline{v}^k/k$ and $\underline{v}^k/k$.

Consider another iteration algorithm which is free of the above-mentioned disadvantages.

Monotonic iterative method of solving matrix games. [Sadovsky (1978)].

We consider a mixed extension $\Gamma_A = (X, Y, K)$ of the matrix game with the $(m \times n)$ matrix A.

Denote by $x^N = (\xi_1^N, \ldots, \xi_m^N) \in X$ the approximation of Player 1's optimal strategy at the Nth iteration, and by $c^N \in R^N$, $c^N = (\gamma_1^N, \ldots, \gamma_n^N)$ an auxiliary vector. Algorithm makes it possible to find (exactly and approximately) an optimal strategy for Player 1 and a value of the game v.

Play No	Player 1's choice	Player 2's choice	Player 1's payoff			Player 2's payoff			$\frac{\overline{v}^k}{k}$	$\frac{\underline{v}^k}{k}$
			α	β	γ	a	b	c		
1	α	a	2	3	1	2	1	3	3	1
2	β	b	3	3	3	5	1	4	3/2	1/2
3	β	b	4	3	5	8	1	5	5/3	1/3
4	γ	b	5	3	7	9	3	6	7/4	3/4
5	γ	b	6	3	9	10	5	7	9/5	5/5
6	γ	b	7	3	11	11	7	8	11/6	7/6
7	γ	b	8	3	13	12	9	9	13/7	9/7
8	γ	c	14	4	14	13	12	10	14/8	10/8
9	γ	c	14	5	15	14	12	11	15/9	11/9
10	γ	c	17	6	16	15	14	12	17/10	12/10
11	α	c	20	7	17	17	15	15	20/11	15/11
12	α	b	21	7	19	19	16	18	21/12	16/12

Table 1:

At the start of the process, Player 1 chooses an arbitrary vector of the form $c^0 = a_{i_0}$, where a_{i_0} is the row of the matrix A having the number i_0.

Iterative process is constructed as follows. Suppose the $N-1$ iteration is performed and vectors x^{N-1}, c^{N-1} are obtained. Then x^N and c^N are computed from the following iterative formulas

$$x^N = (1-\alpha_N)x^{N-1} + \alpha_N \tilde{x}^N, \tag{7.9}$$

$$c^N = (1-\alpha_N)c^{N-1} + \alpha_N \tilde{c}^N, \tag{7.10}$$

where $0 \leq \alpha_N \leq 1$. Vectors $\tilde{x}^N$ and $\tilde{c}^N$ will be obtained below.

We consider the vector $c^{N-1} = (\gamma_1^{N-1}, \ldots, \gamma_n^{N-1})$ and select indices j_k on which the minimum is achieved

$$\min_{j=1,\ldots,n} \gamma_j^{N-1} = \gamma_{j_1}^{N-1} = \gamma_{j_2}^{N-1} = \ldots = \gamma_{j_k}^{N-1}.$$

Denote

$$\underline{v}^{N-1} = \min_{j=1,\ldots,n} \gamma_j^{N-1} \tag{7.11}$$

and $J^{N-1} = \{j_1, \ldots, j_k\}$ be the set of indices on which (7.11) is achieved.

Let $\Gamma^N \subset \Gamma_A$ be a subgame of the game Γ_A with the matrix $A^N = \{a_{ij}^{N-1}\}$, $i = 1, \ldots, m$, and the index $j^{N-1} \in J^{N-1}$. Solve the subgame and find an optimal strategy $\tilde{x}^N \in X$ for Player 1. Let $\tilde{x}^N = (\tilde{\xi}_1^N, \ldots, \tilde{\xi}_m^N)$.

Compute the vector $\tilde{c}^N = \sum_{i=1}^m \tilde{\xi}_i^N a_i$. Suppose the vector $\tilde{c}^N$ has components $\tilde{c}^N = (\tilde{\gamma}_1^N, \ldots, \tilde{\gamma}_n^N)$. Consider the $(2 \times n)$ game with matrix

$$\begin{bmatrix} \gamma_1^{N-1} & \ldots & \gamma_n^{N-1} \\ \tilde{\gamma}_1^N & \ldots & \tilde{\gamma}_n^N \end{bmatrix}.$$

Find Player 1's optimal strategy $(\alpha_N, 1-\alpha_N)$, $0 \le \alpha_N \le 1$ in this subgame.

Substituting the obtained values $\tilde{x}^N, \tilde{c}^N, \alpha_N$ into (7.9), (7.10), we find x^N and c^N. We continue the process until the equality $\alpha_N = 0$ is satisfied or the required accuracy of computations is achieved. Convergence of the algorithm is guaranteed by the following theorem [Sadovsky (1978)].

Theorem 15. *Let $\{\underline{v}^N\}, \{x^N\}$ be the iterative sequences determined by (7.9), (7.11). Then the following assertions are true.*

1. $\underline{v}^N > \underline{v}^{N-1}$, *i.e. the sequence $\{\underline{v}^{N-1}\}$ strictly and monotonically increases*

2.
$$\lim_{N\to\infty} \underline{v}^N = \underline{v} = v \tag{7.12}$$

3. $\lim\limits_{N\to\infty} x^N = x^*$, *where $x^* \in X^*$ is an optimal strategy of Player 1.*

Example 16. By employing a monotonic algorithm, solve the game with the matrix

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 3 & 0 & 1 \\ 1 & 2 & 1 \end{bmatrix}.$$

Iteration 0. Suppose Player 1 chooses the 1st row of the matrix A, i.e. $x^* = (1,0,0)$ and $c^0 = a_1 = (2,1,3)$. Compute $\underline{v}^0 = \min\limits_j \gamma_j^0 = \gamma_2^0 = 1$, $J^0 = 2$.

Iteration 1. Consider the subgame $\Gamma^1 \subset \Gamma$ having the matrix

$$A^1 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}.$$

An optimal strategy $\tilde{x}^1$ of Player 1 is the vector $\tilde{x}^1 = (0,0,1)$. Then $\tilde{c}^1 = a_3 = (1,2,1)$. Solve the (2×3) game with the matrix $\begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 1 \end{bmatrix}$. Note that the 3rd column of the matrix is dominated and so we consider the matrix $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. Because of the symmetry, Player 1's optimal strategy in this game is the vector $(\alpha_N, 1-\alpha_N) = (1/2, 1/2)$.

We compute x^1 and c^1 by formulas (7.9), (7.10). We have

$$\begin{aligned} &x^1 = 1/2x^0 + 1/2\tilde{x}^1 = (1/2, 0, 1/2), \\ &c^1 = 1/2c^0 + 1/2\tilde{c}^1 = (3/2, 3/2, 2), \\ &\underline{v}^1 = \min_j \gamma_j^1 = \gamma_1^1 = \gamma_2^1 = 3/2 > \underline{v}^0 = 1. \end{aligned}$$

The set of indices is of the form $J^1 = \{1, 2\}$.

Iteration 2. Consider the subgame $\Gamma^2 \subset \Gamma$ having the matrix

$$A = \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 1 & 2 \end{bmatrix}.$$

The first row in this matrix is dominated; hence it suffices to examine the submatrix

$$\begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix}.$$

Player 1's optimal strategy in this game is the vector $(1/4, 3/4)$; hence $\tilde{x}^2 = (0, 1/4, 3/4)$.

Compute $\tilde{c}^2 = 1/4a_2 + 3/4a_3 = (3/2, 3/2, 1)$ and consider the (2×3) game with the matrix

$$\begin{bmatrix} 3/2 & 3/2 & 1 \\ 3/2 & 3/2 & 2 \end{bmatrix}.$$

The second strategy of Player 1 dominates the first strategy and hence $\alpha_2 = 0$. This completes the computations $x^* = x^1 = (1/2, 0, 1/2)$; the value of the game is $v = \underline{v}^1 = 3/2$, and Player 2's optimal strategy is of the form $y^* = (1/2, 1/2, 0)$.

References

[Ashmanov (1981)] Ashmanov, S. A. *Linear Programming*, p. 198. Nauka, Moscow, 1981.

[Gale (1960)] Gale, D. *The Theory of Linear Economic Models*, p. 330. McGraw–Hill Book comp., inc., N.Y., London, 1960.

[Hu (1970)] Hu, T. *Integer Programming and Network Flows*, p. 411. Addison-Wesley Publ. Comp., Mento Park, Calif.-London-Dou Hills, 1970.

[Sakaguchi, 1973] *Sakaguchi, M.* Oper. Res. Soc. Jap., 16, 1973.

[Robinson (1950)] Robinson, G. B. *An iteration method of solving a game*, volume P-154, p. 9. RAND Corp., 1950.

[Sadovsky (1978)] Sadovsky, A. L. *Soviet Math. Reports*, 238(3), 1978.

Nonzero-sum games

November 28, 2018

1 Definition of noncooperative game in normal form

The preceding chapters concentrated on zero-sum two-person games, i.e. the games in which the interests of the parties are strictly contradictory. However, a special feature of the actual problems of decisions making in a conflict context is that there are too many persons involved, with the result that the conflict situation is far from being strictly contradictory. As for a two-person conflict and its models, it may be said that such a conflict is not confined to the antagonistic case alone. Although the players' interests may intersect, they are not necessarily contradictory. This, in particular, can involve situations that are of mutual benefit to both players (which is not possible in the antagonistic conflict). Cooperation (selection of an agreed decision) is thus made meaningful and tends to increase a payoff to both players. At the same time, there are conflicts for which the rules of a game do not specify any agreement or cooperation. For this reason, in non zero-sum games, a distinction is made between noncooperative behavior, where the rules do not allow any cooperation and cooperative behavior, where the rules allow cooperation in the joint selection of strategies and side payments making. We shall consider the former case.

Definition 1. *The system*

$$\Gamma = (N, \{X_i\}_{i\in N}, \{H_i\}_{i\in N}),$$

where $N = \{1, 2, \ldots, n\}$ is the set of players, X_i is the strategy set for player i, H_i is the payoff function for player i defined on Cartesian product of the players' strategy sets $X = \prod_{i=1}^{n} X_i$ (the set of situations in the game), is called a noncooperative game.

A noncooperative n-person game is played as follows. Players choose simultaneously and independently their strategies x_i from the strategy sets X_i, $i = 1, 2, \ldots, n$, thereby generating a situation $x = (x_1, \ldots, x_n)$, $x_i \in X_i$. Each player i receives the amount $H_i(x)$, whereupon the game ends.

If the players' pure strategy sets X_i are finite, the game is called a *finite noncooperative n-person game.*

The noncooperative game Γ played by two players is called a *two-person game*. The noncooperative two-person game Γ is then defined by the system $\Gamma = (X_1, X_2, H_1, H_2)$, where X_1 is the strategy set of one player, X_2 is the strategy set of the other player, $X_1 \times X_2$ is the set of situations, while $H_1 : X_1 \times X_2 \to R^1$, $H_2 : X_1 \times X_2 \to R^1$ are the payoff functions to the players 1 and 2, respectively. The finite noncooperative two-person game is called *bimatrix game*. This is due to the fact that, once the pure strategy sets of players have been designated by the numbers $1, 2, \ldots, \overline{m}$ and $1, 2, \ldots, \overline{n}$, the payoff functions can be written in the form of two matrices

$$H_1 \triangleq A \triangleq \begin{bmatrix} \alpha_{11} & \ldots & \alpha_{1\bar{n}} \\ \ldots & \ldots & \ldots \\ \alpha_{\bar{m}1} & \ldots & \alpha_{\bar{m}\bar{n}} \end{bmatrix} \text{ and } H_2 \triangleq B \triangleq \begin{bmatrix} \beta_{11} & \ldots & \beta_{1\bar{n}} \\ \ldots & \ldots & \ldots \\ \beta_{\bar{m}1} & \ldots & \beta_{\bar{m}\bar{n}} \end{bmatrix}.$$

Here the elements α_{ij} and β_{ij} of the matrices A, B are respectively the payoffs to players 1 and 2 in the situation (i, j), $i \in \overline{M}$, $j \in \overline{N}$, $\overline{M} = \{1, \ldots, \overline{m}\}$, $\overline{N} = \{1, \ldots, \overline{n}\}$.

In line with the foregoing, the bimatrix game is played as follows. Player 1 chooses number i (the row) and Player 2 (simultaneously and independently) chooses number j (the column). Then Player 1 receives the amount $\alpha_{ij} \triangleq H_1(x_i, y_j)$ and Player 2 receives the amount $\beta_{ij} \triangleq H_2(x_i, y_j)$.

Note that the bimatrix game with matrices A and B can also be described by the $(\overline{m} \times \overline{n})$ matrix (A, B), where each element is a pair $(\alpha_{ij}, \beta_{ij})$, $i = 1, 2, \ldots, \overline{m}$, $j = 1, 2, \ldots, \overline{n}$. The game determined by the matrix A and B will be denoted as $\Gamma(A, B)$.

If the noncooperative two-person game Γ is such that $H_1(x, y) = -H_2(x, y)$ for all $x \in X_1$, $y \in X_2$, then Γ appears to be a zero-sum two-person game discussed in the preceding chapters. In the special bimatrix game, where there is $\alpha_{ij} = -\beta_{ij}$, we have a matrix game examined in Chapter 1.

Example 1. Battle of the sexes. Consider the bimatrix game determined by

$$(A, B) = \begin{array}{c} \\ \alpha_1 \\ \alpha_2 \end{array} \begin{array}{c} \begin{array}{cc} \beta_1 & \beta_2 \end{array} \\ \begin{bmatrix} (4, 1) & (0, 0) \\ (0, 0) & (1, 4) \end{bmatrix} \end{array}.$$

Although this game has a variety of interpretations, the best known seems to be the following [1]. Husband (Player 1) and his wife (Player 2) may choose one of two evening entertainments: football match (α_1, β_1) or theatre (α_2, β_2). If they have different desires, (α_1, β_2) or (α_2, β_1), they stay at home. The husband shows preference to the football match, while his wife prefers to go to the theatre. However, it is more important for them to spend the evening together than to be alone at the entertainment (though preferable).

Example 2. "Crossroads" game [2]. Two drivers move along two mutually perpendicular routes and simultaneously meet each other at a crossroad. Each driver may make a stop (1st strategy, α_1 or β_1) or continue on his way (2nd strategy, α_2 or β_2).

It is assumed that each player prefers to make a stop in order to avoid an accident, or to continue on his way if the other player has made a stop. This conflict can be formalized by the bimatrix game with the matrix

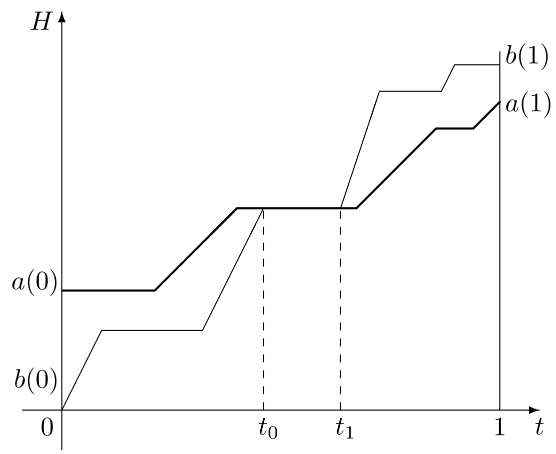
$$(A, B) = \begin{array}{c} \\ \alpha_1 \\ \alpha_2 \end{array} \begin{array}{c} \begin{array}{cc} \quad\beta_1 \quad & \quad \beta_2 \quad \end{array} \\ \begin{bmatrix} (1,1) & (1-\epsilon, 2) \\ (2, 1-\epsilon) & (0,0) \end{bmatrix} \end{array}$$

(the non-negative number ϵ corresponds to the feeling of dissatisfaction that one player has to make a stop and let the other go).

Example 3. Selection of a vehicle for a city tour [2]. Suppose the number of players is large and each of the sets X_i consists of two elements: $X_i = \{0, 1\}$ (for definiteness: 0 is the use of a private vehicle and 1 is the use of a public vehicle). The payoff function is defined as follows:

$$H_i(x_1, \ldots, x_n) = \begin{cases} a(t), & \textit{with } x_i = 1, \\ b(t), & \textit{with } x_i = 0, \end{cases}$$

where $t = 1/n \sum_{j=1}^{n} x_j$.



Plot 1

Let a and b be of the form shown in Plot 1. From the form of the functions $a(t)$ and $b(t)$ it follows that if the number of players choosing 1 is greater than t_1, then the street traffic is light enough to make the driver of a private vehicle more comfortable than the passenger in a public vehicle. However, if the number of motorists is greater than $1 - t_0$, then the traffic becomes so heavy (with the natural priority for public vehicles) that the passenger in a public vehicle compares favourably with the driver of a private vehicle.

Example 4. Allocation of a limited resource taking into account the users' interests. Suppose n users have a good chance of using (accumulating) some resource whose volume is bounded by $A > 0$. Denote by x_i the volume of the resource to be used (accumulated) by the ith user. The users receive a payoff depending on the values of vector $x = (x_1, x_2, \dots, x_n)$. The payoff for the ith user is evaluated by the function $h_i(x_1, x_2, \dots, x_n)$, if the total volume of the used (accumulated) resource does not exceed a given positive value $\Theta < A$, i.e.

$$\sum_{i=1}^{n} x_i \leq \Theta,\ x_i \geq 0.$$

If the inverse inequality is satisfied, the payoff to the ith user is calculated by the function $g_i(x_1, x_2, \dots, x_n)$. Here the resource utility shows a sharp decrease if $\sum_{i=1}^{n} x_i > \Theta$, i.e.

$$g_i(x_1, x_2, \dots, x_n) < h_i(x_1, x_2, \dots, x_n).$$

Consider a nonzero-sum game in normal form

$$\Gamma = (N, \{X_i\}_{i\in N}, \{H_i\}_{i\in N})$$

where the players' payoff functions is

$$H_i(x_1, x_2, \dots, x_n) = \begin{cases} h_i(x_1, \dots, x_n), & \sum_{i=1}^{n} x_i \leq \Theta, \\ g_i(x_1, \dots, x_n), & \sum_{i=1}^{n} x_i > \Theta, \end{cases}$$

$$X_i = [0, a_i],\ 0 \leq a_i \leq A,\ \sum_{i=1}^{n} a_i = A,\ N = \{1, 2, \dots, n\}.$$

The players in this game are the users of the resource.

Example 5. Game-theoretic model for air pollution control. In an industrial area there are n enterprises, each having an emission source. Also, in this area there is an ecologically significant zone Ω whose air pollution must not exceed a maximum permissible level. The time and area-averaged emission from n emitters can be approximately calculated by the formula

$$q = \sum_{i=1}^{n} c_i x_i, \quad 0 \leq x_i \leq a_i,\ i = 1, 2, \dots, n.$$

Let $\Theta < \sum_{i=1}^{n} c_i a_i$ be a maximum emission concentration level.

We shall consider the enterprises to be players and construct the game, modeling an air pollution conflict situation. Suppose each enterprise can reduce its operating expenses by increasing an emission x_i. However, if the air pollution in the area Ω exceeds the maximum emission concentration level, the enterprise incurs a penalty $s_i > 0$.

Suppose player i (enterprise) has an opportunity of choosing the values x_i from the set $X_i = [0, a_i]$. The players' payoff functions are

$$H_i(x_1,\dots,x_n) = \begin{cases} h_i(x_1,x_2,\dots,x_n), & q \le \Theta, \\ h_i(x_1,x_2,\dots,x_n) - s_i, & q > \Theta, \end{cases}$$

where $h_i(x_1,x_2,\dots,x_n)$ are the functions that are continuous and increasing in the variables x_i.

Example 6. Game-theoretic model for bargaining of divisible good [3]. Two players take part in an auction where q units of good with minimal price p_0 are offered. Assumed that players 1,2 have budgets M_1, M_2 respectively. The players demand their quantities of good q_1, q_2 (q_1, q_2, q –integers) and bid their prices p_1, p_2 for unit of the good simultaneously and independently in such a way that

$$q_1 + q_2 > q,\ 0 < q_1 < q,\ 0 < q_2 < q,\ p_1 \in [p_0, \overline{p_1}],\ p_2 \in [p_0, \overline{p_2}],$$

where $\overline{p_1} = M_1/(q-1)$, $\overline{p_2} = M_2/(q-1)$.

According to the bargaining process rules, a player who bids the higher price buys demanded quantity of good at this price. The other buys the rest of good at his own price. If bidden players' prices are equal then Player 1 has an advantage over Player 2. Each player objective is to maximize his profit.

This bargaining process can be described as a nonzero-sum two-person game in normal form $\Gamma = (X, Y, H_1, H_2)$, where sets of the players' strategies are

$$X = \{p_1 | p_1 \in [p_0, \overline{p_1}]\},\ Y = \{p_2 | p_2 \in [p_0, \overline{p_2}]\}$$

and payoff functions are

$$H_1(p_1,p_2) = \begin{cases} (\overline{p_1} - p_1)q_1, & p_1 \ge p_2, \\ (\overline{p_1} - p_1)(q - q_2), & p_1 < p_2, \end{cases}$$

$$H_2(p_1,p_2) = \begin{cases} (\overline{p_2} - p_2)q_2, & p_1 < p_2, \\ (\overline{p_2} - p_2)(q - q_1), & p_1 \ge p_2. \end{cases}$$

2 Optimality principles in noncooperative games

It is well known that for zero-sum games the principles of minimax, maximin and equilibrium coincide (if they are realizable, i.e. there exists an equilibrium, while maximin and minimax are reached and equal to each other). In such a case, they define a unified notion of optimality and game solutions. The theory of nonzero-sum games does not have a unified approach to optimality principles. Although, there are actually many such principles, each of them is based on some additional assumptions of players' behavior and a structure of a game.

It appears natural that each player in the game Γ seeks to reach a situation x in which his payoff function has a maximum value. The payoff function H_i, however, depends not only on the strategy of the ith player, but also on the strategies chosen by the other players. Because of this, the situations $\{x^i\}$ determining a maximum payoff to the ith player may not do the same thing for the other players. As in the case of a zero-sum game, the quest for a maximum payoff involves a conflict, and even formulation of a "good" or optimal behavior in the game becomes highly conjectural. There are many approaches to this problem. One of these is the Nash equilibrium and its various extensions and refinements. When the game Γ is zero-sum, the Nash equilibrium coincides with the notion of optimality (saddle point – equilibrium) that is the basic principle of optimality in a zero-sum game.

Suppose $x = (x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n)$ is an arbitrary situation in the game Γ and x_i is a strategy of player i. We construct a situation that is different from x only in that the strategy x_i of player i has been replaced by a strategy x_i'. As a result we have a situation $(x_1, \dots, x_{i-1}, x_i', x_{i+1}, \dots, x_n)$ denoted by $(x\|x_i')$. Evidently, if x_i and x_i' coincide, then $(x\|x_i') = x$.

Definition 2. *The situation $x^* = (x_1^*, \dots, x_i^*, \dots, x_n^*)$ is called the Nash equilibrium if for all $x_i \in X_i$ and $i = 1, \dots, n$ there is*

$$H_i(x^*) \geq H_i(x^*\|x_i). \tag{2.1}$$

Example 7. Consider the game from Example 3. Here the Nash equilibrium is the situation for which following condition holds

$$t_0 \leq t^* - 1/n, \quad t^* + 1/n \leq t_1 \tag{2.2}$$

where $t^* = 1/n \sum_{j=1}^n x_j^*$. It follows from (2.2) that a payoff to a player remains unaffected when he shifts from one pure strategy to another provided the other players do not change their strategies.

Suppose a play of the game realizes the situation x corresponding to $t = 1/n \sum_{j=1}^n x_j$, $t \in [t_0, t_1]$, and the quantity δ is the share of the players who wish to shift from strategy 0 to strategy 1. Note that if δ is such that $b(t) = a(t) < a(t+\delta)$, then the payoffs to these players tend to increase (with such a strategy shift) provided the strategies of the other players remain unchanged. However, if this shift is actually effected, then the same players may wish to shift from strategy 1 to strategy 0, because the condition $a(t+\delta) < b(t+\delta)$ is satisfied. If this wish is realized, then the share of players, $1/n \sum_{j=1}^n x_j$, decreases and again falls in the interval $[t_0, t_1]$.

Similarly, let δ be the share of players, who decided, for some reason (e.g. because of random errors), to shift from strategy 1 to strategy 0, when $t - \delta < t_0$. Then, by the condition $b(t-\delta) < a(t-\delta)$, the players may wish to shift back to strategy 1. When this wish is realized, the share of the players, $1/n \sum_{j=1}^n x_j$, increases and again comes back to the interval $[t_0, t_1]$.

Example 8. Bertrand Paradox [4]. Assume that two firms produce homogeneous goods and consumers buy from the producer who charges the lowest price. Each firm incurs a cost c per unit of production an always supplies the demand it faces. If the firms charge the same price we assume that each firm faces a demand schedule equal to half of the market demand at the common price. The market demand function is decreasing function $q = d(p)$, $p \geq 0$ and profit function of firm has the form

$$H_1(p_1, p_2) = (p_1 - c)d_1(p_1, p_2),$$

$$H_2(p_1, p_2) = (p_2 - c)d_2(p_1, p_2),$$

where the demand d_i for the output of firm i is given by $(i \neq j)$

$$d_i(p_i, p_j) = \begin{cases} d(p_i), & \text{if } p_i < p_j, \\ d(p_i)/2, & \text{if } p_i = p_j, \\ 0, & \text{if } p_i > p_j, \end{cases}$$

Therefore, the Bertrand duopoly can be described as two-person game $\Gamma_B = (X, Y, H_1, H_2)$, where $X = Y = \{p | p \geq 0\}$.

The Bertrand (1883) paradox states that the unique equilibrium has the two firms charge the equilibrium price: $p_1^* = p_2^* = c$.

The proof is as follows. It is clear, that (c, c) is Nash equilibrium:

$$0 = H_1(c, c) \geq H_1(p_1, c) \quad \text{and}$$

$$0 = H_2(c, c) \geq H_2(c, p_2) \quad \text{for all} \quad p_1 \geq 0, \quad p_2 \geq 0.$$

To prove the uniqueness of equilibrium suppose, for example, that in equilibrium $p_1^* > p_2^* > c$. Then

$$0 = H_1(p_1^*, p_2^*) < H_1(p_2^*, p_2^*) = (p_2^* - c)\frac{d(p_2^*)}{2}$$

and inequality (2.1) is not satisfied.

Therefore, firm 1 cannot be acting in its own best interests if it charges p_1^*. Now suppose that $p_1^* = p_2^* > c$. If firm 1 reduces its price slightly to $p_1^* - \epsilon$, we have

$$(p_1^* - c)\frac{d(p_1^*)}{2} = H_1(p_1^*, p_1^*) < H_1(p_1^* - \epsilon, p_1^*) = (p_1^* - \epsilon - c)d(p_1^* - \epsilon),$$

which is true for small $\epsilon > 0$.

Suppose now that $p_1^* > p_2^* = c$. Then

$$H_2(p_1^*, p_2^*) < H_2(p_1^*, p_2^* + \epsilon) = (p_2^* + \epsilon - c)d(p_2^*)$$

for $0 < \epsilon < p_1^* - p_2^*$ and inequality (2.1) is also not satisfied.

It follows from the definition of Nash equilibrium situation that none of the players i is interested to deviate from the strategy x_i^* appearing in this situation (by (2.1), when such a player uses strategy x_i instead of x_i^*, his payoff may decrease provided the other players follow the strategies generating an equilibrium x^*). Thus, if the players agree on the strategies appearing in the equilibrium x^*, then any individual non-observance of this agreement is disadvantageous to such a player.

For the noncooperative two-person game $\Gamma = (X_1, X_2, H_1, H_2)$ the situation (x^*, y^*) is equilibrium if the inequalities

$$H_1(x, y^*) \leq H_1(x^*, y^*), \quad H_2(x^*, y) \leq H_2(x^*, y^*) \tag{2.3}$$

hold for all $x \in X_1$ and $y \in X_2$

In particular, for the bimatrix $(m \times n)$ game $\Gamma(A, B)$ the pair (i^*, j^*) is the Nash equilibrium if the inequalities

$$\alpha_{ij^*} \leq \alpha_{i^*j^*}, \quad \beta_{i^*j} \leq \beta_{i^*j^*} \tag{2.4}$$

hold for all the rows $i \in M$ and columns $j \in N$. Thus, Example 1 has two equilibria at (α_1, β_1) and (α_2, β_2), whereas Example 2 has equilibria at (α_1, β_2) and (α_2, β_1).

Recall that for the zero-sum game $\Gamma = (X_1, X_2, H)$ the pair $(x^*, y^*) \in X_1 \times X_2$ is an equilibrium if

$$H(x, y^*) \leq H(x^*, y^*) \leq H(x^*, y), \quad x \in X_1,\ y \in X_2.$$

Equilibria in zero-sum games have the following properties:

1^0. A player is not interested to inform the opponent of the strategy (pure or mixed) he wishes to use. Of course, if the player announces in advance of the play, the optimal strategy to be employed, then a payoff to him will not be reduced by the announcement, though he will not win anything.

2^0. Denote by Z(Γ) the set of saddle-points (equilibria) in zero-sum game Γ, then if $(x, y) \in Z(\Gamma)$, $(x', y') \in Z(\Gamma)$ are equilibria in the game Γ, and v is the value of the game, then

$$(x', y) \in Z(\Gamma), \quad (x, y') \in Z(\Gamma), \tag{2.5}$$

$$v = H(x, y) = H(x', y') = H(x, y') = H(x', y). \tag{2.6}$$

3^0. Players are not interested in any intercourse for the purposes of developing joint actions before the game starts.

4^0. If the game Γ has an equilibrium, with x as a maximin and y as a minimax strategy for the players 1 and 2, respectively, then $(x, y) \in Z(\Gamma)$ is an equilibrium, and vice versa.

We shall investigate whether these properties hold for bimatrix games.

Example 9. Consider a "battle of the sexes" game (see example 1). As already noted, this game has two equilibria: (α_1, β_1) and (α_2, β_2). The former is advantageous to Player 1, while the latter is advantageous to Player 2. This contradicts (2.6), since the payoffs to the players in these situations are different. Although the situations (α_1, β_1), (α_2, β_2) are equilibria, the pairs (α_1, β_2) and (α_2, β_1) are not Nash equilibria, i.e. property 2^0 (see (2.5)) is not satisfied.

If Player 1 informs his partner of the strategy α_1 to be employed, and if Player 2 is convinced that he is sure to do it, then he cannot do better than to announce the first strategy β_1. Similar reasoning applies to Player 2. Thus, it is advantageous to each player to announce his strategy, which contradicts property 1^0 for zero-sum games.

Suppose the players establish no contact with each other and make their choices simultaneously and independently (as specified by the rules of a noncooperative game). Let us do the reasoning for Player 1. He is interested in realization of the situation (α_1, β_1), whereas the situation (α_2, β_2) is advantageous to Player 2. Therefore, if Player 1 chooses strategy α_1, then Player 2 can choose strategy β_2, with the result that both players become losers (the payoff vector $(0, 0)$). Then it may be wise of Player 1 to choose strategy α_2, since in the situation (α_2, β_2) he would receive a payoff 1. Player 2, however, may follow a similar line of reasoning and choose β_1, then, in the situation (α_2, β_1) both players again become losers.

Thus, this is the case where the situation is advantageous (but at the same time unstable) to Player 1. Similarly, we may examine the situation (α_2, β_2) (from Player 2's point of view). For this reason, it may be wise of the players to make, in advance of the play, contact and agree on a joint course of action, which contradicts property 3^0. Note that some difficulties may arise when the pairs of maximin strategies do not form an equilibrium.

Thus we have an illustrative example, where none of the properties $1^0 - 4^0$ of a zero-sum game is satisfied.

Payoffs to players may vary with Nash equilibria. Furthermore, unlike the equilibrium set in a zero-sum game, the Nash equilibrium set is not rectangular. If $x = (x_1, \dots, x_i, \dots, x_n)$ and $x' = (x'_1, \dots, x'_i, \dots, x'_n)$ are two different equilibria, then the situation x'' composed of the strategies, which form the situations x and x' and coincides with none of these situations, may not be equilibrium. The Nash equilibrium is a multiple optimality principle in that various equilibria may be preferable to different players to a variable extent. It now remains for us to answer the question: which of the equilibria can be taken as an optimality principle convenient to all players? In what follows it will be shown that the multiplicity of the optimality principle is characteristically and essential feature of an optimal behavior in the controlled conflict processes, with many participants.

Note that, unlike a zero-sum case, the equilibrium strategy x_i^* of the ith player may not always ensure at least the payoff $H_i(x^*)$ in the Nash equilibrium, since this essentially depends on whether the other players choose the strategies appearing in the given Nash equilibrium. For this reason, the equilibrium strategy should not be interpreted as an optimal strategy for the ith player. This interpretation makes sense only for the n-tuples of players' strategies, i.e. for situations.

An important feature of the Nash equilibrium is that any deviation from it made by two or more players may increase a payoff to one of deviating players. Let $S \subset N$ be a subset of the set of players (coalition) and let $x = (x_1, \dots, x_n)$ be a situation in the game Γ. Denote by $(x \| x'_S)$ the situation which is obtained from the situation x by replacing therein the strategies x_i, $i \in S$, with the strategies $x'_i \in X_i$, $i \in S$. In other words, the players appearing in the coalition S replace their strategies x_i by the strategies x'_i. If x^* is the Nash equilibrium, then (2.1) does not necessary imply

$$H_i(x^*) \geq H_i(x^* \| x_S) \quad \text{for all} \quad i \in S. \tag{2.7}$$

In what follows this will be established by some simple examples.

But we may strengthen the notion of a Nash equilibrium by requiring the condition (2.7) or the relaxed condition (2.7) to hold for at least one of the players $i \in S$. Then we arrive at the following definition.

Definition 3. *The situation x^* is called a strong equilibrium if for any coalition $S \subset N$ and $x_S \in \prod_{i \in S} X_i$ there is a player $i_0 \in S$ such that the following inequality is satisfied:*

$$H_{i_0}(x^*) > H_{i_0}(x^* \| x_S). \tag{2.8}$$

Condition (2.8) guarantees that the players' agreement to enter a coalition S is inexpedient because any coalition has a player i_0 who is not interested in this agreement. Any strongly equilibrium situation is a Nash equilibrium.

If the strong equilibrium existed in a broad class of games, then it could be an acceptable principle of optimality in a noncooperative games. However, it happens extremely rare.

Example 10. Prisoners' dilemma. Consider the bimatrix game with payoffs

$$\begin{array}{cc} & \begin{array}{cc} \beta_1 & \quad\beta_2 \end{array} \\ \begin{array}{c} \alpha_1 \\ \alpha_2 \end{array} & \begin{bmatrix} (5,5) & (0,10) \\ (10,0) & (1,1) \end{bmatrix} \end{array}.$$

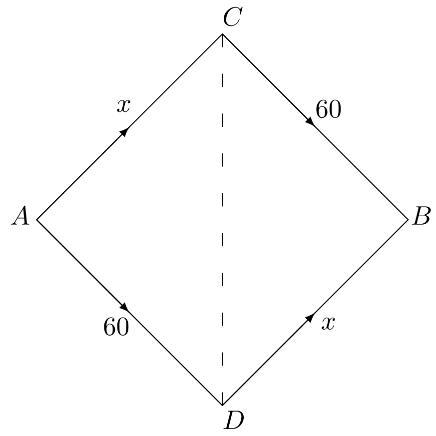
Here we have one equilibrium situation (α_2, β_2) (though not strong equilibrium), which yields the payoff vector $(1, 1)$. However, if both players play (α_1, β_1), they obtain the payoffs $(5, 5)$, which is better to both of them. Zero-sum games have no such paradoxes. As for this particular case, the result is due to the fact that a simultaneous deviation from the equilibrium strategies may further increase the payoff to each player.

Example 11. Braess's paradox [5]. The model was proposed by D. Braess (1968). The road network is shown on Plot 2.

Suppose that 60 cars (players) move from point A to point B. The time for passing from C to B and from A to D equals 60 min. (and does not depend on number of cars on each of arcs AD and CB. On arcs AC and DB the passing time is equal to the number of cars using this arcs. Each player (car) from the set of 60 players (cars) has to go from A to B, and has the possibility to choose one of two roads (strategies) ACB or ADB. It is clear that Nash equilibrium ujhikjh is such allocation of cars in which the time of passing along ACB is equal to the time passing along ADB. If x is the number of cars using ACB and y is the number of cars using ADB we must have in Nash equilibrium

$$60 + x = 60 + y, \quad x + y = 60,$$

which gives us $x = y = 30$ (the proof is clear, since if one car changes his mind and switches, for instance, from road ACB to ADB he will need passing time $60 + 30 + 1 = 91$, but in Nash equilibrium his time is $60 + 30 = 90$).



Plot 2

Suppose now that we connect points C and D by speed way, which each car can pass in time 0. Then we see, that any car, which chooses ACB or ADB will benefit by moving along ACDB spending 60 instead of 90 along ADB or ACB. This means that allocation of cars along ACB and ADB will not be Nash equilibrium after opening a new road CD. It is easily seen, that the new Nash equilibrium in this case will be all cars use ACDB with passing time 120, since if one car deviates she will get the same amount $60 + 60 = 120$.

We observe a paradoxical case – the time passing from A to B increases from 90 to 120 after a new road construction.

Example 10 suggests the possibility of applying other optimality principles to a noncooperative game which may bring about situations that are more advantageous to both players than in the case of equilibrium situations. Such an optimality principle is *pareto-optimality* (optimality Pareto).

Consider a set of vectors $\{H(x)\} = \{H_1(x), \dots, H_n(x)\}$, $x \in X$, $X = \prod_{i=1}^{n} X_i$, i.e. the set of possible values of vector payoffs in all possible situations $x \in X$.

Definition 4. *The situation $\overline{x}$ in the noncooperative game Γ is called pareto-optimal if there is no situation $x \in X$ for which the following inequalities hold:*

$$H_i(x) \geq H_i(\overline{x}) \text{ for all } i \in N \text{ and}$$

$$H_{i_0}(x) > H_{i_0}(\overline{x}) \text{ for at least one } i_0 \in N.$$

The set of all pareto-optimal situations will be denoted by X^P.

The belonging of the situation $\overline{x}$ to the set X^P means that there is no other situation x which might be more preferable to all players than the situation $\overline{x}$.

Following [6], we conceptually distinguish the notion of an equilibrium situation from that of a pareto-optimal situation. In the first case, neither player may individually increase his payoff, while in the second, all the players cannot increase acting as one player (even not strictly) a payoff to each of them.

To be noted also is that the agreement on a fixed equilibrium does not allow each individual player to deviate from it. In the pareto-optimal situation, the deviating player can occasionally obtain an essentially greater payoff. Of course, a strong equilibrium situation is also pareto-optimal. Thus, example 9 provides a situation (α_2, β_2) which is equilibrium, but is not pareto-optimal. Conversely, the situation (α_1, β_1) is pareto-optimal, but not an equilibrium. In the game "battle of the sexes" , both equilibrium situations (α_1, β_1), (α_2, β_2) are strong equilibria and pareto-optimal, but, as already noted in example 8, they are not interchangeable. Similar reasoning also applies to the following example.

Example 12. Consider the "crossroads" game (see Example 2). The situations $(\alpha_2, \beta_1), (\alpha_1, \beta_2)$ form Nash equilibria and are pareto-optimal (the situation (α_1, β_1) is pareto-optimal, but not an equilibrium). For each player the "stop" strategy α_1, β_1 is equilibrium if the other player decides to pass the crossroads and, conversely, it is advantageous for him to choose the "continue" strategy α_2, β_2 if the other player decides to pass the crossroads and, conversely, it is advantageous for him to choose the "continue" strategy α_2, β_2 if the other player makes a stop. However, each player receives a payoff of 2 units only if he chooses the "continue" strategy $\alpha_2(\beta_2)$. This necessarily involves competition for leadership, i.e. each player is interested to be the first to announce the "continue" strategy.

Note that we have reached the same conclusion from examination of the "battle of the sexes" game (see example 9).

We shall now consider behavior of a "leader–follower" type in a two-person game $\Gamma = (X_1, X_2, H_1, H_2)$. Denote by Z^1, Z^2 the sets of best responses for players 1 and 2, respectively, here

$$Z^1 \triangleq \{(x_1, x_2) | H_1(x_1, x_2) = \max_{y_1} H_1(y_1, x_2)\}, \tag{2.9}$$

$$Z^2 \triangleq \{(x_1, x_2) | H_2(x_1, x_2) = \max_{y_2} H_2(x_1, y_2)\} \tag{2.10}$$

(maximum in (2.9) and (2.10) are supposed to be reached, for example, if functions H_i are concave on x_i).

Definition 5. *We call the situation $(x_1, x_2) \in X_1 \times X_2$ the Stakelberg i-equilibrium in the two-person game Γ if*

$$\overline{H}_i \triangleq H_i(x_1, x_2) = \max_{(y_1, y_2) \in Z^j} H_i(y_1, y_2), \tag{2.11}$$

where $i = 1, 2$, $i \neq j$.

The notion of i-equilibrium may be interpreted as follows. Player 1 (Leader) knows the payoff functions of both players H_1, H_2, and hence he learns Player 2's (Follower) set of best responses Z^2 to any strategy x_1 of Player 1. Having this information he then maximizes his payoff by selecting strategy x_1 from condition (2.11). Thus, $\overline{H}_i$ is a payoff to the ith player acting as a "leader" in the game Γ.

Lemma 1. *Let $Z(\Gamma)$ be a set of Nash equilibria in the two-person game Γ. Then*

$$Z(\Gamma) = Z^1 \cap Z^2, \tag{2.12}$$

where Z^1, Z^2 are the sets of the best responses (2.9), (2.10) given by the players 1,2 in the game Γ.

Proof. Let $(x_1, x_2) \in Z(\Gamma)$ be the Nash equilibrium. Then the inequalities

$$H_1(x_1', x_2) \leq H_1(x_1, x_2), \quad H_2(x_1, x_2') \leq H_2(x_1, x_2)$$

hold for all $x_1' \in X_1$ and $x_2' \in X_2$; whence it follows that

$$H_1(x_1, x_2) = \max_{x_1'} H_1(x_1', x_2), \tag{2.13}$$

$$H_2(x_1, x_2) = \max_{x_2'} H_2(x_1, x_2'). \tag{2.14}$$

Thus, $(x_1, x_2) \in Z^1$ and $(x_1, x_2) \in Z^2$, i.e. $(x_1, x_2) \in Z^1 \cap Z^2$.

The inverse inclusion follows immediately from (2.13), (2.14).

If the maximum in (2.9) and (2.10) to be reached in unique point for each x_2 and x_1 respectively, then we can say about reaction functions $R_1(x_2)$, $R_2(x_1)$ and the sets of best responses can rewritten as

$$Z^1 = \{(R_1(x_2), x_2) | x_2 \in X_2\},$$

$$Z^1 = \{(x_1, R_2(x_1)) | x_1 \in X_1\}.$$

This case take place, for example, if payoff functions H_i are strictly concave on x_i.

Example 13. Entry game [4]. Consider a two-firm industry. Firm 1 (the existing firm, leader) chooses a level of capital x_1, which is fixed. Firm 2 (the potential entrant, follower) observes x_1 and then chooses its level of capital x_2.

Assume that payoff functions of the two firms are specified by

$$H_1(x_1, x_2) = x_1(1 - x_1 - x_2),$$

$$H_2(x_1, x_2) = x_2(1 - x_1 - x_2).$$

The game between the two firms is a two-stage game. Firm 1 must foresee the reaction of firm 2 to capital level x_1. Reaction functions $x_2 = R_2(x_1)$, $x_1 = R_1(x_2)$ exists because functions H_i are strictly concave on x_i. Payoff maximization H_2 on x_2 give us first-order condition

$$\frac{\partial H_2}{\partial x_2} = 1 - x_1 - 2x_2 = 0$$

and reaction function
$$x_2 = R_2(x_1) = \frac{1 - x_1}{2}.$$
Therefore, firm 1 maximizes function
$$H_1\left(x_1, \frac{1 - x_1}{2}\right) = x_1\left(1 - x_1 - \frac{1 - x_1}{2}\right)$$
from which we can determine Stakelberg 1-equilibrium:
$$x_1^1 = \frac{1}{2}, \quad x_2^1 = \frac{1}{4}, \quad H_1^1 = \frac{1}{8}, \quad H_2^1 = \frac{1}{16}.$$
If the two firms choose their strategies simultaneously, each would react to other optimally, so that $x_2 = R_2(x_1)$ and $x_1 = R_1(x_2)$, where
$$x_2 = R_2(x_1) = \frac{1 - x_1}{2}, \quad x_1 = R_1(x_2) = \frac{1 - x_2}{2}.$$
Solving the equations we receive the Nash equilibrium $x_1^* = x_2^* = \frac{1}{3}$, $H_1^* = H_2^* = \frac{1}{9}$.

Comparing the Nash and Stakelberg equilibrium we see the leader's advantage.

Definition 6. *[2] We say that the two-person game $\Gamma = (X_1, X_2, H_1, H_2)$ involves competition for leadership if there exists a situation $(x_1, x_2) \in X_1 \times X_2$ such that*
$$\overline{H}_i \leq H_i(x_1, x_2), \quad i = 1, 2. \tag{2.15}$$

Theorem 1. *[2] If the two-person game $\Gamma = (X_1, X_2, H_1, H_2)$ has at least two pareto-optimal and Nash equilibrium situations (x_1, x_2), (y_1, y_2) with different payoff vectors*
$$(H_1(x_1, x_2), H_2(x_1, x_2)) \neq (H_1(y_1, y_2), H_2(y_1, y_2)), \tag{2.16}$$
then the game Γ involves competition for leadership.

Proof. By (2.12), for any Nash equilibrium $(z_1, z_2) \in Z(\Gamma)$ we have
$$H_i(z_1, z_2) \leq \overline{H}_i, \quad i = 1, 2.$$
Suppose the opposite is true, i.e. the game Γ does not involve competition for leadership. Then there is a situation $(z_1, z_2) \in X_1 \times X_2$ for which
$$H_i(x_1, x_2) \leq \overline{H}_i \leq H_i(z_1, z_2), \tag{2.17}$$
$$H_i(y_1, y_2) \leq \overline{H}_i \leq H_i(z_1, z_2), \tag{2.18}$$
$i = 1, 2$. But (x_1, x_2), (y_1, y_2) are pareto-optimal situations, and hence the inequalities (2.17), (2.18) are satisfied as equalities, which contradicts (2.16). This completes the proof of the theorem.

In conclusion we may say that the games "battle of the sexes" and "crossroads" satisfy the condition of the theorem and hence involve competition for leadership.

Example 14. Auction of an indivisible goods [7]. A seller has one indivisible unit of an object for sale with reservation price c. There are n potential bidders with valuations

$$c \leq v_n \leq ... \leq v_2 \leq v_1$$

and these valuations and the reservation price are common knowledge. The bidders simultaneously submit bids $x_i \geq c$. The highest bidder wins the object.
a) *First-price auction.* In this case the winner pays the highest bid. Set of strategies for each bidder is $X_i = X = [c, +\infty)$. Suppose that profile $x = (x_1, ..., x_n)$ is realized. Denote the set of highest bidders as

$$w(x) = \{i | x_i = \max_j x_j\}.$$

Then the payoff function for bidder i have the following form

$$H_i(x_1, ..., x_i, ..., x_n) = \begin{cases} v_i - x_i, & i = \min_{j \in w(x)} j, \\ 0, & i \neq \min_{j \in w(x)} j. \end{cases}$$

b) *Second-price auction.* In second-price auction winner i pays the second bid, i.e.

$$x_i^+ = \max_{j \neq i} x_j.$$

Then the payoff function for bidder i is

$$H_i(x_1, ..., x_i, ..., x_n) = \begin{cases} v_i - x_i^+, & i = \min_{j \in w(x)} j, \\ 0, & i \neq \min_{j \in w(x)} j. \end{cases}$$

There are many Nash equilibria in the models. The structure of Nash equilibrium is the following: $x^* = (v_2, v_2, x_3, ..., x_n)$, where $x_i \in [c, v_i]$, $i = 3, 4, ..., n$.

But if bidders' valuations are not common knowledge and bidder i knows its own valuation, then the first-price auction does not have Nash equilibrium in pure strategies. In this case second-price auction has Nash equilibrium in truthful strategies: $x^* = (v_1, v_2, ..., v_n)$. However the Nash equilibrium is paretodominated in both cases.

3 Mixed extension of noncooperative game

We shall examine a noncooperative two-person game $\Gamma = (X_1, X_2, H_1, H_2)$. In the nonzero-sum case, we have already seen that an equilibrium in pure strategies generally does not exist. The matrix games have an equilibrium in mixed strategies. For this reason, it would appear natural that the Nash equilibrium in a noncooperative game would be sought in the class of mixed strategies.

As in the case of zero-sum games, we identify the player's mixed strategy with the probability distribution over the set of pure strategies. For simplicity, we assume that the sets of strategies X_i are finite, and introduce the notion of a mixed extension of the game. Let

$$\Gamma = (N, \{X_i\}_{i\in N}, \{H_i\}_{i\in N}) \tag{3.1}$$

be an arbitrary finite noncooperative game. For definitiveness, suppose the player i in the game Γ has m_i strategies.

Denote by μ_i an arbitrary mixed strategy of the player i, i.e. some probability distribution over the set of strategies X_i to be referred to as pure strategies. Also, denote by $\mu_i(x_i)$ the probability prescribed by strategy μ_i to the particular pure strategy $x_i \in X_i$. The set of all mixed strategies of player i will be denoted by $\overline{X}_i$.

Suppose each player $i \in N$ uses his mixed strategy μ_i, i.e. he chooses pure strategies with probabilities $\mu_i(x_i)$. The probability that a situation $x = (x_1, \ldots, x_n)$ may arise is equal to the product of the probabilities of choosing its component strategies, i.e.

$$\mu(x) = \mu_1(x_1) \times \mu_2(x_2) \times \ldots \times \mu_n(x_n). \tag{3.2}$$

Formula (3.2) defines the probability distribution over the set of all situations $X = \prod_{i=1}^n X_i$ determined by mixed strategies $\mu_1, \mu_2, \ldots, \mu_n$. The n-tuple $\mu = (\mu_1, \ldots, \mu_n)$ is called a situation in mixed strategies. The situation in mixed strategies μ realizes various situations in pure strategies with some probabilities; hence the value of the payoff function for each player turns out to be a random variable. The value of the payoff function for the ith player in the situation μ is taken to be the mathematical expectation of this random variable:

$$K_i(\mu) \triangleq \sum_{x\in X} H_i(x)\mu(x)$$

$$= \sum_{x_1\in X_1} \ldots \sum_{x_n\in X_n} H_i(x_1, \ldots, x_n) \times \mu_1(x_1) \times \ldots \times \mu_n(x_n),$$

$$i \in N,\ x = (x_1, \ldots, x_n) \in X. \tag{3.3}$$

We introduce the notation

$$K_i(\mu\|x'_j) \triangleq \sum_{x_1\in X_1} \ldots \sum_{x_{j-1}\in X_{j-1}} \sum_{x_{j+1}\in X_{j+1}} \ldots \sum_{x_n\in X_n} H_i(x\|x'_j) \prod_{k\neq j} \mu_k(x_k). \tag{3.4}$$

Let μ'_j be an arbitrary mixed strategy for player j in the game Γ. Multiplying (3.4) by $\mu'(x'_j)$ and summing over all $x'_j \in X_j$, we obtain

$$\sum_{x'_j\in X_j} K_i(\mu\|x'_j)\mu'_j(x'_j) \triangleq K_i(\mu\|\mu'_j).$$

Definition 7. *The game $\overline{\Gamma} = (N, \{\overline{X}_i\}_{i\in N}\{K_i\}_{i\in N})$, in which N is the set of players, $\overline{X}_i$ is the set of mixed strategies of each player i, and the payoff function is defined by (3.3), is called a mixed extension of the game Γ.*

If the inequality $K_j(\mu\|x_i) \le a$ holds for any pure strategy x_i of player i, then the inequality $K_j(\mu\|\mu_i^*) \le a$ holds for any mixed strategy μ_i^*. The truth of this assertion follows from (3.3) and (3.4) by a standard shift to mixed strategies.

For the bimatrix $(m \times n)$ game $\Gamma(A, B)$ we may define the respective sets of mixed strategies X_1, X_2 for players 1 and 2 as

$$X_1 = \{x \mid xu = 1,\ x \ge 0,\ x \in R^m\},$$

$$X_2 = \{y \mid yw = 1,\ y \ge 0,\ y \in R^n\},$$

where $u = (1, \dots, 1) \in R^m$, $w = (1, \dots, 1) \in R^n$. We also define the players' payoffs K_1 and K_2 at (x, y) in mixed strategies as payoff expectations

$$K_1(x, y) \triangleq xAy, \quad K_2(x, y) \triangleq xBy, \quad x \in X_1,\ y \in X_2.$$

Thus, we have constructed formally a mixed extension $\overline{\Gamma}(A, B)$ of the game $\Gamma(A, B)$, i.e. the noncooperative two-person game $\overline{\Gamma}(A, B) = (X_1, X_2, K_1, K_2)$.

For the bimatrix game (just as for the matrix game) the set $M_x = \{i|\xi_i > 0\}$ will be called Player 1's *spectrum of mixed strategy* $x = (\xi_1, \dots, \xi_m)$, while the strategy x, for which $M_x = M$, $M = \{1, 2, \dots, m\}$, will be referred to as *completely mixed*. Similarly, $N_y = \{j|\eta_j > 0\}$ will be Player 2's spectrum of mixed strategy $y = \{\eta_1, \dots, \eta_n\}$ in the bimatrix $(m \times n)$ game $\Gamma(A, B)$. The situation (x, y), in which both strategies x and y are completely mixed, will be referred to as *completely mixed*.

We shall now use the "battle of the sexes" game to demonstrate that the difficulties encountered in examination of a noncooperative game (Example 9) are not resolved through introduction of mixed strategies.

Example 15. Suppose Player 1 in the "battle of the sexes" game wishes to maximize his guaranteed payoff. This means that he is going to choose a mixed strategy $x^0 = (\xi^0, 1 - \xi^0)$, $0 \le \xi^0 \le 1$ so as to maximize the least of the two quantities $K_1(x, \beta_1)$ and $K_1(x, \beta_2)$, i.e.

$$\max_x \min\{K_1(x, \beta_1), K_1(x, \beta_2)\} = \min\{K_1(x^0, \beta_1), K_1(x^0, \beta_2)\}.$$

The maximin strategy x^0 of Player 1 is of the form $x^0 = (1/5, 4/5)$ and guarantees him, on the average, a payoff of 4/5. If Player 2 chooses strategy β_1, then the players' payoffs are $(4/5, 1/5)$. However, if he uses strategy β_2, then the players' payoffs are $(4/5, 16/5)$.

Thus, if Player 2 suspects that his partner will use the strategy x^0, then he will choose β_2 and receive a payoff of 16/5. (If Player 1 can justify the choice of β_2 for Player 2, then he may also improve his own choice.) Similarly, suppose Player 2 uses a maximin strategy that is $y^0 = (4/5, 1/5)$. If Player 1 chooses strategy α_1 then the players' payoffs are $(16/5, 4/5)$. However, if he chooses α_2, then the players' payoffs are $(1/5, 4/5)$. Therefore, it is advantageous for him to use his strategy α_1 against the maximin strategy y^0.

If both players follow this line of reasoning, they will arrive at a situation (α_1, β_2), in which the payoff vector is $(0, 0)$. Hence the situation (x^0, y^0) in maximin mixed strategies is not a Nash equilibrium.

Definition 8. *The situation μ^* is called a Nash equilibrium in mixed strategies in the game Γ if for any player i, and for any mixed strategies μ_i the following inequality holds:*

$$K_i(\mu^* \| \mu_i) \le K_i(\mu^*),\ i = 1, \dots, n.$$

Example 15 shows that a situation in maximin mixed strategies is not necessarily a Nash equilibrium in mixed strategies.

Example 16. The game of "crossroads" (see Example 12) has two Nash equilibria in pure strategies: (α_1, β_2) and (α_2, β_1). These situations are also pareto-optimal. The mixed extension of the game gives rise to one more equilibrium situation, namely the pair (x^*, y^*):

$$x^* = y^* = \frac{1-\epsilon}{2-\epsilon} u_1 + \frac{1}{2-\epsilon} u_2,$$

where $u_1 = (1, 0)$, $u_2 = (0, 1)$ or $x^* = y^* = ((1-\epsilon)/(2-\epsilon), 1/(2-\epsilon))$.

Indeed, we have

$$K_1(\alpha_1, y^*) = \frac{1-\epsilon}{2-\epsilon} + \frac{1-\epsilon}{2-\epsilon} = 1 - \frac{\epsilon}{2-\epsilon},$$

$$K_1(\alpha_2, y^*) = 2\frac{1-\epsilon}{2-\epsilon} = 1 - \frac{\epsilon}{2-\epsilon}.$$

Furthermore, since for any pair of mixed strategies $x = (\xi, 1-\xi)$ and $y = (\eta, 1-\eta)$, we have

$$K_1(x, y^*) = \xi K_1(\alpha_1, y^*) + (1-\xi) K_1(\alpha_2, y^*) = 1 - \frac{\epsilon}{2-\epsilon},$$

$$K_2(x^*, y) = \eta K_2(x^*, \beta_1) + (1-\eta) K_2(x^*, \beta_2) = 1 - \frac{\epsilon}{2-\epsilon},$$

then we get

$$K_1(x, y^*) = K_1(x^*, y^*),\ K_2(x^*, y) = K_2(x^*, y^*)$$

for all mixed strategies $x \in X_1$ and $y \in X_2$. Therefore, (x^*, y^*) is a Nash equilibrium. Furthermore, it is a completely mixed equilibrium. But the situation (x^*, y^*) is not pareto-optimal, since the vector $K(x^*, y^*) = (1 - \epsilon/(2-\epsilon), 1 - \epsilon/(2-\epsilon))$ is strictly (component-wise) smaller than the payoff vector $(1, 1)$ in the situation (α_1, β_1).

Let $K(\mu^*) = \{K_i(\mu^*)\}$ be a payoff vector in some Nash equilibrium. Denote $v_i = K_i(\mu^*)$ and $v = \{v_i\}$. While the zero-sum games have the same value v of the payoff function in all equilibrium points and hence this value was uniquely defined for each zero-sum game, which had such an equilibrium, in the nonzero-sum games there is a whole set of vectors v. Thus every vector v is connected with a special equilibrium point μ^*, $v_i = K_i(\mu^*)$, $\mu^* \in \overline{X}$, $\overline{X} = \prod_{i=1}^n \overline{X}_i$.

In the game of "crossroads", the equilibrium payoff vector (v_1, v_2) at the equilibrium point (α_1, β_2) is of the form $(1-\epsilon, 2)$, whereas at (x^*, y^*) it is equal to $(1 - \epsilon/(2-\epsilon), 1 - \epsilon/(2-\epsilon))$ (see Example 12).

If the strategy spaces in the noncooperative game $\Gamma = (X_1, X_2, H_1, H_2)$ are infinite, e.g. $X_1 \subset R^m$, $X_2 \subset R^n$, then as in the case of zero-sum infinite game, the mixed strategies of the players are identified with the probability measures given on Borel σ-algebras of the sets X_1 and X_2. If μ and ν are respectively the mixed strategies of Players 1 and 2, then a payoff to player i in this situation $K_i(\mu, \nu)$ is the mathematical expectation of payoff, i.e.

$$K_i(\mu, \nu) = \int_{X_1}\int_{X_2} H_i(x, y) d\mu(x) d\nu(y), \tag{3.5}$$

where the integrals are taken to be Lebesgue-Stieltjes integrals. Note that, the payoffs to the players at (x, ν) and (μ, y) are

$$K_i(x, \nu) = \int_{X_2} H_i(x, y) d\nu(y),$$

$$K_i(\mu, y) = \int_{X_1} H_i(x, y) d\mu(x), \; i = 1, 2.$$

(All integrals are assumed to exist.)

Formally, the mixed extension of the noncooperative two-person game Γ can be defined as a system $\overline{\Gamma} = (\overline{X}_1, \overline{X}_2, K_1, K_2)$, where $\overline{X}_1 = \{\mu\}$, $\overline{X}_2 = \{\nu\}$ with K_1 and K_2 determined by (3.5). The game $\overline{\Gamma}$ is a noncooperative two-person game, and hence the situation (μ^*, ν^*) is equilibrium if and only if the inequalities (as in (2.3)) are satisfied.

4 Existence of Nash equilibrium

In the theory of zero-sum games, the continuity of 1 payoff function and the compactness of strategy sets is sufficient for the existence of an equilibrium in mixed strategies. It turns out that these conditions also suffice for the existence of a Nash equilibrium in mixed strategies where a noncooperative two-person game is concerned.

First we prove the existence of an equilibrium in mixed strategies for a bimatrix game. This proof is based on the familiar Kakutani's fixed point theorem. This theorem will be given without proof.

Theorem 2. *Let S be a convex compact set in R^n and ψ be a multi-valued map which corresponds to each point of S the convex compact subsets of S and satisfies the condition: if $x_n \in S$, $x_n \to x$, $y_n \in \psi(x_n)$, $y_n \to y$ and $y \in \psi(x)$. Then there exists $x^* \in S$ such that $x^* \in \psi(x^*)$.*

Theorem 3. *Let $\Gamma(A, B)$ be a bimatrix $(m \times n)$ game. Then there are mixed strategies $x^* \in X_1$ and $y^* \in X_2$ for Players 1 and 2, respectively, such that the pair (x^*, y^*) is a Nash equilibrium.*

Proof. The mixed strategy sets X_1 and X_2 of players 1 and 2 are convex polyhedra. Hence the set of situations $X_1 \times X_2$ is a convex compact set.

Let ψ be a multi-valued map,

$$\psi : X_1 \times X_2 \to X_1 \times X_2,$$

determined by the relationship

$$\psi : (x_0, y_0) \to \left\{ \quad (x', y') \;\middle|\; \begin{array}{l} K_1(x', y_0) = \max_{X_1} K_1(x, y_0), \\ K_2(x_0, y') = \max_{X_2} K_2(x_0, y), \end{array} \right.$$

i.e. the image of the map ψ consists of the pairs of the players' best responses to the strategies y_0 and x_0, respectively.

The functions K_1 and K_2 as the mathematical expectations of the payoffs in the situation (x, y) are bilinear in x and y, and hence the image $\psi(x_0, y_0)$ of the situation (x_0, y_0) with ψ as a map represents a convex compact subset in $X_1 \times X_2$. Furthermore, if the sequence of pairs $\{(x_0^n, y_0^n)\}$, $(x_0^n, y_0^n) \in X_1 \times X_2$ and $\{(x'_n, y'_n)\}$, $(x'_n, y'_n) \in \psi(x_0^n, y_0^n)$ have limit points, i.e.

$$\lim_{n\to\infty} (x_0^n, y_0^n) = (x_0, y_0), \;\; \lim_{n\to\infty} (x'_n, y'_n) = (x', y'),$$

then by the bilinearity of the functions K_1 and K_2, and because of the compactness of the sets X_1 and X_2, we have $(x', y') \in \psi(x_0, y_0)$. Then, by the Kakutani's theorem, there exists a situation $(x^*, y^*) \in X_1 \times X_2$ for which $(x^*, y^*) \in \psi(x^*, y^*)$, i.e.

$$K_1(x^*, y^*) \geq K_1(x, y^*), \;\; K_2(x^*, y^*) \geq K_2(x^*, y)$$

for all $x \in X_1$ and $y \in X_2$. This completes the proof of the theorem.

The preceding theorem can be extended to the case of continuous payoff functions H_1 and H_2. To prove this result, we have to use the well-known Brouwer fixed point theorem [8].

Theorem 4. *Let S be a convex compact set in R^n which has an interior. If φ is a continuous self-map of S, then there exists a fixed point x^* of the map φ, i.e. $x^* \in S$ and $x^* = \varphi(x^*)$.*

Theorem 5. *Let $\Gamma = (X_1, X_2, H_1, H_2)$ be a noncooperative two-person game, where the strategy spaces $X_1 \subset R^m$, $X_2 \subset R^n$ are convex compact subsets and the set $X_1 \times X_2$ has an interior. Also, let the payoff functions $H_1(x, y)$ and $H_2(x, y)$ be continuous in $X_1 \times X_2$, with $H_1(x, y)$ being concave in x at every fixed y and $H_2(x, y)$ being concave in y at every fixed x.*

Then the game Γ has the Nash equilibrium (x^*, y^*).

Proof. Let $p \triangleq (x, y) \in X_1 \times X_2$ and $q \triangleq (\overline{x}, \overline{y}) \in X_1 \times X_2$ be two situations in the game Γ. Consider the function

$$\theta(p, q) = H_1(x, \overline{y}) + H_2(\overline{x}, y).$$

First we show that there exists a situation $q^* = (x^*, y^*)$ for which

$$\max_{p \in X_1 \times X_2} \theta(p, q^*) = \theta(q^*, q^*).$$

Suppose this is not the case. Then for each $q \in X_1 \times X_2$ there is a $p \in X_1 \times X_2$, $p \neq q$, such that $\theta(p,q) > \theta(q,q)$. Introduce the set

$$G_p = \{q | \theta(p,q) > \theta(q,q)\}.$$

Since the function θ is continuous (H_1 and H_2 are continuous in all their variables) and $X_1 \times X_2$ is a convex compact set, then the sets G_p are open. Furthermore, by the assumptions, $X_1 \times X_2$ is covered by the sets from the family G_p.

It follows from the compactness of $X_1 \times X_2$ that there is a finite collection of these sets which covers $X_1 \times X_2$. Suppose these are the sets $G_{p_1}, \ldots, G_{p_k}$. Denote

$$\varphi_j(q) = \max\{\theta(p_j, q) - \theta(q,q), 0\}.$$

The functions $\varphi_j(q)$ are non-negative, and, by the definition of G_{p_j}, at least one of the functions φ_j takes a positive value at every point q.

We shall now define the self-map ψ of the set $X_1 \times X_2$ as follows:

$$\psi(q) \triangleq \frac{1}{\varphi(q)} \sum_j \varphi_j(q) p_j,$$

where $\varphi(q) = \sum_j \varphi_j(q)$. The functions φ_j are continuous and hence ψ is a continuous self-map of $X_1 \times X_2$. By the Brouwer's fixed point theorem, there is a point $\overline{q} \in X_1 \times X_2$ such that $\psi(\overline{q}) = \overline{q}$, i.e.

$$\overline{q} = (1/\varphi(\overline{q})) \sum_j \varphi_j(\overline{q}) p_j.$$

Consequently,

$$\theta(\overline{q}, \overline{q}) = \theta\left(\frac{1}{\varphi(\overline{q})} \sum_j \varphi_j(\overline{q}) p_j, \overline{q}\right).$$

But the function $\theta(p,q)$ is concave in p, with q fixed, and hence

$$\theta(\overline{q}, \overline{q}) \geq \frac{1}{\varphi(\overline{q})} \sum_j \varphi_j(\overline{q}) \theta(p_j, \overline{q}). \tag{4.1}$$

On the other hand, if $\varphi_j(\overline{q}) > 0$, then $\theta(\overline{q}, \overline{q}) < \theta(p_j, \overline{q})$, and if $\varphi_j(\overline{q}) = 0$, then $\varphi_j(\overline{q})\theta(p_j, \overline{q}) = \varphi_j(\overline{q})\theta(\overline{q}, \overline{q})$. Since $\varphi_j(\overline{q}) > 0$ for some j, we get the inequality

$$\theta(\overline{q}, \overline{q}) < \frac{1}{\varphi(\overline{q})} \sum_j \varphi_j(\overline{q}) \theta(p_j, \overline{q}),$$

which contradicts (4.1).

Thus, there always exists q^* for which

$$\max_{p \in X_1 \times X_2} \theta(p, q^*) = \theta(q^*, q^*).$$

Which means that

$$H_1(x, y^*) + H_2(x^*, y) \le H_1(x^*, y^*) + H_2(x^*, y^*)$$

for all $x \in X_1$ and $y \in X_2$. Setting successively $x = x^*$ and $y = y^*$ in the last inequality, we obtain the inequalities

$$H_2(x^*, y) \le H_2(x^*, y^*), \quad H_1(x, y^*) \le H_1(x^*, y^*),$$

which hold for all $x \in X_1$ and $y \in X_2$. This completes the proof of the theorem.

The result given below holds for the noncooperative two-person games played on compact sets (specifically, on a unit square) with a continuous payoff function.

Example 17. Cournot duopoly [4]. Consider the Cournot model of duopoly producing a homogeneous good. The strategies are quantities. Firm 1 and firm 2 simultaneously choose their strategies q_i from feasible sets $X = Y = \{q | q \ge 0\}$. Suppose that market demand function $q = d(c)$, $p \ge 0$ is strictly decreasing. Then exist strictly decreasing inverse function

$$p = p(q) = d^{-1}(q), \quad q \ge 0.$$

Firms sell their good at the market-clearing price $p(q)$, where $q = q_1 + q_2$. Firm i's cost of production is increasing function $c_i(q_i)$ and firm i's total profit is

$$H_1(q_1, q_2) = q_1 p(q) - c_1(q_1),$$

$$H_2(q_1, q_2) = q_2 p(q) - c_2(q_2).$$

The Cournot duopoly can be described as nonzero-sum two person game in normal form $\Gamma_C = (X, Y, H_1, H_2)$.

The existence of a pure-strategy Nash equilibrium in Cournot duopoly follows from concavity profit functions H_i on q_i. Calculate partial derivatives

$$\frac{\partial H_i}{\partial q_i} = p(q_i + q_j) - c_i'(q_i) + q_i p'(q_i + q_j)$$

and

$$\frac{\partial^2 H_i}{\partial q_i^2} = 2p' + q_i p'' - c_i''.$$

Recall that $p' < 0$. For the payoff function to be concave $\left(\frac{\partial^2 H_i}{\partial q_i^2} < 0\right)$, it suffices that the firm's cost function be convex ($c'' \ge 0$) and that the inverse-demand function be concave ($p'' \ge 0$).

The Nash equilibrium is easily derived in case of linear demand and cost. Suppose that $d(p) = a - p$ (or $p(q) = a - q$) and $c_i(q_i) = c_i q_i$, $c_i > 0$. Let us consider the first-order condition to find Nash equilibrium

$$\begin{cases} \frac{\partial H_1}{\partial q_1} = a - 2q_1 - q_2 - c_1 = 0, \\ \frac{\partial H_2}{\partial q_2} = a - q_1 - 2q_2 - c_2 = 0. \end{cases}$$

Hence, The Nash equilibrium is given by

$$q_1^* = \frac{a - 2c_1 + c_2}{3}, \quad q_2^* = \frac{a + c_1 - 2c_2}{3}$$

and the profits are

$$\Pi_1^* = \frac{(a - 2c_1 + c_2)^2}{9}, \quad \Pi_2^* = \frac{(a + c_1 - 2c_2)^2}{9}.$$

Example 18. Hotelling competition [9]. A city of length 1 lies on the line, and consumers are uniformly distributed with density 1 along the interval $[0, 1]$ on this line. There are two stores located at the two extremes of city, which sell the same product. p_1, p_2 are prices of the unit of good in stores 1 and 2 correspondingly. Suppose that store 1 is at $x = 0$ and store 2 at $x = 1$.

Consumers incur a transportation cost t per unit of distance and have unit demands.

A consumer who is indifferent between the two stores is located at x, where x can be calculated from the equation

$$p_1 + tx = p_2 + t(1 - x).$$

The demand for store 1 is equal to the number of consumers who find it cheaper to buy from store 1 and is given by

$$d_1(p_1, p_2) = x = \frac{p_2 - p_1 + t}{2t}$$

and

$$d_2(p_1, p_2) = 1 - x = \frac{p_1 - p_2 + t}{2t}.$$

Firm's profits are

$$H_1(p_1, p_2) = (p_1 - c)\frac{(p_2 - p_1 + t)}{2t},$$

$$H_2(p_1, p_2) = (p_2 - c)\frac{(p_1 - p_2 + t)}{2t}.$$

This game is a strictly concave, that is why we can calculate thee Nash equilibrium from first-order conditions

$$\begin{cases} p_2 + c + t - 2p_1 = 0, \\ p_1 + c + t - 2p_2 = 0. \end{cases}$$

Using the symmetry of the problem, we obtain the solution:

$$p_1^* = p_2^* = c + t, \quad \Pi_1^* = \Pi_2^* = \frac{t}{2} > 0.$$

Theorem 6. *Let $\Gamma = (X_1, X_2, H_1, H_2)$ be a noncooperative two-person game, where H_1 and H_2 are continuous functions on $X_1 \times X_2$; X_1, X_2 are compact subsets in finite-dimensional Euclidean spaces. Then the game Γ has an equilibrium (μ, ν) in mixed strategies.*

This theorem is given without proof, since it is based on the continuity and bilinearity of the function

$$K_i(\mu, \nu) = \int_{X_1}\int_{X_2} H_i(x, y) d\mu(x) d\nu(y), \quad i = 1, 2.$$

over the set $\overline{X}_1 \times \overline{X}_2$ and almost exactly repeats the proof of the preceding theorem.

We shall discuss in more detail the construction of mixed strategies in noncooperative n-person games with an infinite number of strategies. Note that if the players' payoff functions $H_i(x)$ are continuous on the Cartesian product $X = \prod_{i=1}^n X_i$ of the compact sets of pure strategies, then in such a noncooperative game there always exists a Nash equilibrium in mixed strategies. As for the existence of pareto-optimal situations, it suffices to ensure the compactness of the set $\{H(x)\}$, $x \in X$, which in turn can be ensured by the compactness in some topology of the set of all situations X and the continuity in this topology of all the payoff functions K_i, $i = 1, 2, \dots, n$. It is evident that this is always true for finite noncooperative games.

5 Kakutani fixed-point theorem and proof of existence of an equilibrium in n-person games

The reader can read Section 5 without referring to the previous Section 4. Given any game $\Gamma = \langle N, \{X_i\}_{i\in N}, \{H_i\}_{i\in N}\rangle$ with finite sets of strategies X_i in normal form ($|N| = n$), a mixed strategy for any player i is a probability distribution over X_i. We let $\overline{X}_i$ denote the set of all possible mixed strategies for player i. To underline the distinction from mixed strategies, the strategies in X_i will be called pure strategies.

A mixed strategy profile is any vector that specifies one mixed strategy for each player, so the set of all mixed strategy profiles (situations in mixed strategies) is a Cartesian product $\overline{X} = \prod_{i=1}^n \overline{X}_i$. $\mu = (\mu_1, \dots, \mu_n)$ is a mixed-strategy profile in $\prod_{i=1}^n \overline{X}_i$ if and only if, for each player i and each pure strategy $x_i \in X_i$, μ prescribes a non-negative real number $\mu_i(x_i)$, representing the probability that player i would choose x_i, such that

$$\sum_{x_i \in X_i} \mu_i(x_i) = 1, \text{ for all } i \in N.$$

If the players choose their pure strategies independently, according to the mixed strategy profile μ, then the probability, that they will choose the pure strategy profile $x = (x_1, \dots, x_i, \dots, x_n)$ is $\prod_{i=1}^n \mu_i(x_i)$, the multiplicative product of the individual strategy probabilities.

For any mixed strategy profile μ, let $K_i(\mu)$ denote the mathematical expectation of payoff that player i would get when the players independently choose their pure strategies according to μ. Denote $X = \prod_{i=1}^n X_i$ (X_i is the set of all possible situation in pure strategies), then

$$K_i(\mu) = \sum_{x \in X} \Big(\prod_{j \in N} \mu_j(x_j) \Big) H_i(x), \text{ for all } i \in N.$$

For any $\tau_i \in \overline{X}_i$, we denote by $(\mu \| \tau_i)$ the mixed strategy profile in which the i-component is τ_i and all other components are as in μ. Thus

$$K_i(\mu \| \tau_i) = \sum_{x \in X} \Big(\prod_{j \in N \setminus \{i\}} \mu_j(x_j) \Big) \tau_i(x_i) H_i(x).$$

We shall not use any special notation for the mixed strategy $\mu_i \in \sum_i$ that puts probability 1 on the pure strategy x_i, denoting this mixed strategy by x_i (in the same manner as the corresponding pure strategy).

If player i used the pure strategy x_i, while all other players behaved independently according to the mixed-strategy profile μ, then player i's mathematical expectation of payoff would be

$$K_i(\mu \| x_i) = \sum_{x_j \in X_j,\ j \neq i} \Big(\prod_{j \in N \setminus \{i\}} \mu_j(x_j) \Big) H_i(\mu \| x_i).$$

Definition 9. *The mixed strategy profile μ is Nash equilibrium in mixed strategies if*

$$K_i(\mu \| \tau_i) \leq K_i(\mu), \text{ for all } \tau_i \in \overline{X}_i,\ i \in N.$$

Lemma 2. *For any $\mu \in \prod_{i=1}^n \overline{X}_i$ and any player $i \in N$,*

$$\max_{x_i \in X_i} K_i(\mu \| x_i) = \max_{\tau_i \in \overline{X}_i} K_i(\mu_i \| \tau_i).$$

Furthermore, $p_i \in \arg\max_{\tau_i \in \overline{X}_i} K_i(\mu \| \tau_i)$ if and only if $p_i(x_i) = 0$ for every x_i such that $x_i \notin \arg\max_{x_i \in X_i} K_i(\mu \| x_i)$.

Proof. Notice that for any $\tau_i \in \overline{X}_i$

$$K_i(\mu \| \tau_i) = \sum_{x_i \in X_i} \tau_i(x_i) K_i(\mu \| x_i).$$

$K_i(\mu \| \tau_i)$ is a mathematical expectation of terms $K_i(\mu \| x_i)$. This mathematical expectation cannot be greater than the maximum value of the random variable $K_i(\mu \| x_i)$, and is strictly less than this maximum value, whenever any nonmaximal value of $K_i(\mu \| x_i)$ gets a positive probability ($\tau_i(x_i) \geq 0$, $\sum_{x_i \in X_i} \tau_i(x_i) = 1$).

So the highest expected payoff that player i can get against any combination of other player's mixed strategies is the same whether he uses a mixed strategy or not.

As we have seen in the two-person case, the Kakutani fixed-point theorem is a useful mathematical tool for proving existence of solution concepts in game theory including Nash equilibrium. To state the Kakutani fixed-point theorem we first develop some terminology.

The set S of a finite dimensional vector space R^m is closed if for every convergent sequence of vectors $\{x^j\}$, $j = 1, \ldots, \infty$, if $x^j \in S$ for every j, then $\lim_{j\to\infty} x^j \in S$.

The set S is bounded if there exists some positive number K such that for every $x \in S$, $\sum_{i=1}^{m} \xi_i^2 \leq K$ (here $x = \{\xi_i\}$, ξ_i are the components of x).

A point-to-set correspondence $F : X \to Y$ is any mapping that specifies, for every point x in X, a set $F(x)$ that is subset of Y. Suppose that X and Y are any metric spaces, so the notion of convergence and limit are defined for sequences of points in X and in Y. A correspondence $F : X \to Y$ is upper-semicontinuous if, for every sequence x^j, y^j, $j = 1, \ldots, \infty$, if $x^j \in S$ and $y^j \in F(x^j)$ for every j, and the sequence $\{x^j\}$ converges to some point $\overline{x}$, and the sequence $\{y^j\}$ converges to some point $\overline{y}$, then $\overline{y} \in F(\overline{x})$. Thus $F : X \to Y$ is upper-semicontinuous, if the set $\{(x, y) : x \in X, y \in F(x)\}$ is a closed subset of $X \times Y$.

A fixed point of a correspondence $F : S \to S$ is any x in S such that $x \in F(x)$.

Theorem 7. *[Kakutani fixed-point theorem.]*

Let S be any nonempty, convex, bounded, and closed subset of a finite-dimensional vector space R^m. Let $F : S \to S$ be any upper-semicontinuous point-to-set correspondence such that, for every x in S, $F(x)$ is a nonempty convex subset of S. Then there exists some $\overline{x}$ in S such that $\overline{x} \in F(\overline{x})$.

Proofs of the Kakutani fixed-point Theorem can be found in Burger (1963).

With the help of the previous theorem we shall prove the following fundamental result.

Theorem 8. *Given any finite n-person game Γ in normal form, there exists at least one equilibrium in mixed strategies (in $\prod_{i\in N} \overline{X}_i$).*

Proof. Let Γ be any finite game in normal form

$$\Gamma = (N, \{X_i\}_{i\in N}, \{H_i\}_{i\in N}).$$

The set of mixed-strategy profiles $\prod_{i=1}^{n} \overline{X}_i$ is a nonempty, convex, closed, and bounded subset of a finite dimensional vector space. This set satisfies the above definition of boundedness with $K = |N|$ and it is a subset of R^m, where $m = \sum_{i=1}^{n} |X_i|$ (here $|A|$ means the number of elements in a finite set A).

For any $\mu \in \prod_{i=1}^{n} \overline{X}_i$ and any player $j \in N$, let

$$R_j(\mu) = \arg \max_{\tau_j \in \overline{X}_j} K_i(\mu \| \tau_j).$$

That is, $R_j(\mu)$ is the set of best responses in $\overline{X}_j$ to the combination of independently mixed strategies $(\mu_1, \ldots, \mu_{j-1}, \mu_{j+1}, \ldots, \mu_N)$ of other players. By previous lemma $R_j(\mu)$ is the set of all probability distributions ρ_j over X_j such that

$$\rho_j(x_j) = 0 \text{ for every } x_j \text{ such that } x_j \notin \arg \max_{y_j \in \overline{X}_j} K_j(\mu \| y_j).$$

Thus, $R_j(\mu)$ is convex, because it is a subset of $\overline{X}_j$ that is defined by a collection of linear equalities. Furthermore, $R_j(\mu)$ is nonempty, because it includes x_j from the set $\arg\max_{y_j \in \overline{X}_j} K_j(\mu \| y_j)$, which is nonempty.

Let $R : \prod_{i=1}^n \overline{X}_i \to \prod_{i=1}^n \overline{X}_i$ be the point-to-set correspondence such that

$$R(\mu) = \prod_{i=1}^n R_j(\mu), \text{ for all } \mu \in \prod_{i=1}^n \overline{X}_i.$$

That is, $\tau \in R(\mu)$ if and only if $\tau_j \in R_j(\mu)$ for every $j \in N$. For each μ, $R(\mu)$ is nonempty and convex, because it is the Cartesian product of nonempty convex sets.

To show that R is upper-semicontinuous, suppose that $\{\mu^k\}$ and $\{\tau^k\}$, $k = 1, \dots, \infty$ are convergent sequences, $\mu^k \in \prod_{i \in N} \overline{X}_i$, $k = 1, 2, \dots$; $\tau_k \in R(\mu_k)$, $k = 1, 2, \dots$; $\overline{\mu} = \lim_{k\to\infty} \mu^k$, $\overline{\tau} = \lim_{k\to\infty} \tau^k$.

We have to show that $\overline{\tau} \in R(\overline{\mu})$. For every player $j \in N$ and every $\rho_j \in \overline{X}_j$

$$K_j(\mu^k \| \tau_j^k) \geq K_j(\mu^k \| \rho_j), \quad k = 1, 2, \dots.$$

By continuity of the mathematical expectation $K_j(\mu)$ on $\prod_{i=1}^n \overline{X}_i$, this in turn implies that, for every $j \in N$ and $\rho_j \in \overline{X}_j$,

$$K_j(\overline{\mu} \| \overline{\tau}_j) \geq K_j(\overline{\mu} \| \rho_j).$$

Thus $\overline{\tau}_j \in R_j(\overline{\mu})$ for every $j \in N$, and by the definition of $R(\overline{\mu})$, $\overline{\tau} \in R(\overline{\mu})$. And we have proved that $R : \prod_{i \in N} \overline{X}_i \to \prod_{i \in N} \overline{X}_i$ is an upper-semicontinuous correspondence.

By the Kakutani fixed-point theorem, there exists some mixed strategy profile μ in $\prod_{i \in N} \overline{X}_i$ such that $\mu \in R(\mu)$. That is $\mu_j \in R_j(\mu)$ for every $j \in N$ thus $K_j(\mu) \geq K_j(\mu \| \tau_j)$ for all $j \in N$, $\tau_j \in \overline{X}_j$, and so μ is a Nash equilibrium of Γ.

6 Refinements of Nash equilibria

Many attempts were made to choose a particular Nash equilibrium from the set of all possible Nash equilibria profiles. There are some approaches, but today it is very difficult to distinguish among them to find out the most perspective ones. We shall introduce only some of them, but for the better understanding of this topic we refer to the book of Eric van Damme (1991).

One of the ideas is that each player with a small probability makes mistakes, and as a consequence every pure strategy is chosen with a positive (although small) probability. This idea is modelled through perturbed games, i.e. games in which players have to use only completely mixed strategies.

Let $\Gamma = \langle N, X_1, \dots, X_n, H_1, \dots, H_n \rangle$ be n-person game in normal form. Denote as before $\overline{X}_i$ – the set of mixed strategies of player i, K_i – mathematical expectation of the payoff of player i in mixed strategies. For $i \in N$, let $\eta_i(x_i)$, $x_i \in X_i$ and $\overline{X}_i(\eta_i)$ be defined by

$$\overline{X}_i(\eta_i) = \{\mu_i \in \overline{X}_i : \mu_i(x_i) \geq \eta_i(x_i), \text{ for all } x_i \in X_i\},$$

where $\eta_i(x_i) > 0$, $\sum_{x_i \in X_i} \eta_i(x_i) < 1$.

Let $\eta(x) = (\eta_1(x_1), \dots, \eta_n(x_n))$, $x_i \in X_i$, $i = 1, \dots, n$ and $\overline{X}[\eta(x)] = \prod_{i=1}^n \overline{X}_i(\eta_i(x_i))$. The perturbed game (Γ, η) is the infinite game in normal form

$$\Gamma = \langle N, \overline{X}_1(\eta_1(x_1)), \dots, \overline{X}_n(\eta_n(x_n)), K_1(\mu_1, \dots, \mu_n), \dots, K_n(\mu_1, \dots, \mu_n) \rangle$$

defined over the strategy sets $\overline{X}_i(\eta_i(x_i))$ with payoffs $K_i(\mu_1, \dots, \mu_n)$, $\mu_i \in \overline{X}_i(\eta_i(x_i))$, $i = 1, \dots, n$.

It is easily seen that a perturbed game (Γ, η) satisfies the conditions under which the Kakutani fixed point theorem can be used and so such a game possesses at least one equilibrium. It is clear that in such an equilibrium a pure strategy which is not a best reply has to be chosen with a minimum probability. And we have the following lemma.

Lemma 3. *A strategy profile $\mu \in \overline{X}(\eta)$ is an equilibrium of (Γ, η) if and only if the following condition is satisfied:*

$$\textit{if } K_i(\mu \| x_k) < K_i(\mu \| x_l), \textit{ then } \mu_i(x_k) = \eta_i(x_k), \textit{ for all } i, x_k, x_l.$$

Definition 10. *Let Γ be a game in normal form. An equilibrium μ of Γ is a perfect equilibrium of Γ if μ is a limit point of a sequence $\{\mu(\eta)\}_{\eta \to \infty}$ with $\mu(\eta)$ being Nash equilibrium in a perturbed game (Γ, η) for all η.*

For an equilibrium μ of Γ to be perfect it is sufficient that some perturbed games (Γ, η) with η close to zero possess an equilibrium close to μ and that it is not required that all perturbed games (Γ, η) with η close to zero possess such an equilibrium. Let $\{(\Gamma, \eta^k)\}$, $k = 1, \dots, \infty$ be a sequence of perturbed games for which $\eta^k \to 0$ as $k \to \infty$. Since every game (Γ, η^k) possesses at least one equilibrium μ^k, and since μ is an element of compact set $\overline{X} = \prod_{j=1}^n \overline{X}_i$, there exists one limit point of $\{\mu^k\}$. It can be easily seen that this limit point is an equilibrium of Γ and this will be a perfect equilibrium. Thus the following theorem holds.

Theorem 9. *[10] Every game in normal form possesses at least one perfect equilibrium.*

Example 19. Consider a bimatrix game Γ

$$\begin{array}{cc} & \begin{array}{cc} L_2 & R_2 \end{array} \\ \begin{array}{c} L_1 \\ R_1 \end{array} & \begin{bmatrix} (1,1) & (0,0) \\ (0,0) & (0,0) \end{bmatrix} \end{array}.$$

This game has two equilibria (L_1, L_2) and (R_1, R_2). Consider a perturbed game (Γ, η). In the situation (R_1, R_2) in the perturbed game the strategies R_1 and R_2 will be chosen with probabilities $1 - \eta_1(L_1)$ and $1 - \eta_2(L_2)$ respectively and the strategies L_1 and L_2 will be chosen with probabilities $\eta_1(L_1)$ and $\eta_2(L_2)$. Thus the payoff $K_1^\eta(R_1 R_2)$ in (Γ, η) will be equal

$$K_1^\eta(R_1, R_2) = \eta_1(L_1) \cdot \eta_2(L_2).$$

In the situation (R_1, L_2) the strategies R_1 and L_2 will be chosen with probabilities $(1-\eta_1(L_1))$ and $(1-\eta_2(R_2))$, and

$$K_1^\eta(L_1, R_2) = \eta_1(L_1)(1-\eta_2(L_2)).$$

Since η is small we get

$$K_1^\eta(L_1, R_2) > K_1^\eta(R_1, R_2).$$

Then we see that in perturbed game (R_1, R_2) is not an equilibrium, from this it follows that (R_1, R_2) is not a perfect equilibrium in the original game. It is easily seen that (L_1, L_2) is a perfect equilibrium.

Consider now the game with matrix

$$\begin{array}{cc} & \begin{array}{cc} L_2 & \qquad R_2 \end{array} \\ \begin{array}{c} L_1 \\ R_1 \end{array} & \left[\begin{array}{cc} (1,1) & (10,0) \\ (0,10) & (10,10) \end{array}\right] \end{array}.$$

In this game we can see that a perfect equilibrium (L_1, L_2) is payoff dominated by a non-perfect one. This game also has two different equilibria, (L_1, L_2) and (R_1, R_2). Consider the perturbed game (Γ, η). Show that (L_1, L_2) is a perfect equilibrium in (Γ, η)

$$K_1(L_1, L_2) = (1-\eta_1(R_1))(1-\eta_2(R_2)) + 10(1-\eta_1(R_1))\eta_2(R_2) + 10\eta_1(R_1)\eta_2(R_2),$$

$$K_1(R_1, L_2) = 10(1-\eta_1(L_1))\eta_2(R_2) + 1\cdot\eta_1(L_1)(1-\eta_2(R_2)) + 10\eta_1(L_1)\eta_2(R_2).$$

For η_1, η_2 small we have

$$K_1(L_1, L_2) > K_2(R_1, L_2).$$

In the similar way we can show that

$$K_2(L_1, L_2) > K_2(L_1, R_2).$$

Consider now (R_1, R_2) in (Γ, η)

$$K_1(R_1, R_2) = 10(1-\eta_1(L_1))(1-\eta_2(L_2)) + 10(1-\eta_2(L_2))\eta_1(L_1) + \eta_1(L_1)\eta_2(L_2)$$

$$= 10(1-\eta_2(L_2)) + \eta_1(L_1)\eta_2(L_2),$$

$$K_1(L_1, R_2) = 10(1-\eta_1(R_1))(1-\eta_2(L_2)) + 10\eta_1(R_1)(1-\eta_2(L_2)) + (1-\eta_1(R_1))\eta_2(L_2)$$

$$= 10(1-\eta_2(L_2)) + (1-\eta_1(R_1))\eta_2(L_2).$$

For small η we have $K_1^\eta(L_1, R_2) > K_1^\eta(R_1, R_2)$. Thus (R_1, R_2) is not an equilibrium in (Γ, η) and it cannot be a perfect equilibrium in Γ.

It can be seen that (L_1, L_2) equilibrium in (Γ, η), and the only perfect equilibrium in Γ, but this equilibrium is payoff dominated by (R_1, R_2). We see that the perfectness refinement eliminates equilibria with attractive payoffs. At the same time the perfectness concept does not eliminate all intuitively unreasonable equilibria.

As it is seen from the example of [11]

$$\begin{array}{cc} & \begin{array}{ccc} L_2 & R_2 & A_2 \end{array} \\ \begin{array}{c} L_1 \\ R_1 \\ A_1 \end{array} & \left[\begin{array}{ccc} (1,1) & (0,0) & (-1,-2) \\ (0,0) & (0,0) & (0,-2) \\ (-2,-1) & (-2,0) & (-2,-2) \end{array} \right] \end{array} .$$

It can be seen that an equilibrium (R_1, R_2) in this game is also perfect. Namely if the players have agreed to play (R_1, R_2) and if each player expects, that the mistake A will occur with a larger probability than the mistake L, then it is optimal for each player to play R. Hence adding strictly dominated strategies may change the set of perfect equilibria.

There is another refinement of equilibria concept introduced by Myerson (1978), which exclude some "unreasonable" perfect equilibria like (R_1, R_2) in the last example.

This is the so-called *proper equilibrium*. The basic idea underlying the properness concept is that a player when making mistakes, will try much harder to prevent more costly mistakes than he will try to prevent the less costly ones, i.e. that there is some rationality in the mechanism of making mistakes. As a result, a more costly mistake will occur with a probability which is of smaller order than the probability of a less costly one.

Definition 11. *Let $\langle N, \overline{X}_1, \ldots, \overline{X}_n, K_1, \ldots, K_n \rangle$ be an n-person normal form game in mixed strategies. Let $\epsilon > 0$, and $\mu^\epsilon \in \prod_{i=1}^n \overline{X}_i$. We say that the strategy profile μ^ϵ is an ϵ-proper equilibrium of Γ if μ^ϵ is completely mixed and satisfies*

$$\text{if } K_i(\mu^\epsilon \| x_k) < K_i(\mu^\epsilon \| x_l), \text{ then } \mu_i^\epsilon(x_k) < \epsilon \mu_i^\epsilon(x_l) \text{ for all } i, k, l.$$

$\mu \in \prod_{i=1}^n \overline{X}_i$ is a proper equilibrium of Γ if μ is a limit point of a sequence $\mu^\epsilon(\epsilon \to 0)$, where μ^ϵ is an ϵ-proper equilibrium of Γ.

The following theorem holds.

Theorem 10. *[11] Every normal form game possesses at least one proper equilibrium.*

When introducing perfectness and properness concepts we considered refinements of the Nash equilibrium which are based on the idea that a reasonable equilibrium should be stable against slight perturbations in the equilibrium strategies. There are refinements based on the idea that a reasonable equilibrium should be stable against perturbations in the payoffs of the game. But we do not cover all possible refinements. We recommend the readers to the book of Eric van Damme (1991) for a complete investigation of the problem.

7 Properties of optimal solutions

We shall now present some of the equilibrium properties which may be helpful in finding a solution of a noncooperative two-person game.

Theorem 11. *In order for a mixed strategy situation* (μ^*, ν^*) *in the game* $\Gamma = (X_1, X_2, H_1, H_2)$ *to be an equilibrium, it is necessary and sufficient that for all the players' pure strategies* $x \in X_1$ *and* $y \in X_2$ *the following inequalities be satisfied:*

$$K_1(x, \nu^*) \leq K_1(\mu^*, \nu^*), \tag{7.1}$$

$$K_2(\mu^*, y) \leq K_2(\mu^*, \nu^*). \tag{7.2}$$

Proof. The necessity is evident, since every pure strategy is a special case of a mixed strategy, and hence inequalities (7.1), (7.2) must be satisfied. To prove the sufficiency, we need to shift to the mixed strategies of Players 1 and 2, respectively, in inequalities (7.1), (7.2).

This theorem (as in the case of zero-sum games) shows that, for the proof that the situation forms an equilibrium in mixed strategies it only suffices to verify inequalities (7.1), (7.2) for opponent's pure strategies. For the bimatrix $(m \times n)$ game $\Gamma(A, B)$ these inequalities become

$$K_1(i, y^*) \triangleq a_i y^* \leq x^* A y^* = K_1(x^*, y^*), \tag{7.3}$$

$$K_2(x^*, j) \triangleq x^* b^j \leq x^* B y^* = K_2(x^*, y^*), \tag{7.4}$$

where $a_i(b^j)$ are rows (columns) of the matrix $A(B)$, $i = 1, \dots, m$, $j = 1, \dots, n$.

Recall that, for matrix games, each essentially pure strategy equalizes any optimal strategy of the opponent. A similar result is also true for bimatrix games.

Theorem 12. *Let* $\Gamma(A, B)$ *be a bimatrix* $(m \times n)$ *game and let* $(x, y) \in Z(\Gamma)$ *be a Nash equilibrium in mixed strategies. Then the equations*

$$K_1(i, y) = K_1(x, y), \tag{7.5}$$

$$K_2(x, j) = K_2(x, y) \tag{7.6}$$

hold for all $i \in M_x$ *and* $j \in N_y$*, where* $M_x(N_y)$ *is the spectrum of a mixed strategy* $x(y)$*.*

Proof. By the Theorem 11, we have

$$K_1(i, y) \leq K_1(x, y) \tag{7.7}$$

for all $i \in M_x$. Suppose that at least one strict inequality in (7.7) is satisfied. That is

$$K_1(i_0, y) < K_1(x, y), \tag{7.8}$$

where $i_0 \in M_x$. Denote by ξ_i the components of the vector $x = (\xi_1, \ldots, \xi_m)$. Then $\xi_{i_0} > 0$ and

$$K_1(x,y) = \sum_{i=1}^{m} \xi_i K_1(i,y) = \sum_{i\in M_x} \xi_i K_1(i,y) < K_1(x,y) \sum_{i\in M_x} \xi_i = K_1(x,y).$$

The contradiction proves the validity of (7.5). Equations (7.6) can be proved in the same way.

This theorem provides a means of finding equilibrium strategies of players in the game $\Gamma(A,B)$. Indeed, suppose we are looking for an equilibrium (x,y), with the strategy spectra M_x, N_y being given. The optimal strategies must then satisfy a system of linear equations

$$ya_i = v_1, \quad xb^j = v_2, \tag{7.9}$$

where $i \in M_x$, $j \in N_y$, v_1, v_2 are some numbers. If, however, the equilibrium (x,y) is completely mixed, then the system (7.9) becomes

$$Ay = v_1 u, \quad xB = v_2 w, \tag{7.10}$$

where $u = (1,\ldots,1)$, $w = (1,\ldots,1)$ are the vectors of suitable dimensions composed of unit elements, and the numbers $v_1 = xAy$, $v_2 = xBy$ are the players' payoffs in the situation (x,y).

Theorem 13. *Let $\Gamma(A,B)$ be a bimatrix $(m \times m)$ game, where A, B are nonsingular matrices. If the game Γ has a completely mixed equilibrium, then it is unique and is defined by formulas*

$$x = v_2 u B^{-1}, \tag{7.11}$$

$$y = v_1 A^{-1} u, \tag{7.12}$$

where

$$v_1 = 1/(uA^{-1}u), \quad v_2 = 1/(uB^{-1}u). \tag{7.13}$$

Conversely, if $x \geq 0$, $y \geq 0$ hold for the vectors $x, y \in R^m$ defined by (7.11)–(7.13), then the pair (x,y) forms an equilibrium in mixed strategies in the game $\Gamma(A,B)$ with the equilibrium payoff vector (v_1, v_2).

Proof. If (x,y) is a completely mixed equilibrium, then x and y necessarily satisfy system (7.10). Multiplying the first of the equations (7.10) by A^{-1}, and the second by B^{-1}, we obtain (7.11), (7.12). On the other hand, since $xu = 1$ and $yu = 1$, we find values for v_1 and v_2. The uniqueness of the completely mixed situation (x,y) follows from the uniqueness of the solution of system (7.10) in terms of the theorem.

We shall now show that the reverse is also true. By the construction of the vectors x, y in terms of (7.11)–(7.13), we have $xu = yu = 1$. From this, and from the conditions $x \geq 0$, $y \geq 0$, it follows that (x,y) is a situation in mixed strategies in the game Γ.

By Theorem 11, for the situation (x, y) to be an equilibrium in mixed strategies in the game $\Gamma(A, B)$, it suffices to satisfy the conditions

$$a_i y = K_1(i, y) \leq xAy, \ i = \overline{1, m},$$

$$xb^j = K_2(x, j) \leq xBy, \ j = \overline{1, m},$$

or

$$Ay \leq (xAy)u, \quad xB \leq (xBy)u.$$

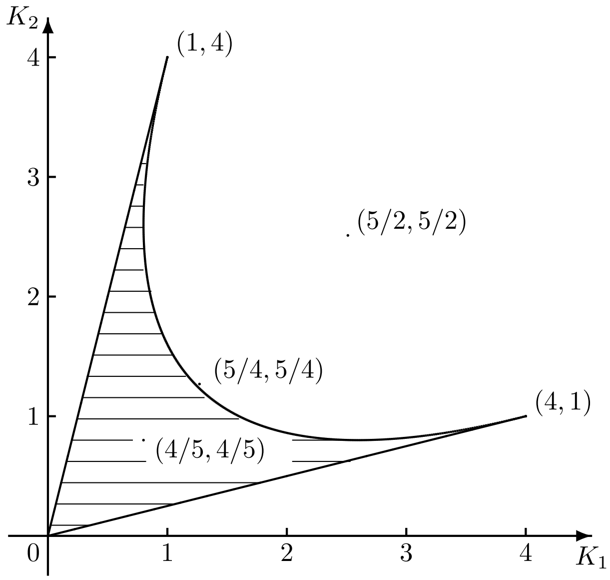
Let us check the validity of these relations for $x = \dfrac{uB^{-1}}{uB^{-1}u}$ and $y = \dfrac{A^{-1}u}{uA^{-1}u}$. We have

$$Ay = \frac{u}{uA^{-1}u} = \frac{(uB^{-1}AA^{-1}u)u}{(uB^{-1}u)(uA^{-1}u)} = (xAy)u,$$

$$xB = \frac{u}{uB^{-1}u} = \frac{(uB^{-1}BA^{-1}u)u}{(uB^{-1}u)(uA^{-1}u)} = (xBy)u,$$

which proves the statement.

We shall now demonstrate an application of the theorem with the example of a "battle of the sexes" game. Consider a mixed extension of the game. The set of points representing the payoff vectors in mixed strategies can be represented graphically (Plot 3, Exercise 6). It can be easily seen that the game



Plot 3

satisfies the conditions of the theorem; therefore, it has a unique, completely mixed equilibrium (x, y) which can be computed by the formulas (7.11)–(7.13): $x = (4/5, 1/5)$, $y = (1/5, 4/5)$, $(v_1, v_2) = (4/5, 4/5)$.

We shall now consider the properties of various optimality principles. The theorem of competition for leadership holds for the two-person game:

$$Z(\overline{\Gamma}) = \overline{Z}^1 \cup \overline{Z}^2,$$

where $Z(\overline{\Gamma})$ is the set of Nash equilibria, $\overline{Z}^1$ and $\overline{Z}^2$ are the sets of the best responses to be given by the players 1 and 2, respectively, in the game $\overline{\Gamma}$.

Things become more complicated where the Nash equilibria and pareto-optimal situations are concerned. The examples given before suggest the possibility of the cases where the situation is Nash equilibrium, but not pareto-optimal, and vice versa. However, the same situation can be optimal in both senses.

Example 16 shows that an additional equilibrium arising in the mixed extension of the game Γ is not pareto-optimal in the mixed extension of Γ. This appears to be a fairly common property of bimatrix games.

Theorem 14. *Let $\Gamma(A, B)$ be a bimatrix $(m \times n)$ game. Then the following assertion is true for almost all $(m \times n)$ games (except for no more than a countable set of games).*

Nash equilibrium situations in mixed strategies, which are not equilibrium in the original game, are not pareto-optimal in the mixed extension.

For the proof of this theorem, see [2].

In conclusion of this section, we examine an example of the solution of bimatrix games.

Example 20. Bimatrix (2×2) *games* [2]. Consider the game $\Gamma(A, B)$, in which each player has two pure strategies. Let

$$(A, B) = \begin{array}{c c} & \begin{array}{cc} \tau_1 & \tau_2 \end{array} \\ \begin{array}{c} \delta_1 \\ \delta_2 \end{array} & \begin{bmatrix} (\alpha_{11}, \beta_{11}) & (\alpha_{12}, \beta_{12}) \\ (\alpha_{21}, \beta_{21}) & (\alpha_{22}, \beta_{22}) \end{bmatrix} \end{array}.$$

Here the indices δ_1, δ_2, τ_1, τ_2 denote pure strategies of Players 1 and 2, respectively.

For simplicity, assume that the numbers α_{11}, α_{12}, α_{21}, α_{22}, (β_{11}, β_{12}, β_{21}, β_{22}) are different.

Case 1. In the original game Γ, at least one player, say Player 1, has a strictly dominant strategy, say δ_1. Then the game Γ and its mixed extension $\overline{\Gamma}$ have a unique Nash equilibrium. In fact, inequalities $\alpha_{11} > \alpha_{21}$, $\alpha_{12} > \alpha_{22}$ cause the pure strategy δ_1 in the game $\overline{\Gamma}$ to dominate strictly all the other mixed strategies of Player 1. Therefore, an equilibrium is represented by the pair (δ_1, τ_1) if $\beta_{11} > \beta_{12}$, or by the pair (δ_1, τ_2) if $\beta_{11} < \beta_{12}$.

Case 2. The game Γ does not have a Nash equilibrium in pure strategies. Here two mutually exclusive cases *a*) and *b*) are possible:

$$a)\ \alpha_{21} < \alpha_{11},\ \alpha_{12} < \alpha_{22},\ \beta_{11} < \beta_{12},\ \beta_{22} < \beta_{21},$$

$$b)\ \alpha_{11} < \alpha_{21},\ \alpha_{22} < \alpha_{12},\ \beta_{12} < \beta_{11},\ \beta_{21} < \beta_{22},$$

where $\det A \neq 0$, $\det B \neq 0$ and hence the conditions of Theorem 13 are satisfied. The game, therefore, has the equilibrium (x^*, y^*), where

$$x^* = \left(\frac{\beta_{22} - \beta_{21}}{\beta_{11} + \beta_{22} - \beta_{21} - \beta_{12}}, \frac{\beta_{11} - \beta_{12}}{\beta_{11} + \beta_{22} - \beta_{21} - \beta_{12}} \right), \tag{7.14}$$

$$y^* = \left(\frac{\alpha_{22} - \alpha_{12}}{\alpha_{11} + \alpha_{22} - \alpha_{21} - \alpha_{12}}, \frac{\alpha_{11} - \alpha_{21}}{\alpha_{11} + \alpha_{22} - \alpha_{21} - \alpha_{12}} \right) \tag{7.15}$$

while the corresponding equilibrium payoffs v_1 and v_2 are determined by

$$v_1 = \frac{\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}}{\alpha_{11} + \alpha_{22} - \alpha_{21} - \alpha_{12}}, \quad v_2 = \frac{\beta_{11}\beta_{22} - \beta_{12}\beta_{21}}{\beta_{11} + \beta_{22} - \beta_{12} - \beta_{21}}.$$

Case 3. The game Γ has two Nash equilibria. This occurs when one of the following conditions is satisfied:

$$a)\ \alpha_{21} < \alpha_{11},\ \alpha_{12} < \alpha_{22},\ \beta_{12} < \beta_{11},\ \beta_{21} < \beta_{22},$$

$$b)\ \alpha_{11} < \alpha_{21},\ \alpha_{22} < \alpha_{12},\ \beta_{11} < \beta_{22},\ \beta_{12} < \beta_{21}.$$

In case a), the situations (δ_1, τ_1), (δ_2, τ_2) are found to be equilibrium, whereas in case b), the situations (δ_1, τ_2), (δ_2, τ_1) form an equilibrium. The mixed extension, however, has one more completely mixed equilibrium (x^*, y^*) determined by (7.14), (7.15).

The above cases provide an exhaustive examination of a (2×2) game with the matrices having different elements.

8 Network games

Games in information networks form a modern branch of game theory. Their development was connected with expansion of the global information network (Internet), as well as with organization of parallel computations on supercomputers. Here the key paradigm concerns the non-cooperative behavior of a large number of players acting independently (still, their payoffs depend on the behavior of the rest participants). Each player strives for transmitting or acquiring maximum information over minimum possible time. Therefore, the payoff function of players is determined either as the task time or as the packet transmission time over a network (to-be-minimized). Another definition of the payoff function lies in the transmitted volume of information or channel capacity (to-be-maximized). An important aspect is comparing the payoffs of players with centralized (cooperative) behavior and their equilibrium payoffs under non-cooperative behavior. Such comparison provides an answer to the following question. Should one organize management in a system (thus, incurring some costs)? If this sounds inefficient, the system has to be self-organized. Interesting effects arise naturally in the context of equilibration. Generally speaking, in an equilibrium players may obtain non-maximal payoffs. Perhaps, the most striking result covers Braess's (1968) paradox (network expansion reduces the equilibrium payoffs of different players). There exist two approaches to network games analysis. According to the first one, a player chooses a route for packet transmission; a packet is treated as an indivisible quantity. Here we mention the works by Papadmitrious and Koutsoupias (1999). Accordingly, such models will be called the KP-models. The second approach presumes that a packet can be divided into segments and transmitted by different routes. It utilizes the equilibrium concept suggested by J.G. Wardrop (1952).

8.1 The KP-model of optimal routing with indivisible traffic. The price of anarchy.

We begin with an elementary information network representing m parallel channels (see Fig. 1).

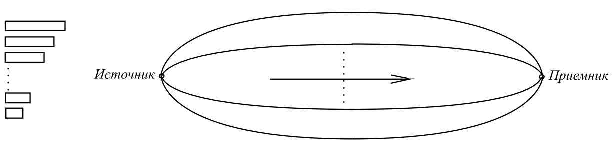


Figure 1: A network of parallel channels.

Consider a system of n users (players). Player i ($i = 1, \dots, n$) intends to send traffic of some volume w_i through a channel. Each channel $l = 1, \dots, m$ has a given capacity c_l. When traffic of a volume w is transmitted by a channel with a capacity c, the channel delay equals w/c.

Each user pursues individual interests, endeavoring to occupy the minimal delay channel. The pure strategy of player i is the choice of channel l for his traffic. Consequently, the vector $L = (l_1, \dots, l_n)$ makes the pure strategy profile of all users; here l_i means the number of the channel selected by user i. His mixed strategy represents the probabilistic distribution $p_i = (p_i^1, \dots, p_i^m)$, where p_i^l stands for the probability of choosing channel l by user i. The matrix P composed of the vectors p_i is the mixed strategy profile of the users.

In the case of pure strategies for user i, the traffic delay in the channel l_i is determined by $\lambda_i = \dfrac{\sum\limits_{k:l_k=l_i} w_k}{c_{l_i}}$.

Definition 12. *A pure strategy profile* $(l_1, \dots, l_n)$ *is called a Nash equilibrium, if for each user* i *we have* $\lambda_i = \min\limits_{j=1,\dots,m} \dfrac{w_i + \sum\limits_{k \neq i: l_k = j} w_k}{c_j}$.

In the case of mixed strategies, it is necessary to introduce the expected traffic delay for user i employing channel l. This characteristic makes up

$\lambda_i^l = \frac{w_i + \sum\limits_{k=1,k\neq i}^{n} p_k^l w_k}{c_l}$. The minimal expected delay of user i equals $\lambda_i = \min\limits_{l=1,\dots,m} \lambda_i^l$.

Definition 13. *A strategy profile P is called a Nash equilibrium, if for each user i and any channel adopted by him the following condition holds true: $\lambda_i^l = \lambda_i$, if $p_i^l > 0$, and $\lambda_i^l > \lambda_i$, if $p_i^l = 0$.*

Definition 14. *A mixed strategy equilibrium P is said to be a completely mixed strategy equilibrium, if each user selects each channel with a positive probability, i.e., for any $i = 1, \dots, n$ and any $l = 1, \dots, m$ $p_i^l > 0$.*

The quantity λ_i describes the minimum possible individual costs of user i to send his traffic. Pursuing personal goals, each user chooses strategies ensuring this value of the expected delay. The so-called *social costs* characterize the general costs of the system due to channels operation. One can involve the following social costs functions $SC(w, c, L)$ for a pure strategy profile:

1. the linear costs $LSC(w, c, L) = \sum\limits_{l=1}^{m} \frac{\sum\limits_{k:l_k=l} w_k}{c_l}$;
2. the quadratic costs $QSC(w, c, L) = \sum\limits_{l=1}^{m} \frac{\left(\sum\limits_{k:l_k=l} w_k\right)^2}{c_l}$;
3. the maximal costs $MSC(w, c, L) = \max\limits_{l=1,\dots,m} \frac{\sum\limits_{k:l_k=l} w_k}{c_l}$.

Definition 15. *The social costs for a mixed strategy profile P are the expected social costs $SC(w, c, L)$ for a random pure strategy profile L:*

$$SC(w, c, P) = E\left(SC(w, c, L)\right) = \sum_{L=(l_1,\dots,l_n)} \left(\prod_{k=1}^{n} p_k^{l_k} \cdot SC(w, c, L) \right).$$

Denote by $OPT = \min_P SC(w, P)$ the optimal social costs. The global optimum in the model considered follows from social costs minimization. Generally, the global optimum is found by enumeration of all admissible pure strategy profiles. However, in a series of cases, it results from solving the continuous conditional minimization problem for social costs, where the mixed strategies of users (the vector P) act as variables.

Definition 16. *The price of anarchy is the ratio of the social costs in the worst-case Nash equilibrium and the optimal social costs:*

$$PoA = \sup_{P-equilibrium} \frac{SC(w, P)}{OPT}.$$

Moreover, if *sup* affects equilibrium profiles composed of pure strategies only, we mean the pure price of anarchy. Similarly, readers can state the notion of the mixed price of anarchy. The price of anarchy defines how much the social costs under centralized control differ from the social costs when each player acts according to his individual interests. Obviously, $PoA \geq 1$ and the actual deviation from 1 reflects the efficiency of centralized control.

8.2 Pure strategy equilibrium. Braess's paradox.

Study several examples of systems, where the behavior of users represents pure strategy profiles only. As social costs, we select the maximal social costs function. Introduce the notation $(w_{i_1}, \dots, w_{i_k}) \to c_l$ for a situation when traffic segments $w_{i_1}, \dots, w_{i_k}$, belonging to users $i_1, \dots, i_k \in \{1, \dots, n\}$ are transmitted through the channel with the capacity c_l.

Example 21. Actually, it illustrates Braess's paradox under elimination of one channel. Consider the following set of users and channels: $n = 5$, $m = 3$, $w = (20, 10, 10, 10, 5)$, $c = (20, 10, 8)$ (see Fig. 2). In this case, there exist several Nash equilibria. One of them consists in the strategy profile

$$\{(10, 10, 10) \to 20, 5 \to 10, 20 \to 8) = 2.5\}.$$

Readers can easily verify that any deviation of a player from this profile increases his delay. However, such equilibrium maximizes the social costs:

$$MSC(w; c; (10, 10, 10) \to 20, 5 \to 10, 20 \to 8) = 2.5) = 2.5.$$

We call this equilibrium the **worst-case equilibrium**.

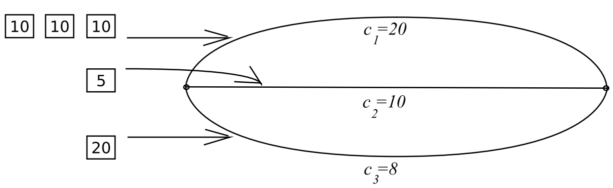


Figure 2: The worst-case Nash equilibrium with the delay of 2.5.

Interestingly, the global optimum of the social costs is achieved in the strategy profile $(20, 10) \to 20, (10, 5) \to 10, 10 \to 8$; it makes up 1.5. Exactly this value represents the bestcase pure strategy Nash equilibrium. If we remove channel 8 (see Fig. 3), the worst-case social costs become

$$MSC(w; c; (20, 10, 10) \to 20, (10, 5) \to 10) = 2.$$

This strategy profile forms the best-case pure strategy equilibrium and the global optimum.

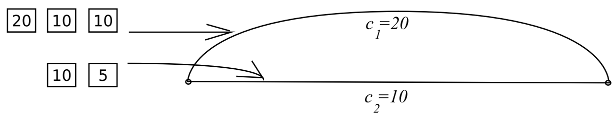


Figure 3: Delay reduction owing to channel elimination.

Example 22. Braess's paradox. This modelwas proposed by D. Braess in 1968. Consider a road network shown in Fig. 4. Suppose that 60 automobiles move from point A to point B. The delay on the segments (C, B) and (A, D) does not depend on the number of automobiles (it equals 1 h). On the segments (A, C) and (D, B) the delay is proportional to the number of moving automobiles (measured in mins). Obviously, here an equilibrium lies in the equal distribution of automobiles between the routes (A, C, B) and (A, D, B), i.e., 30 automobiles per route. In this case, for each automobile the trip consumes 1.5 h.

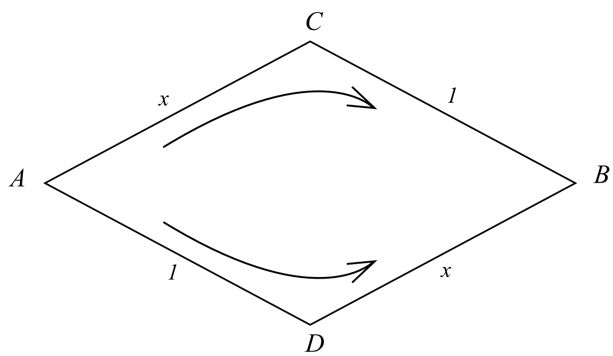


Figure 4: In the equilibrium, players are equally distributed between the routes.

Now, imagine that we have connected points C and D by a speedway, where each automobile has zero delay (see Fig. 5). Then automobiles that have previously selected the route (A, D, B), benefit from moving along the route (A, C, D, B). This applies to automobiles that have previously chosen the route (A, C, B), they should move along the route (A, C, D, B) as well. Hence, the Nash equilibrium (worst case) is the strategy profile, where all automobiles move along the route (A, C, D, B). However, each automobile spends 2 h for the trip.

Therefore, we observe a self-contradictory situationthe costs of each participant have increased as the result of highway construction. This makes Braess's paradox.

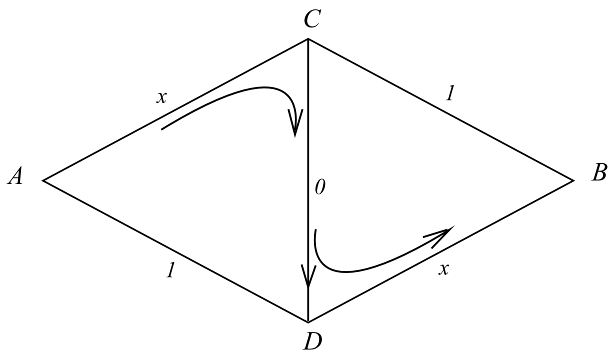


Figure 5: $ACDB$.

8.3 The Wardrop optimal routing model with divisible traffic.

The routing model studied in this section is based on theWardrop model with divisible traffic suggested in 1952. Here the optimality criterion lies in traffic delay minimization.

The optimal traffic routing problem is treated as a game $\Gamma = \langle n, G, w, Z, f\rangle$, where n users transmit their traffic by network channels; the network has the topology described by a graph $G = (V, E)$. For each user i, there exists a certain set Z_i – of routes from s_i to t_i via channels G and a given volume of traffic w_i. Next, each channel $e \in E$ possesses some capacity $c_e > 0$. All users pursue individual interests and choose routes for their traffic to minimize the maximal delay during traffic transmission from s to t. Each user selects a specific strategy $x_i = \{x_{iR_i} \geq 0\}_{R_i \in Z_i}$. The quantity x_{iR_i} determines the volume of traffic sent by user i through route R_i, and $\sum_{R_i \in Z_i} x_{iR_i} = w_i$. Then $x = (x_1, \dots, x_n)$ represents a strategy profile of all users. For a strategy profile x, we again introduce the notation $(x_{-i}, x'_i) = (x_1, \dots, x_{i-1}, x'_i, x_{i+1}, \dots, x_n)$. It indicates that user i has modified his strategy from x_i to x'_i, while the rest users keep their strategies invariable.

For each channel $e \in E$, define its load (the total traffic through this channel) by

$$\delta_e(x) = \sum_{i=1}^{n} \sum_{R_i \in Z_i : e \in R_i} x_{iR_i}.$$

The traffic delay on a given route depends on the loads of channels in this route. The continuous latency function $f_{iR_i}(x) = f_{iR_i}(\{\delta_e(x)\}_{e \in R_i})$ is specified for each user i and each route R_i engaged by him. Actually, it represents a non-decreasing function with respect to the loads of channels in a route (*ergo*, with respect to x_{iR_i}).

Each user i strives to minimize the maximal traffic delay over all channels in his route:

$$PC_i(x) = \max_{R_i \in Z_i : x_{iR_i} > 0} f_{iR_i}(x).$$

This function represents the individual costs of user i.

A Nash equilibrium is defined as a strategy profile such that none of the players benefit by unilateral deviation from this strategy profile (provided that the rest players still follow their strategies). In terms of the current model, the matter concerns a strategy profile such that none of the players can reduce his individual costs by modifying his strategy.

Definition 17. *A strategy profile x is called a Nash equilibrium, if for each user i and any strategy profile $x' = (x_{-i}, x'_i)$ we have $PC_i(x) \le PC_i(x')$.*

Within the framework of network models, an important role belongs to the concept of a Wardrop equilibrium.

Definition 18. *A strategy profile x is called a Wardrop equilibrium, if for each i and any $R_i, \rho_i \in Z_i$ the condition $x_{iR_i} > 0$ leads to $f_{iR_i}(x) \le f_{i\rho_i}(x)$.*

This definition can be restated similarly to the definition of a Nash equilibrium.

Definition 19. *A strategy profile x is a Wardrop equilibrium, if for each i the following condition holds true: the inequality $x_{iR_i} > 0$ leads to $f_{iR_i}(x) = \min\limits_{\rho_i \in Z_i} f_{i\rho_i}(x) = \lambda_i$ and the equality $x_{iR_i} = 0$ yields $f_{iR_i}(x) \ge \lambda_i$.*

Such explicit definition provides a system of equations and inequalities for evaluating Wardrop equilibrium strategy profiles. Strictly speaking, the definitions of a Nash equilibrium and a Wardrop equilibrium are not equivalent. Their equivalence depends on the type of latency functions in channels.

Theorem 15. *If a strategy profile x represents a Wardrop equilibrium, then x is a Nash equilibrium.*

Proof. Let x be a strategy profile such that for all i we have the following condition: the inequality $x_{iR_i} > 0$ brings to $f_{iR_i}(x) = \min\limits_{\rho_i \in Z_i} f_{i\rho_i}(x) = \lambda_i$ and the equality $x_{iR_i} = 0$ implies $f_{iR_i}(x) \ge \lambda_i$. Then for all i and R_i one obtains

$$\max_{\rho_i \in Z_i : x_{i\rho_i} > 0} f_{i\rho_i}(x) \le f_{iR_i}(x).$$

Suppose that user i modifies his strategy from x_i to x'_i. In this case, denote by $x' = (x_{-i}, x'_i)$ a strategy profile such that, for user i, the strategies on all his routes $R_i \in Z_i$ change to $x'_{iR_i} = x_{iR_i} + \Delta_{R_i}$, where $\sum_{R_i \in Z_i} \Delta_{R_i} = 0$. The rest users $k \ne i$ adhere to the same strategies as before, i.e., $x'_k = x_k$.

If all $\Delta_{R_i} = 0$, then $PC_i(x) = PC_i(x')$. Assume that $x \ne x'$, there exists a route R_i such that $\Delta_{R_i} > 0$. This route meets the condition $f_{iR_i}(x) \le f_{iR_i}(x')$, since $f_{iR_i}(x)$ is a non-decreasing function in x_{iR_i}. As far as $x'_{iR_i} > 0$, we get

$$f_{iR_i}(x') \le \max_{\rho_i \in Z_i : x_{i\rho_i} > 0} f_{i\rho_i}(x').$$

Finally

$$\max_{\rho_i \in Z_i : x_{i\rho_i} > 0} f_{i\rho_i}(x) \leq \max_{\rho_i \in Z_i : x_{i\rho_i} > 0} f_{i\rho_i}(x'),$$

or $PC_i(x) \leq PC(x')$. Hence, due to the arbitrary choice of i and x'_i we conclude that the strategy profile x forms a Nash equilibrium.

Any Nash equilibrium in the model considered represents a Wardrop equilibrium under the following sufficient condition imposed on all latency functions. For a given user, it is possible to redistribute a small volume of his traffic from any route to other (less loaded) routes for this user such that the traffic delay on this route becomes strictly smaller.

Example 23. Consider a simple example explaining the difference between the definitions of a Nash equilibrium and a Wardrop equilibrium. A system contains one user, who sends traffic of volume 1 from node s to node t via two routes (see Fig. 6).

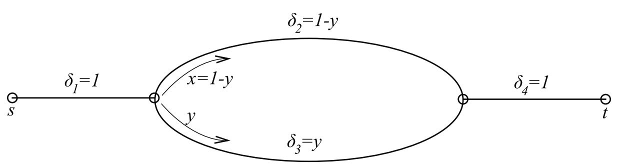


Figure 6: A Nash equilibrium mismatches a Wardrop equilibrium.

Suppose that the latency functions on route 1 (which includes channels (1,2,4)) and on route 2 (which includes channels (1,3,4)) have the form $f_1(x) = \max\{1, x, 1\} = 1$ and $f_2(y) = \min\{1, y, 1\} = y$, respectively; here $x = 1 - y$. Both functions are continuous and nondecreasing in x and y, respectively. The inequality $f_1(x) > f_2(y)$ takes place for all feasible strategy profiles (x, y), such that $x + y = 1$. However, any reduction in x (the volume of traffic through channel 1) does not affect $f_1(x)$. In the described model, a Nash equilibrium is any strategy profile $(x, 1 - x)$, where $0 \leq x \leq 1$. Still, the delays in both channels coincide only for the strategy profile $(0, 1)$.

Definition 20. *Let x indicate some strategy profile. The social costs are the total delay of all players under this strategy profile:*

$$SC(x) = \sum_{i=1}^{n} \sum_{R_i \in Z_i} x_{iR_i} f_{iR_i}(x).$$

Note that, if x represents a Wardrop equilibrium, then (by definition) for each player i the delays on all used routes R_i equal $\lambda_i(x)$. Therefore, the social costs in the equilibrium acquire the form

$$SC(x) = \sum_{i=1}^{n} w_i \lambda_i(x).$$

Designate by $OPT = \min_x SC(x)$ the minimal social costs.

Definition 21. *We call the price of anarchy the maximal value of the ratio $SC(x)/OPT$, where the social costs are evaluated only in Wardrop equilibria.*

8.4 The optimal routing model with parallel channels. The Pigou model. Braess's paradox

We analyze the Wardrop model for a network with parallel channels.

Example 24. The Pigou model (1920). Consider a simple network with two parallel channels (see Fig. 7). One channel possesses the fixed capacity of 1, whereas the second channel has the capacity proportional to traffic. Imagine very many users transmitting their traffic from node s to node t such that the total load is 1. Each user seeks to minimize his costs. Then a Nash equilibrium lies in employing the lower channel for each user. Indeed, if the upper channel comprises a certain quantity of players, the lower channel always guarantees a smaller delay than the upper one. Therefore, the costs of each player in the equilibrium make up 1. Furthermore, the social costs constitute 1 too.

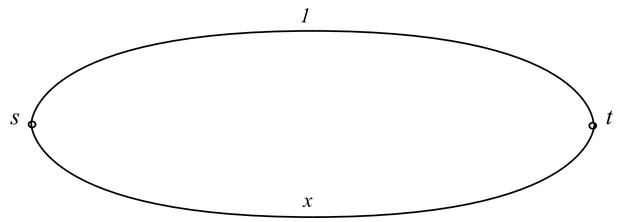


Figure 7: .

Now, assume that some share x of users utilize the upper channel, and the rest users (the share $1-x$) employ the lower channel. Then the social costs become $x \cdot 1 + (1-x) \cdot (1-x) = x^2 - x + 1$. The minimal social costs of $3/4$ correspond to $x = 1/2$. Obviously, the price of anarchy in the Pigou model is $PoA = 4/3$.

Example 25. Consider the same two-channel network, but set the delay in the lower channel equal to x^p, where p means a certain parameter. A Nash equilibrium also consists in sending the traffic of all users through the lower channel (the social costs make up 1). Next, send some volume ϵ of traffic by the upper channel. The corresponding social costs $\epsilon \cdot 1 + (1-\epsilon)^{p+1}$ possess arbitrary small values as $\epsilon \to 0$ and $p \to \infty$. And so, the price of anarchy can have arbitrary large values.

Example 26. Braess's paradox. Recall that we have explored this phenomenon in the case of indivisible traffic. Interestingly, Braess's paradox arises in models with divisible traffic. Select a network composed of four nodes, see Fig. 4. There are two routes from node s to node t with the identical delays of $1 + x$. Suppose that the total traffic of all users equals 1. Owing to the symmetry of this network, all users get partitioned into two equal groups with the identical costs of 3/2. This forms a Nash equilibrium.

To proceed, imagine that we have constructed a new superspeed channel (CD) with zero delay. Then, for each user, the route $A \to C \to D \to B$ is always not worse than the route $A \to C \to B$ or $A \to D \to B$. Nevertheless, the costs of all players increase up to 2 in the new equilibrium. This example shows that adding a new channel may raise the costs of individual players and the social costs.

References

[1] Luce, R. D. and H. Raiffa. *Games and decisions. Introduction and critical survey*, p. 509. Wiley, N.Y., 1957.

[2] Moulin, H. *Theorie des jeux pour l'economie et la politique*, p. 200. Hermann, Paris, 1981.

[3] Zenkevich, N. A. and S. N. Voznyuk. In V. V. Mazalov and L. A. Petrosjan, editors, *Year-book on Game Theory and Applications, vol. 2.* Nauka (Novosibirsk) and Nova Science Publ. (N.Y.), 1994.

[4] Tirole, J. *The theory of industrial organization*, p. 325. Cambridge, Mass., MIT Press, 2003.

[5] Braess, D. *Unternehmensforschung*, 12, 1968.

[6] Vorobjev, N. N. *Game theory lectures for economists and system scientists*, p. 178. Springer, N.Y., 1977.

[7] Moulin, H. *Game Theory for the Social Sciences*, p. 465. N.Y. Univ. Press, N.Y., 2nd edition, 1986.

[8] Parthasarathy, T. and T.E.S. Raghavan. *Some topics in two-person ganes*, p. 259. Amer. Elsevier, N.Y., 1971.

[9] Fudenberg, D. and J. Tirole. *Game theory*, p. 580. The MIT Press, Cambridge, 1992.

[10] Selten, R. *International Journal of game theory*, 4, 1975.

[11] Myerson, R. B. *International Journal of Game Theory*, 7, 1978.

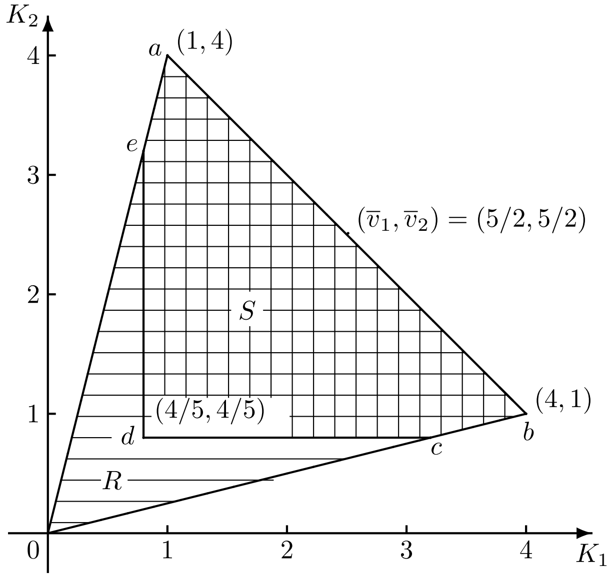
Cooperative games

November 28, 2018

1 The bargaining problem

This section deals with the question: how rational players can come to an agreement on a joint choice by negotiations. Before stating the problem, we return to the game "battle of the sexes" once again.

Example 1. Consider the set R corresponding to possible payoff vectors in joint mixed strategies for the game "battle of the sexes" (this region is shaded in Plot 1). Acting together, the players can ensure any payoff in mixed strategies in



Plot 1.

the region R. However, this does not mean that they can agree on any outcome of the game. Thus, the point $(4,1)$ is preferable to Player 1 whereas the point $(1,4)$ is preferable to Player 2. Neither of the two players can agree with the results of negotiations if his payoff is less than the maximin value, since he can receive this payoff independently of his partner. Maximin mixed strategies for the players in this game are respectively $x^0 = (1/5, 4/5)$ and $y^0 = (4/5, 1/5)$, while the payoff vector in maximin strategies (v_1^0, v_2^0) is $(4/5, 4/5)$. Therefore, the set S, which can be used in negotiations, is bounded by the points a, b, c, d, e (see Plot 1). This set will be called a *bargaining set of the game*. Furthermore, acting jointly, the players can always agree to choose points on the line segment $\overline{ab}$, since this is advantageous to both of them (the line segment $\overline{ab}$ corresponds to pareto-optimal situations).

The problem of choosing the points $(\overline{v}_1, \overline{v}_2)$ from S by bargaining will be a *bargaining problem*. This brings us to the following consideration. Let the bargaining set S and the maximin payoff vector (v_1^0, v_2^0) be given for the bimatrix game $\Gamma(A, B)$. We need to find the device capable for solving bargaing problem, i.e. to find a function φ such that

$$\varphi(S, v_1^0, v_2^0) = (\overline{v}_1, \overline{v}_2), \tag{1.1}$$

where $(\overline{v}_1, \overline{v}_2) \in S$ is the solution. The point (v_1^0, v_2^0) is called a disagreement point.

It appears that, under some reasonable assumptions, it is possible to construct such a function $\varphi(S, v_1^0, v_2^0)$.

Theorem 1. *Let S be a convex compact set in R^2, and let (v_1^0, v_2^0) be a maximin payoff vector in the game $\Gamma(A, B)$. The set S, the pair $(\overline{v}_1, \overline{v}_2)$ and the function φ satisfy the following conditions:*

1. $(\overline{v}_1, \overline{v}_2) \geq (v_1^0, v_2^0)$.
2. $(\overline{v}_1, \overline{v}_2) \in S$.
3. *If* $(v_1, v_2) \in S$ *and* $(v_1, v_2) \geq (\overline{v}_1, \overline{v}_2)$, *then* $(v_1, v_2) = (\overline{v}_1, \overline{v}_2)$.
4. *If* $(\overline{v}_1, \overline{v}_2) \in \overline{S} \subset S$ *and* $(\overline{v}_1, \overline{v}_2) = \varphi(S, v_1^0, v_2^0)$, *then* $(\overline{v}_1, \overline{v}_2) = \varphi(\overline{S}, v_1^0, v_2^0)$.
5. *Let T be obtained from S by linear transformation* $v_1' = \alpha_1 v_1 + \beta_1$, $v_2' = \alpha_2 v_2 + \beta_2$; $\alpha_1 > 0$, $\alpha_2 > 0$. *If* $\varphi(S, v_1^0, v_2^0) = (\overline{v}_1, \overline{v}_2)$, *then* $\varphi(T, \alpha_1 v_1^0 + \beta_1, \alpha_2 v_2^0 + \beta_2) = (\alpha_1 \overline{v}_1 + \beta_1, \alpha_2 \overline{v}_2 + \beta_2)$.
6. *If for any* $(v_1, v_2) \in S$, *also* $(v_2, v_1) \in S$, $v_1^0 = v_2^0$ *and* $\varphi(S, v_1^0, v_2^0) = (\overline{v}_1, \overline{v}_2)$, *then* $\overline{v}_1 = \overline{v}_2$.

There exists a unique function φ, satisfying 1–6 such that

$$\varphi(S, v_1^0, v_2^0) = (\overline{v}_1, \overline{v}_2).$$

The function φ, which maps the bargaining game (S, v_1^0, v_2^0) into the payoff vector set $(\overline{v}_1, \overline{v}_2)$ and satisfies conditions 1–6, is called a Nash bargaining scheme [1], conditions 1–6 are called *Nash axioms*, and the vector $(\overline{v}_1, \overline{v}_2)$ is called a *bargaining solution vector*. Thus, the bargaining scheme is a realizable optimality principle in the bargaining game.

Before going to prove the theorem we will discuss its conditions using the game "battle of the sexes" as an example (see Plot 1). Axioms 1 and 2 imply that the payoff vector $(\overline{v}_1, \overline{v}_2)$ is contained in the set bounded by the points a, b, c, d, e. The axiom 3 implies that $(\overline{v}_1, \overline{v}_2)$ is pareto-optimal. Axiom 4 shows that the function φ is independent of irrelevant alternatives. This says that if the solution outcome of a given problem remains feasible for a new problem obtained from it by contraction, then it should also be the solution outcome of this new problem. Axiom 5 is the scale invariance axiom and axiom 6 shows that the two players possess equal rights.

The proof of the Theorem 1 is based on the following auxiliary results.

Lemma 1. *If there are points $(v_1, v_2) \in S$ such that $v_1 > v_1^0$ and $v_2 > v_2^0$, then there exists a unique point $(\overline{v}_1, \overline{v}_2)$ which maximizes the function*

$$\theta(v_1, v_2) = (v_1 - v_1^0)(v_2 - v_2^0)$$

over a subset $S_1 \subset S$, $S_1 = \{\ (v_1, v_2) \mid (v_1, v_2) \in S,\ v_1 \geq v_1^0\}$.

Proof. By condition, S_1 is a nonempty compact set while θ is a continuous function, and hence achieves its maximum $\overline{\theta}$ on this set. By assumption, $\overline{\theta}$ is positive.

Suppose there are two different points of maximum (v_1', v_2') and (v_1'', v_2'') for the function θ on S_1. Note that $v_1' \neq v_1''$; otherwise the form of the function θ would imply $v_2' = v_2''$.

If $v_1' < v_1''$, then $v_2' > v_2''$. Since the set S_1 is convex, then $(\overline{v}_1, \overline{v}_2) \in S_1$, where $\overline{v}_1 = (v_1' + v_1'')/2$, $\overline{v}_2 = (v_2' + v_2'')/2$. We have

$$\theta(\overline{v}_1, \overline{v}_2) = \frac{(v_1' - v_1^0) + (v_1'' - v_1^0)}{2} \frac{(v_2' - v_2^0) + (v_2'' - v_2^0)}{2}$$

$$= \frac{(v_1' - v_1^0)(v_2' - v_2^0)}{2} + \frac{(v_1'' - v_1^0)(v_2'' - v_2^0)}{2} + \frac{(v_1' - v_1'')(v_2'' - v_2')}{4}.$$

Each of the first two summands in the last sum is equal to $\overline{\theta}/2$, while the third summand is positive, which is impossible, because $\overline{\theta}$ is the maximum of the function θ. Thus, the point $(\overline{v}_1, \overline{v}_2)$, which maximizes the function θ over the set S_1, is unique.

Lemma 2. *Suppose that S satisfies the conditions of Lemma 1, while $(\overline{v}_1, \overline{v}_2)$ is the point of maximum for the function $\theta(v_1, v_2)$. Define*

$$\delta(v_1, v_2) \equiv (\overline{v}_2 - v_1^0)v_1 + (\overline{v}_1 - v_1^0)v_2.$$

If $(v_1, v_2) \in S$, then the following inequality holds:

$$\delta(v_1, v_2) \leq \delta(\overline{v}_1, \overline{v}_2).$$

Proof. Suppose there exists a point $(v_1, v_2) \in S$ such that $\delta(v_1, v_2) > \delta(\overline{v}_1, \overline{v}_2)$. From the convexity of S we have: $(v_1', v_2') \in S$, where $v_1' = \overline{v}_1 + \epsilon(v_1 -$

$\overline{v}_1)$ and $v_2' = \overline{v}_2 + \epsilon(v_2 - \overline{v}_2)$, $0 < \epsilon < 1$. By linearity, $\delta(v_1 - \overline{v}_1, v_2 - \overline{v}_2) > 0$. We have

$$\theta(v_1', v_2') = \theta(\overline{v}_1, \overline{v}_2) + \epsilon\delta(v_1 - \overline{v}_1, v_2 - \overline{v}_2) + \epsilon^2(v_1 - \overline{v}_1)(v_2 - \overline{v}_2).$$

For a sufficiently small $\epsilon > 0$ we obtain the inequality $\theta(v_1', v_2') > \theta(\overline{v}_1, \overline{v}_2)$, but this contradicts the maximality of $\theta(\overline{v}_1, \overline{v}_2)$.

We shall now prove Theorem 1. To do this, we shall show that the point $(\overline{v}_1, \overline{v}_2)$, which maximizes $\theta(v_1, v_2)$, is a solution of the bargaining problem.

Proof. Suppose the conditions of Lemma 1 are satisfied. Then the point $(\overline{v}_1, \overline{v}_2)$ maximizing $\theta(v_1, v_2)$ is defined. It is easy to verify that $(\overline{v}_1, \overline{v}_2)$ satisfies conditions 1–4 of Theorem 1. This point also satisfies condition 5 of this theorem, because if $v_1' = \alpha_1 v_1 + \beta_1$ and $v_2' = \alpha_2 v_2 + \beta_2$, then

$$\theta'(v_1', v_2') = [v_1' - (\alpha_1 v_1^0 + \beta_1)][v_2' - (\alpha_2 v_2^0 + \beta_2)] = \alpha_1\alpha_2\theta(v_1, v_2),$$

and if $(\overline{v}_1, \overline{v}_2)$ maximizes $\theta(v_1, v_2)$, then $(\overline{v}_1', \overline{v}_2')$ maximizes $\theta'(v_1', v_2')$. Suppose that the set S is symmetric in the sense of condition 6 and $v_1^0 = v_2^0$. Then $(\overline{v}_2, \overline{v}_1) \in S$ and $\theta(\overline{v}_1, \overline{v}_2) = \theta(\overline{v}_2, \overline{v}_1)$. Since $(\overline{v}_1, \overline{v}_2)$ is a unique point, which maximizes $\theta(v_1, v_2)$ over S_1, then $(\overline{v}_1, \overline{v}_2) = (\overline{v}_2, \overline{v}_1)$, i.e. $\overline{v}_1 = \overline{v}_2$.

Thus, the point $(\overline{v}_1, \overline{v}_2)$ satisfies conditions 1–6. Show that this is a unique solution to the bargaining problem. Consider the set

$$R \equiv \{\ (v_1, v_2) \mid \delta(v_1, v_2) \le \delta(\overline{v}_1, \overline{v}_2)\ \}. \tag{1.2}$$

By Lemma 2, the inclusion $S \subset R$ holds. Suppose T is obtained from R by transformation

$$v_1' = \frac{v_1 - v_1^0}{\overline{v}_1 - v_1^0}, \quad v_2' = \frac{v_2 - v_2^0}{\overline{v}_2 - v_2^0}. \tag{1.3}$$

Expressing v_1 and v_2 in terms of (1.3) and substituting into (1.2), we obtain

$$T = \{\ (v_1', v_2') \mid v_1' + v_2' \le 2\ \}$$

and $v_1'^0 = v_2'^0 = 0$. Since T is symmetric, it follows from property 6 that a solution (if any) must lie on a straight line $v_1' = v_2'$, and, by condition 3, it must coincide with the point (1,1), i.e. $(1, 1) = \varphi(T, 0, 0)$. Reversing the transform (1.3) and using property 5, we obtain $(\overline{v}_1, \overline{v}_2) = \varphi(R, v_1^0, v_2^0)$. Since $(\overline{v}_1, \overline{v}_2) \in S$ and $S \subset R$, then by property 4, the pair $(\overline{v}_1, \overline{v}_2)$ is a solution of (S, v_1^0, v_2^0).

Now suppose that the conditions of Lemma 1 are not satisfied, i.e. there are no points $(v_1, v_2) \in S$ for which $v_1 > v_1^0$ and $v_2 > v_2^0$. Then the following cases are possible:

a) There are points, at which $v_1 > v_1^0$ and $v_2 = v_2^0$. Then $(\overline{v}_1, \overline{v}_2)$ is taken to be the point in S, which maximizes v_1 under constraint $v_2 = v_2^0$.

b) There are points, at which $v_1 = v_1^0$ and $v_2 > v_2^0$. In this case, $(\overline{v}_1, \overline{v}_2)$ is taken to be the point in S, which maximizes v_2 under constraint $v_1 = v_1^0$.

c) The bargaining set S degenerates into the point (v_1^0, v_2^0) of maximin payoffs (e.g., the case of matrix games). Set $\overline{v}_1 = v_1^0$, $\overline{v}_2 = v_2^0$.

It can be immediately verified that these solutions satisfy properties 1–6, and properties 1–3 imply uniqueness. This completes the proof of the theorem.

In the game "battle of the sexes", the Nash scheme yields bargaining payoff $(\overline{v}_1, \overline{v}_2) = (5/2, 5/2)$ (see Plot 1).

In this section we survey the axiomatic theory of bargaining for n players. Although alternatives to the Nash solution were proposed soon after the publication of Nash's paper, it is fair to say until the mid-1970s, the Nash solution was often seen by economists and game theorists as the main, if not the only, solution to the bagaining problem. Since all existing solutions are indeed invariant under parallel transformation of the origin and since our own formulation will also assume this invariance, it is convenient to take as admissible only problems that have already been subjected to parallel transformation bringing their disagreement point to the origin. Consequently, $v^0 = (v_1^0, \ldots, v_n^0) = (0, \ldots, 0) \in R^n$ always, and a typical problem is simply denoted by S instead of $(S, 0)$. Finally, all problems are taken to be subsets of R_+^n (instead of R^n). This means that all alternatives that would give any player less than what he gets at the disagreement point $v^0 = 0$ are disregarded.

Definition 1. *The Nash solution N is defined by setting, for all convex, compact, comprehensive subsets $S \subset R_+^n$ containing at least one vector with all positive coordinates (denote $S \in \sum^n$), $N(S)$ equal to the maximizer in $v \in S$ of the "Nash product" $\prod_{i=1}^n v_i$.*

Nash's theorem is based on the following axioms:

1^0. *Pareto-optimality.* For all $S \in \sum^n$, for all $v \in R^n$, if $v \geq \varphi(S)$ and $v \neq \varphi(S)$, then $v \notin S$ [denote $\varphi(S) \in PO(S)$].

A slightly weaker condition is:

2^0. *Weak pareto-optimality.* For all $S \in \sum^n$, for all $v \in R^n$, if $v > \varphi(S)$, then $v \notin S$.

Let $\Pi^n : \{1, \ldots, n\} \to \{1, \ldots, n\}$ be the *class of permutations of order n*. Given $\Pi \in \Pi^n$, and $v \in R^n$, let $\pi(v) \equiv (v_{\pi(1)}, \ldots, v_{\pi(n)})$. Also, given $S \subset R^n$, let $\pi(S) \equiv \{v' \in R^n \mid \exists v \in S \text{ with } v' = \pi(v)\}$.

3^0. *Symmetry.* For all $S \in \sum^n$, if for all $\pi \in \Pi^n, \pi(S) = S$, then $\varphi_i(S) = \varphi_j(S)$ for all i, j (note that $\pi(S) \in \sum^n$).

Let $L^n : R^n \to R^n$ be the class of *positive, independent person-by-person, and linear transformations of order n*. Each $l \in L^n$ is characterised by n positive numbers α_i such that given $v \in R^n, l(v) = (\alpha_1 v_1, \ldots, \alpha_n v_n)$. Now, given $S \subset R^n$, let $l(S) \equiv \{v' \in R^n | \exists v \in S \text{ with } v' = l(v)\}$.

4^0. *Scale invariance.* For all $S \in \sum^n$, for all $l \in L^n$, $\varphi(l(S)) = l(\varphi(S))$ [note that $l(S) \in \sum^n$].

5^0. *Independence of irrelevant alternatives.* For all $S, S' \subset \sum^n$, if $S' \subset S$ and $\varphi(S) \in S'$ then $\varphi(S') = \varphi(S)$.

In previous section we showed the Nash theorem for $n = 2$, i.e. only one solution satisfies these axioms. This result extends directly to arbitrary n.

Theorem 2. *A solution $\varphi(S)$, $S \in \sum^n$ satisfies 1^0, 3^0, 4^0, 5^0 if and only if it is the Nash solution.*

This theorem constitutes the foundation of the axiomatic theory of bargaining. It shows that a *unique* point can be identified for each problem, representing an equitable compromise.

In the mid-1970s, Nash's result become the object of a considerable amount of renewed attention, and the role played by each axiom in the characterization was scrutinized by several authors.

6^0. *Strong individual rationality.* For all $S \in \sum^n$, $\varphi(S) > 0$.

Theorem 3. *[2]. A solution $\varphi(S)$, $S \in \sum^n$ satisfies $3^0, 4^0, 5^0, 6^0$ if and only if it is the Nash solution.*

If 3^0 is dropped from the list of axioms in Theorem 2, a somewhat wider but still small family of additional solutions become admissible.

Definition 2. *Given $a = (\alpha_1, \ldots, \alpha_n)$, $\alpha_i > 0$, $i = 1, \ldots, n$, $\sum_{i=1}^n \alpha_i = 1$, the asymmetric Nash solution with weights a, N^a, is defined by setting, for all $S \in \sum^n$, $N^a(S) \equiv \arg\max \prod_{i=1}^n v_i^{\alpha_i}$, $v \in S$.*

These solutions were introduced by Harsanyi and Selten (1972).

Theorem 4. *A solution $\varphi(S)$, $S \in \sum^n$ satisfies $4^0, 5^0, 6^0$ if and only if it is an asymmetric Nash solution.*

If 6^0 is not used, a few other solutions became available.

Definition 3. *Given $i \in \{1, \ldots, n\}$ the i-th Dictatorial solution D^i is defined by setting, for all $S \in \sum^n, D^i(S)$ equals to the maximal point of S in the direction of the ith unit vector.*

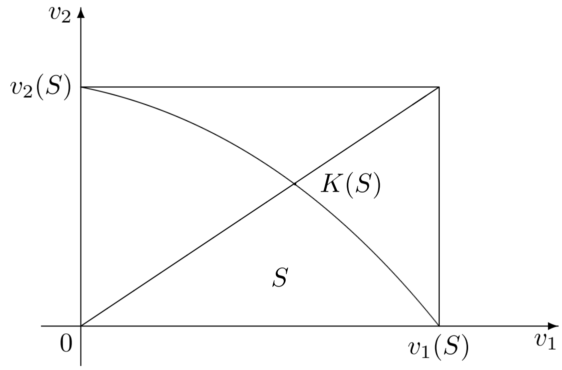
Note that all D^i satisfy $4^0, 5^0$, and 2^0 (but not 1^0). To recover full optimality, one may proceed as follows. First, select an ordering π of the n players. Then given $S \in \sum^n$, pick $D^{\pi(1)}(S)$ if the point belongs to pareto-optimal subset of S; otherwise, among the points whose $\pi(1)$th coordinate is equal to $D^{\pi(1)}(S)$, find the maximal point in the direction of the unit vector pertaining to player $\pi(2)$. Pick this point if it belongs to pareto-optimal subset of S; otherwise, repeat the operation with $\pi(3), \ldots$. This algorithm is summarized in the following definition.

Definition 4. *Given an ordering π of $\{1, \ldots, n\}$, the lexicographic Dictatorial solution relative to π, D^π, if defined by setting, for all $S \in \sum^n, D^\pi(S)$ to be the lexicographic maximizer over $v \in S$ of $v_{\pi(1)}, v_{\pi(2)}, \ldots, v_{\pi(n)}$.*

All of these solutions satisfy $1^0, 4^0, 5^0$, and there are no others if $n = 2$.

The Kalai-Smorodinsky solution. A new impulse was given to the axiomatic theory of bargaining when Kalai and Smorodinsky (1975) provided a characterization of the following solution (see Plot 2).

Definition 5. *The Kalai-Smorodinsky solution K is defined by setting, for all $S \in \sum^n, K(S)$ to be the maximal point of S on the segment connecting the origin to $a(S)$, the ideal point of S, defined by $v_i(S) \equiv \max\{v_i \mid v \in S\}$ for each i.*



Plot 2.

An important distinguishing feature between the Nash solution and the Kalai-Smorodinsky solution is that the latter responds much more satisfactorily to expansions and contractions of the feasible set. In particular, it satisfies the following axiom.

7^0. *Individual monotonicity*: For all $S, S' \in \sum^2$, for all i, if $v_j(S) = v_j(S')$ and $S' \supset S$, then $\varphi_i(S') \geq \varphi_i(S)$.

Theorem 5. *A solution $\varphi(S)$, $S \in \sum^2$ satisfies $1^0, 3^0, 7^0$ if and only if it is the Kalai-Smorodinsky solution.*

Although the extension of the definition of the Kalai-Smorodinsky solution to the n-person case itself causes no problem, the generalization of the preceding results to the n-person case is not as straightforward as was the case of the extensions of the results concerning the Nash solution from $n = 2$ to arbitrary n. First of all, for $n > 2$, the n-person Kalai-Smorodinsky solution satisfies 2^0 only. This is not a serious limitations since, for most problems S, $K(S)$ in fact is (fully) pareto-optimal. But it is not the only change that has to be made in the axioms of Theorem to extend the characterization of the Kalai-Smorodinsky solution to the case $n > 2$.

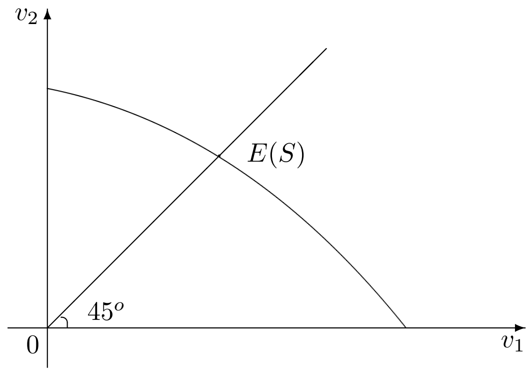
The Egalitarian solution. We now turn to a third solution, which is the main distinguishing feature from the previous two.

Definition 6. *The Egalitarian solution E is defined by setting, for all $S \in \sum^n$, $E(S)$ to be the maximal point of S of equal coordinates (see Plot 3).*

The most striking feature of this solution is that it satisfies the following monotonicity condition, which is very strong, since no restriction are imposed in its hypotheses on the sort of expansions that take S into S'. In fact, this axiom can serve to provide an easy characterization of the solution.

8^0. *Strong monotonicity.* For all $S, S' \in \sum^n$, if $S \subset S'$, then $\varphi(S) \leq \varphi(S')$.

The following characterization result is a variant of a theorem due to Kalai (1977).

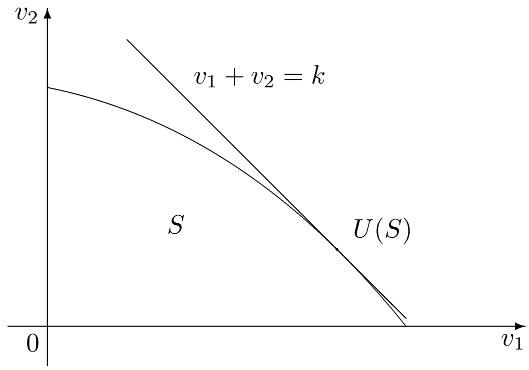


Plot 3.

Theorem 6. *A solution $\varphi(S)$, $S \in \sum^n$ satisfies $2^0, 3^0, 8^0$ if and only if it is the Egalitarian solution.*

The Utilitarian solution. We close this review with a short discussion of the Utilitarian solution.

Definition 7. *A Utilitarian solution U is defined by choosing, for each $S \in \sum^n$ among the maximizers of $\sum_{i=1}^n v_i$ for $v \in S$ (see Plot 4).*



Plot 4.

Obviously, all Utilitarian solutions satisfy 1^0. They also satisfy 3^0 if appropriate selections are made. However, no Utilitarian solution satisfies 4^0. Also, no Utilitarian solution satisfies 5^0, because of the impossibility of performing appropriate selections. The Utilitarian solution has been characterized by Myerson (1981).

Other solutions have been discussed in the literature by Luce and Raiffa (1957), and Perles and Mashler (1981). In this section we follow Thomson and Lensberg (1989), where the reader can find proofs of the theorems.

2 Games in characteristic function form

Previously we demonstrated how the players can arrive at a mutually acceptable decisions on the arising conflict by an agreed choice of strategies (strategic cooperation). We now suppose that the conditions of a game admit the players' joint actions and redistribution of a payoffs. This implies that the utilities of various players can be evaluated by a single scale (transferable payoffs), and hence the mutual redistribution of payoffs does not affect the conceptual statement of the original problem. It appears natural that, from the point of view of each player, the best results may also be produced by uniting players into a maximal coalition (the coalition composed of all players). In this case, we are interested not only in the ways the coalition of players ensures its total payoff, but also in the ways it is distributed among the members of this coalition (cooperative approach).

The chapter deals with the cooperative theory of n-person games. This theory is focused on the conditions under which integration of players into a maximal coalition is advisable and individual players are not interested in forming smaller groups or act individually.

Let $N = \{1, \dots, n\}$ be a set of all players. Any nonempty subset $S \subset N$ is called a *coalition.*

Definition 8. *The real-valued function v defined on coalitions $S \subset N$ is called a characteristic function of the n-person game. Here the inequality*

$$v(T) + v(S) \leq v(T \cup S), \quad v(\oslash) = 0 \tag{2.1}$$

holds for any nonintersecting coalitions T, S ($T \subset N$, $S \subset N$).

Property (2.1) is called a *superadditivity property.* This property is necessary for the number $v(T)$ to be conceptually interpreted as a guaranteed payoff to a coalition T when this coalition is acting independently of other players. This interpretation of inequality (2.1) implies that the coalition $S \cup T$ has no less opportunities then the two nonintersecting coalitions S and T when they act independently.

From the superadditivity of v it follows that for any system of nonintersecting coalitions $S_1, \dots, S_k$ there is

$$\sum_{i=1}^{n} v(S_i) \leq v(N).$$

This, in particular, implies that there is no decomposition of the set N into coalitions such that the guaranteed total payoff to these coalitions exceeds the maximum payoff to all players $v(N)$.

Example 2. Investment fund [3]. Three investment fund managers consider investment possibilities for a year. Fund manager 1 has \$3000000 to invest, manager 2 has \$1000000 and manager 3 has \$2000000. There is an investment scheme

	Deposit	Interest rate
1	less than \$2000000	8%
2	\$2000000 up to \$5000000	9%
3	\$5000000 and more	10%

Table 1:

This situation can be readily translated into a characteristic function v for a three-player game, $N = \{1, 2, 3\}$ (in \$10000 units):

$$v(N) = 60, \quad v(\{1,2\}) = 36, \quad v(\{1,3\}) = 50, \quad v(\{2,3\}) = 27,$$

$$v(\{1\}) = 27, \quad v(\{2\}) = 8, \quad v(\{3\}) = 18.$$

It's simply to test that characteristic function $v(S)$, $S \subset N$ has superadditivity property.

Example 3. Farmer, manufacturer, subdivider [4]. A farmers land is worth \$100000 to him for agricultural use; to a manufacture it is worth \$200000 as a plant site; a subdivider would pay up to \$300000.

Denoting the farmer as player 1, the manufacture as player 2 and the subdivider as player 3, one obtains the player set $N = \{1, 2, 3\}$. The characteristic function v (in \$100000 units) follows immediately from the description of the situation:

$$v(N) = 3, \quad v(\{1,2\}) = 2, \quad v(\{1,3\}) = 3, \quad v(\{2,3\}) = 0,$$

$$v(\{1\}) = 1, \quad v(\{i\}) = 0, \text{for } i = 2, 3.$$

Example 4. Weighted majority voting [5]. Consider a group of n shareholders of a company that has to decide on some investment project. Denote by w_i the number of shares shareholder i owns. The company charter assigns one vote to each share and requires a minimum of q votes for the adoption of a project. Suppose the investment project yields a return R to be distributed equally per share.

This problem can be modeled as a game in characteristic function form with player set $N = \{1, 2, ..., n\}$ and characteristic function v:

$$v(S) = \begin{cases} \left(\sum\limits_{i \in S} w_i\right) \frac{R}{w}, & \text{if } \sum\limits_{i \in S} w_i \geq q, \\ 0, & \text{if } \sum\limits_{i \in S} w_i < q, \end{cases}$$

where $w = \sum_{i \in N} w_i$.

Example 5. Water treatment system [6]. Three communities shear a large city consider developing a water treatment system. So far these communities dispose of their sewage by sending it to a central treatment plant operated by city authorities at a monthly fee.

A cost-benefit study estimates the present value of these payments over the usual lifetime of water treatment plant at \$100 per household. To build and operate a water treatment plant for the same period is estimated at present value cost of \$500000 for up to 5000 households; \$600000 for up to 10000 households, and \$700000 for up to 15000 households.

Community 1 is estimated to serve 5000 households on average during the period under consideration, community 2 has to serve 3000 households and community 3 has 4000 households.

The decision problem of the three communities can be modeled as a game in characteristic function form with set of players $N = \{1, 2, 3\}$ and characteristic function v (measured in \$100000 units):

$$v(N) = 5, \quad v(\{1,2\}) = 2, \quad v(\{1,3\}) = 3, \quad v(\{2,3\}) = 1,$$

$$v(\{1\}) = v(\{2\}) = v(\{3\}) = 0.$$

The net benefit of a coalition is calculated as difference between the joint benefit (number of households times present value of fees saved) minus the cost of a water treatment system of appropriate size. A single community can of couse still send its sewage to the central treatment plant. This guarantees it a net benefit of zero.

Example 6. "Jazz band" game [7]. Manager of a club promises singer S, pianist P, and drummer D to pay \$100 for a joint performance. He values a singer-pianist duet at \$80, a drummer-pianist duet at \$65 and a pianist at \$30.

A singer-drummer duet may earn \$50 and a singer, on the average, \$20 for doing an evening performance. A drummer may not earn anything by playing alone.

Designating players S, P, and D by numbers 1,2,3, respectively, we are facing a cooperative game (N, v), where $N = \{1, 2, 3\}$, $v(1,2,3) = 100$, $v(1,3) = 50$, $v(1) = 20$, $v(1,2) = 80$, $v(2,3) = 65$, $v(2) = 30$, $v(3) = 0$.

We shall now consider a noncooperative game $\Gamma = (N, \{X_i\}_{i \in N}, \{H_i\}_{i \in N})$.

Suppose the players appearing in a coalition $S \subset N$ unite their efforts for the purpose of increasing their total payoff. Let us find the largest payoff they can guarantee themselves. The joint actions of the players from a coalition S imply that this coalition S acting for all its members as one player (call him Player 1) takes the set of pure strategies to be the set of all possible combinations of strategies for its constituent players from S, i.e. the elements of the Cartesian product

$$X_S = \prod_{i \in S} X_i.$$

The community of interests of the players from S means that a payoff to the coalition S (Player 1) is the sum of payoffs to the players from S, i.e.

$$H_S(x) \equiv \sum_{i \in S} H_i(x),$$

where $x \in X_N$, $x = (x_1, \ldots, x_n)$ is a situation in pure strategies.

We are interested in the largest payoff the players from S can guarantee themselves. In the worst case for Player 1, the remaining players from $N \setminus S$ may also unite into a collective Player 2 with the set of strategies $X_{N\setminus S} = \prod_{i \in N\setminus S} X_i$, where interests are diametrically opposite to those of Player 1 (i.e. Player 2's payoff at x is $-H_S(x)$). As a result of this reasoning, the question of the largest guaranteed payoff to the coalition S becomes the issue of the largest guaranteed payoff to Player 1 in the zero-sum game $\Gamma_S = (X_S, X_{N\setminus S}, H_S)$. In the mixed extension $\overline{\Gamma}_S = (\overline{X}_S, \overline{X}_{N\setminus S}, K_S)$ of the game Γ_S, the guaranteed payoff $v(S)$ to Player 1 can merely be increased in comparison with that in the game Γ_S. For this reason, the following discussion concentrates on the mixed extension of Γ_S. In particular, it should be noted that, according to this interpretation, $v(S)$ coincides with the value of the game $\overline{\Gamma}_S$ (if any), while $v(N)$ is the maximum total payoff to the players. Evidently, $v(S)$ only depends on the coalition S (and on the original noncooperative game itself, which remains unaffected in our reasoning) and is a function of S. We shall verify that this function is a characteristic function of a noncooperative game. To do this, it suffices to show that the conditions (2.1) is satisfied.

Note that $v(\oslash) = 0$ for every noncooperative game constructed above.

Lemma 3. Superadditivity. *For the noncooperative game that* $\Gamma = (N, \{X_i\}_{i \in N}, \{H_i\}_{i \in N})$, *we shall construct the function* $v(S)$ *as*

$$v(S) = \sup_{\mu_S} \inf_{\nu_{N\setminus S}} K_S(\mu_S, \nu_{N\setminus S}), \quad S \subset N, \tag{2.2}$$

where $\mu_S \in \overline{X}_S$, $\nu_{N\setminus S} \in \overline{X}_{N\setminus S}$ *and* $\overline{\Gamma}_S = (\overline{X}_S, \overline{X}_{N\setminus S}, K_S)$ *is a mixed extension of the zero-sum game* Γ_S. *Then for all* $S, T \subset N$ *for which* $S \cap T = \oslash$, *the following inequality holds:*

$$v(S \cap T) \geq v(S) + v(T). \tag{2.3}$$

Proof. Note that

$$v(S \cup T) = \sup_{\mu_{S\cup T}} \inf_{\nu_{N\setminus(S\cup T)}} \sum_{i \in S\cup T} K_i(\mu_{S\cup T}, \nu_{N\setminus(S\cup T)}),$$

where $\mu_{S\cup T}$ is the mixed strategy of coalition $S \cup T$, i.e. arbitrary probability measures on $X_{S\cup T}$, $\nu_{N\setminus(S\cup T)}$ is probability measure on $X_{N\setminus(S\cup T)}$, K_i is a payoff to player i in mixed strategies. If we restrict ourselves to those probability measures on $X_{S\cup T}$, which are the products of independent distributions μ_S and ν_T over the Cartesian product $X_S \times X_T$, then the range of the variable, in terms of which maximization is taken, shrinks and supremum merely decreases. Thus we have

$$v(S \cup T) \geq \sup_{\mu_S} \sup_{\mu_T} \inf_{\nu_{N\setminus(S\cup T)}} \sum_{i\in S\cup T} K_i(\mu_S \times \mu_T, \nu_{N\setminus(S\cup T)}).$$

Hence

$$v(S \cup T) \geq \inf_{\nu_{N\setminus(S\cup T)}} \sum_{i\in S\cup T} K_i(\mu_S \times \mu_T, \nu_{N\setminus(S\cup T)})$$

$$= \inf_{\nu_{N\setminus(S\cup T)}} \Big(\sum_{i\in S} K_i(\mu_S \times \mu_T, \nu_{N\setminus(S\cup T)}) + \sum_{i\in T} K_i(\mu_S \times \mu_T, \nu_{N\setminus(S\cup T)}) \Big).$$

Since the sum of infimums does not exceed the infimum of the sum, we have

$$v(S\cup T) \geq \inf_{\nu_{N\setminus(S\cup T)}} \sum_{i\in S} K_i(\mu_S\times\mu_T, \nu_{N\setminus(S\cup T)}) + \inf_{\nu_{N\setminus(S\cup T)}} \sum_{i\in T} K_i(\mu_S\times\mu_T, \nu_{N\setminus(S\cup T)}).$$

Minimizing the first addend on the right-hand side of the inequality over μ_T, and the second addend over μ_S (for uniformity, these will be renamed as ν_T and ν_S), we obtain

$$v(S\cup T) \geq \inf_{\nu_T} \inf_{\nu_{N\setminus(S\cup T)}} \sum_{i\in S} K_i(\mu_S\times\nu_T, \nu_{N\setminus(S\cup T)}) + \inf_{\nu_S} \inf_{\nu_{N\setminus(S\cup T)}} \sum_{i\in T} K_i(\nu_S\times\mu_T, \nu_{N\setminus(S\cup T)})$$

$$\geq \inf_{\nu_{N\setminus S}} \sum_{i\in S} K_i(\mu_S, \nu_{N\setminus S}) + \inf_{\nu_{N\setminus T}} \sum_{i\in T} K_i(\mu_T, \nu_{N\setminus T}).$$

The last inequality holds for any values of measures μ_S and μ_T. Consequently, these make possible the passage to suprema

$$v(S \cup T) \geq \sup_{\mu_S} \inf_{\nu_{N\setminus S}} \sum_{i\in S} K_i(\mu_S, \nu_{N\setminus S}) + \sup_{\mu_T} \inf_{\nu_{N\setminus T}} \sum_{i\in T} K_i(\mu_T, \nu_{N\setminus T}),$$

whence, using (2.2), we obtain

$$v(S \cup T) \geq v(S) + v(T).$$

The superadditivity is proved.

Note that inequality (2.3) also holds if the function $v(S)$ is constructed by the rule

$$v(S) = \sup_{x_S} \inf_{x_{N\setminus S}} H_S(x_S, x_{N\setminus S}), \quad S \subset N,$$

where $x_S \in X_S$, $x_{N\setminus S} \in X_{N\setminus S}$, $\Gamma_S = (X_S, X_{N\setminus S}, H_S)$. In this case, the proof literarily repeats the one given above.

Definition 9. *The noncooperative game* $\Gamma = (N, \{X_i\}_{i\in N}, \{H_i\}_{i\in N})$ *is called a constant sum game if*

$$\sum_{i\in N} H_i(x) = c = const$$

for all $x \in X_N$, $X_N = \prod_{i\in N} X_i$.

Lemma 4. *Let $\Gamma = (N, \{X_i\}_{i\in N}, \{H_i\}_{i\in N})$ be a noncooperative constant sum game, the function $v(S)$, $S \subset N$, be defined as in Lemma 3, and the games Γ_S, $S \subset N$, have the values in mixed strategies. Then*

$$v(N) = v(S) + v(N \setminus S), \quad S \subset N.$$

Proof. The definition of the constant sum game implies that

$$v(N) = \sum_{i\in N} H_i(x) = \sum_{i\in N} K_i(\mu) = c$$

for all situations x in pure strategies and all situations μ in mixed strategies.

On the other hand,

$$v(S) = \sup_{\mu_S} \inf_{\nu_{N\setminus S}} \sum_{i\in S} K_i(\mu_S, \nu_{N\setminus S}) = \sup_{\mu_S} \inf_{\nu_{N\setminus S}} \left(c - \sum_{i\in N\setminus S} K_i(\mu_S, \nu_{N\setminus S})\right)$$

$$= c - \inf_{\nu_{N\setminus S}} \sup_{\mu_S} \sum_{i\in N\setminus S} K_i(\mu_S, \nu_{N\setminus S}) = c - v(N \setminus S),$$

which is what we set out to prove.

In what follows, by the cooperative game is meant a pair (N, v), where v is the characteristic function satisfying inequality (2.1). The conceptual interpretation of the characteristic function justifying property (2.1) is not essential for what follows (see examples 2–6).

The main problem in the cooperative theory of n-person games is to construct realizable principles for optimal distribution of a maximum total payoff $v(N)$ among players.

Let α_i be an amount the player i receives by distribution of maximum total payoff $v(N)$, $N = \{1, 2, \ldots, n\}$.

Definition 10. *The vector $\alpha = (\alpha_1, \ldots, \alpha_n)$, which satisfies the conditions*

$$\alpha_i \geq v(\{i\}), \ i \in N, \tag{2.4}$$

$$\sum_{i=1}^{n} \alpha_i = v(N), \tag{2.5}$$

where $v(\{i\})$ is the value of characteristic function for a single-element coalition $S = \{i\}$ is called an imputation.

Condition (2.4) is called an individual rationality condition and implies that every member of coalition received at least the same amount he could ensure by acting alone, without any support of other players. Furthermore, condition (2.5) must be satisfied, since in the case $\sum_{i\in N} \alpha_i < v(N)$ there is a distribution α', on which every player $i \in N$ receives more than his share α_i. However, if $\sum_{i\in N} \alpha_i > v(N)$, then players from N distribute among themselves an unrealized payoff. For this reason, the vector α can be taken to be admissible only if condition (2.5) is satisfied. This condition is called a collective (or group) rationality condition.

By (2.4), (2.5), for the vector $\alpha = (\alpha_1, \dots, \alpha_n)$ to be an imputation in the cooperative game (N, v), it is necessary and sufficient that it could be represented as

$$\alpha_i = v(\{i\}) + \gamma_i, \quad i \in N,$$

and

$$\gamma_i \geq 0,\ i \in N, \quad \sum_{i \in N} \gamma_i = v(N) - \sum_{i \in N} v(\{i\}).$$

Definition 11. *The game (N, v) is called essential if*

$$\sum_{i \in N} v(\{i\}) < v(N), \tag{2.6}$$

otherwise it is called nonessential.

For any imputation α, we denote the quantity $\sum_{i \in S} \alpha_i$ by $\alpha(S)$ and the set of all imputations by D. The nonessential game has a unique imputation $\alpha = (v(\{1\}), v(\{2\}), \dots, v(\{n\}))$.

In any essential game with more than one player, the imputation set is infinite. We shall examine such games by using a dominance relation.

Definition 12. *Imputation α dominates imputation β in coalition S (denoted as $\alpha \overset{S}{\succ} \beta$) if*

$$\alpha_i > \beta_i,\ i \in S, \quad \alpha(S) \leq v(S). \tag{2.7}$$

The first condition in (2.7) implies that imputation α is more advantageous to all members of coalition S than imputation β, while the second condition accounts for the fact that imputation α can be realized by coalition S (that is, coalition S can actually offer an amount α_i to every player $i \in S$).

Definition 13. *Imputation α is said to dominanate imputation β if there is a coalition S for which $\alpha \overset{S}{\succ} \beta$. Dominance of imputation β by imputation α is denoted as $\alpha \succ \beta$.*

Dominance is not possible in a single element coalition and in the set of all players N. Indeed, $\alpha \overset{i}{\succ} \beta$ had to imply $\beta_i < \alpha_i \leq v(\{i\})$ which contradicts condition (2.5).

Combining cooperative games into one or another class may substantially simplify their subsequent examination. We may examine equivalency classes of games.

Definition 14. *The cooperative game (N, v) is called equivalent to the game (N, v') if there exists a positive number k and n arbitrary real numbers c_i, $i \in N$, such that for any coalition $S \subset N$ there is*

$$v'(S) = kv(S) + \sum_{i \in S} c_i. \tag{2.8}$$

The equivalence of the game (N, v) to (N, v') will be denoted as $(N, v) \sim (N, v')$ or $v \sim v'$.

It is obvious that $v \sim v$. This can be verified by setting $c_i = 0$, $k = 1$, $v' = v$ in (2.8). This property is called reflexity.

We shall prove the symmetry of the relation, i.e. that the condition $v \sim v'$ implies $v' \sim v$. In fact, setting $k' = 1/k$, $c'_i = -c_i/k$ we obtain

$$v(S) = k'v'(S) + \sum_{i \in S} c'_i,$$

i.e. $v' \sim v$.

Finally, if $v \sim v'$ and $v' \sim v''$, then $v \sim v''$. This property is called transitivity. This can be verified by successively applying (2.8).

Since the equivalence relation is reflexive, symmetric and transitive, it decomposes the set of all n-person games into mutually nonintersecting classes of equivalent games.

Theorem 7. *If two games v and v' are equivalent, then the map $\alpha \to \alpha'$, where*

$$\alpha'_i = k\alpha_i + c_i, \; i \in N,$$

establishes the one-to-one mapping of the set of all imputations in the game v onto the imputation set in the game v', so that $\alpha \overset{S}{\succ} \beta$ implies $\alpha' \overset{S}{\prec} \beta'$.

Proof. Let us verify that α' is an imputation in the game (N, v'). Indeed,

$$\alpha'_i = k\alpha_i + c_i \geq kv(\{i\}) + c_i = v(\{i\}),$$

$$\sum_{i \in N} \alpha'_i = \sum_{i \in N} (k\alpha_i + c_i) = kv(N) + \sum_{i \in N} c_i = v'(N).$$

It follows that conditions (2.4), (2.5) hold for α'. Furthermore, if $\alpha \overset{S}{\succ} \beta$, then

$$\alpha_i > \beta_i, \; i \in S, \quad \sum_{i \in S} \alpha_i \leq v(S),$$

and hence

$$\alpha'_i = k\alpha_i + c_i > k\beta_i + c_i = \beta'_i \; (k > 0),$$

$$\sum_{i \in S} \alpha'_i = k \sum_{i \in S} \alpha_i + \sum_{i \in S} c_i \leq kv(S) + \sum_{i \in S} c_i = v'(S),$$

i.e. $\alpha' \overset{S}{\succ} \beta'$. The one-to-one correspondence follows from the existence of the inverse mapping (it was used in the proof of the symmetry of the equivance relation). This completes the proof of the theorem.

When decomposing the set of cooperative games into mutually disjoint classes of equivalence, we are faced with the problem of choosing the simplest representative from each class.

Definition 15. *The game (N, v) is called the game in (0–1) – reduced form, if for all $i \in N$*

$$v(\{i\}) = 0, \quad v(N) = 1.$$

Theorem 8. *Every essential cooperative game is equivalent to some game in (0–1) – reduced form.*

Proof. Let

$$k = \frac{1}{v(N) - \sum_{i \in N} v(\{i\})} > 0,$$

$$c_i = -\frac{v(\{i\})}{v(N) - \sum_{i \in N} v(\{i\})}, \quad v'(S) = kv(S) + \sum_{i \in S} c_i.$$

Then $v'(\{i\}) = 0$, $v'(N) = 1$. This completes the proof of the theorem.

This theorem implies that the game theoretic properties involving the notion of dominance can be examined on the games in (0–1) - reduced form. If v is the characteristic function of an arbitrary essential game (N, v), then

$$v'(S) = \frac{v(S) - \sum_{i \in S} v(\{i\})}{v(N) - \sum_{i \in N} v(\{i\})}, \tag{2.9}$$

is (0–1) – normalization corresponding to the function v. In this case, a imputation is found to be any vector $\alpha = (\alpha_1, \dots, \alpha_n)$ whose components satisfy the conditions

$$\alpha_i \geq 0, \; i \in N, \quad \sum_{i \in N} \alpha_i = 1, \tag{2.10}$$

i.e. imputations can be regarded as the points of the $(n-1)$ – dimensional simplex generated by the unit vectors $w_j = (0, \dots, 0, 1, 0, \dots, 0)$, $j = 1, ..., n$ in the space R^n.

3 The core and NM-solution

We shall now turn to the principles of optimal behavior in cooperative games. As already previously noted, we are dealing with the principles of optimal distribution of a maximum total payoff among players.

The following approach is possible. Suppose the players in the cooperative game (N, v) have come to an agreement on distribution of a payoff to the whole coalition N (imputation α^*), under which none of the imputations dominates α^*. Then such a distribution is stable in that it is disadvantageous for any coalition S to separate from other players and distribute a payoff $v(S)$ among its members. This suggests that it may be wise to examine the set of nondominant imputations.

Definition 16. *The set of nondominant imputations in the cooperative game (N, v) is called core.*

Then we have the theorem which characterizes core.

Theorem 9. *For the imputation α to belong to core, it is necessary and sufficient that*

$$v(S) \leq \alpha(S) = \sum_{i \in S} \alpha_i \tag{3.1}$$

hold for all $S \subset N$.

Proof. This theorem is straightforward for nonessential games, and, by Theorem 8, it suffices to prove it for the games in (0–1) - reduced form.

Prove first that the statement of the theorem is sufficient. Suppose that condition (3.1) holds for the imputation α. Show that the imputation α belongs to the core. Suppose this is not so. Then there is an imputation β such that $\beta \overset{S}{\succ} \alpha$, i.e. $\beta(S) > \alpha(S)$ and $\beta(S) \leq v(S)$ which contradicts (3.1).

We shall now prove the necessity of (3.1). For any imputation α, which does not satisfy (3.1), there exists a coalition S for which $\alpha(S) < v(S)$. Let

$$\beta_i = \alpha_i + \frac{v(S) - \alpha(S)}{|S|},\ i \in S,$$

$$\beta_i = \frac{1 - v(S)}{|N| - |S|},\ i \notin S,$$

where $|S|$ is the number of elements of the set S. It can be easily seen that $\beta(N) = 1$, $\beta_i \geq 0$ and $\beta \overset{S}{\succ} \alpha$. Then it follows that α does not belong to the core.

Theorem 9 implies that core is a closed convex subset of the set of all imputations (core may also be an empty set).

Suppose the players are negotiating the choice of a cooperative agreement. It follows from the superadditivity of v that such an agreement brings about the formation of the coalition N of all players. The question is tackled as to the way of distributing the total payoff $v(N)$, i.e. the way of choosing a vector $\alpha \in R^n$ for which $\sum_{i \in N} \alpha_i = v(N)$.

The minimum requirement for obtaining the players' consent to choose a vector α is the individual rationality of this vector, i.e. the condition $\alpha_i \geq v(\{i\})$, $i \in N$. Suppose the players are negotiating the choice of the particular imputation α. Some coalition S demanding a more advantageous imputation may raise an objection against the choice of this imputation. The coalition S lays down this demand, threatening to break up general cooperation (this threat is quite real, since the payoff $v(N)$ can only be ensured by unanimous consent on the part of all players). Suppose the other players $N \setminus S$ respond to this threat by uniting their efforts against the coalition S. The maximum guaranteed payoff to the coalition S is evaluated by the number $v(S)$. Condition (3.1) implies that there exists a stabilizing threat to the coalition S from the coaltion N. Thus, a core of the game (N, v) is the set of distributions of the maximum total payoff $v(N)$ which is immune to cooperative threats.

We shall bring forward one more criterion to judge whether an imputation belongs to the core.

Lemma 5. *Let α be an imputation in the game (N, v). Then α belongs to the core if and only if the inequality*

$$\sum_{i\in S}\alpha_i \le v(N) - v(N\setminus S) \tag{3.2}$$

holds for all coalitions $S\subset N$.

Proof. Since $\sum_{i\in N}\alpha_i = v(N)$, the above inequality can be written as

$$v(N\setminus S) \le \sum_{i\in N\setminus S}\alpha_i.$$

Now the assertion of the lemma follows from (3.1).

Condition (3.1) shows that if the imputation α belongs to the core, then no coalition S can guarantee itself the amount exceeding $\sum_{i\in S}\alpha_i = \alpha(S)$, i.e. the total payoff ensured by the coalition members using the imputation α. This makes unreasonable the existence of coalitions S other than the maximal coalition N.

Theorem 9 provides enough reason to use core as an important optimality principle in the cooperative theory. However, in many cases the core appears to be empty, whereas in the other cases it represents a multiple optimality principle and the question as to which of the imputation are to be chosen from the core in the particular case is still open.

Example 7. Consider the "jazz band" game. The total receipts of three musicians is maximum (\$100) when they do performance jointly. If the singer does performance separately from the pianist and drummer, they receive \$65+\$20 all together. If the pianist does performance separately from the singer and drummer, they receive \$30+\$50 all together. Finally, if the pianist and singer do performance without the drummer, their total receipts amount to \$80. What is the distribution of the maximum total receipts to be considered rational in terms of the above-mentioned partial cooperation and individual behavior?

The vector $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ in the "jazz band" game belongs to the core if and only if

$$\begin{cases} \alpha_1 \ge 20,\ \alpha_2 \ge 30,\ \alpha_3 \ge 0, \\ \alpha_1+\alpha_2+\alpha_3 = 100, \\ \alpha_1+\alpha_2 \ge 80,\ \alpha_2+\alpha_3 \ge 65,\ \alpha_1+\alpha_3 \ge 50. \end{cases}$$

This set is a convex hull of the following three imputations: (35,45,20), (35,50,15), (30,50,20). Thus, the payoffs of each player in different imputations differs on the amount not more than 5 rubles. The typical representative of the core is the arithmetical mean of extreme points of core, namely $\alpha^* = (33.3, 48.3, 18.3)$. The characteristic feature of the imputation a^* is that all two-component coalitions have the same additional receipts: $\alpha_i+\alpha_j-v(\{i,j\}) = 1.6$. The imputation α^* is a fair compromise from the interior of the core.

Example 8. Recall the characteristic function of example 4:

$$v(N) = 3, \quad v(\{1,2\}) = 2, \quad v(\{1,3\}) = 3, \quad v(\{2,3\}) = 0,$$

$$v(\{1\}) = 1, \quad v(\{i\}) = 0, \text{for } i = 2, 3.$$

After substituting the $\alpha_3 = 3 - \alpha_1 - \alpha_2$, the core of this game is given by the following inequalities:

$$\alpha_1 + \alpha_2 \geq 2, \quad \alpha_2 \leq 0, \quad \alpha_1 \leq 3, \quad \alpha_1 \geq 1, \quad \alpha_2 \geq 0, \quad \alpha_1 + \alpha_2 \leq 3.$$

Obviously, player 2 has no power in this example.

Each imputation in the core will assign zero to player 2. Player 1, however, plays a major role, since the coalition of player 3 with player 1 guarantees player 1 at least 2. Hence, $C(v) = \{(\alpha, 0, 3 - \alpha) | \alpha \in [2, 3]\}$ is the set of core imputations.

Example 9. In the problem of cost sharing in the example 5 the following characteristic function v can be established:

$$v(N) = 5, \quad v(\{1, 2\}) = 2, \quad v(\{1, 3\}) = 3, \quad v(\{2, 3\}) = 1,$$

$$v(\{1\}) = v(\{2\}) = v(\{3\}) = 0.$$

Substituting $\alpha_3 = 5 - \alpha_1 - \alpha_2$ into inequalities (3.1) defining the core, one derives the following inequalities:

$$\alpha_1 + \alpha_2 \geq 2, \quad \alpha_2 \leq 2, \quad \alpha_1 \leq 4, \quad \alpha_1 \geq 0, \quad \alpha_2 \geq 0, \quad \alpha_1 + \alpha_2 \leq 5.$$

It's clear, that the core of the problem is not empty.

The fact that the core is empty does not mean that the cooperation of all players N is impossible. This simply means that no imputation can be stabilized with the help of simple threats as above. The kernel is empty when intermediate coalitions are too strong. This assertion can be explained as follows.

Example 10. Symmetric games [7]. In a symmetric game, coalitions with the same number of players have the same payoffs. The characteristic function v is

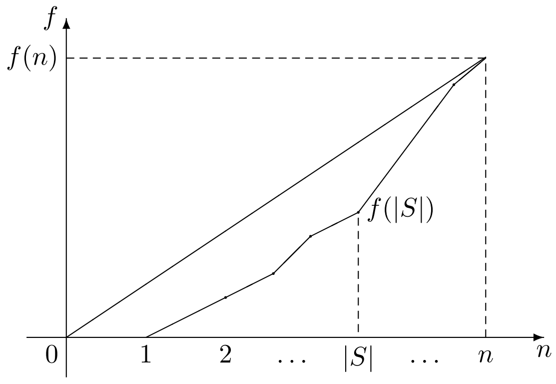
$$v(S) = f(|S|)$$

for all $S \subset N$, where $|S|$ is the number of elements of the set S.

We may assume, without loss of generality, that $f(1) = 0$ and $N = \{1, \ldots, n\}$. Then the imputation set in the game (N, v) is the following simplex in R^n

$$\sum_{i=1}^{n} \alpha_i = f(n) = v(N), \ \alpha_i \geq 0, \ i = 1, \ldots, n.$$

The core is a subset of the imputation set defined by linear inequalities (3.1), i.e. a convex polyhedron. By the symmetry of $v(S)$, the core is also symmetric, i.e. invariant under any permutation of components $\alpha_1, \ldots, \alpha_n$. Furthermore, by the convexity of the core, it can be shown that the core is nonempty if and only if it contains the center α^* of the set of all distributions ($\alpha_i^* = f(n)/n$, $i = 1, \ldots, n$). Returning to system (3.1), we see that the core is nonempty if and only if the inequality $(1/|S|)f(|S|) \leq (1/n)f(n)$ holds for all $|S| = 1, \ldots, n$. Thus, the core is nonempty if and only if there is no intermediate coalition S,



Plot 5.

in which the average payment to each player exceeds the corresponding amount in the coalition N. Plot 5 (Plot 6) corresponds to the case where the core is nonempty (empty).

Example 11 [8]. Consider a general three-person game in (0–1) - reduced form. For its characteristic function we have $v(\oslash) = v(1) = v(2) = v(3) = 0$, $v(1,2,3) = 1$, $v(1,2) = c_3$, $v(1,3) = c_2$, $v(2,3) = c_1$, where $0 \le c_i \le 1$, $i = 1,2,3$. By the Theorem 9, for the imputation α to belong to the core, it is necessary and sufficient that

$$\alpha_1 + \alpha_2 \ge c_3,\ \alpha_1 + \alpha_3 \ge c_2,\ \alpha_2 + \alpha_3 \ge c_1$$

or

$$\alpha_3 \le 1 - c_3,\ \alpha_2 \le 1 - c_2,\ \alpha_1 \le 1 - c_1. \tag{3.3}$$

Summing inequalities (3.3) we obtain

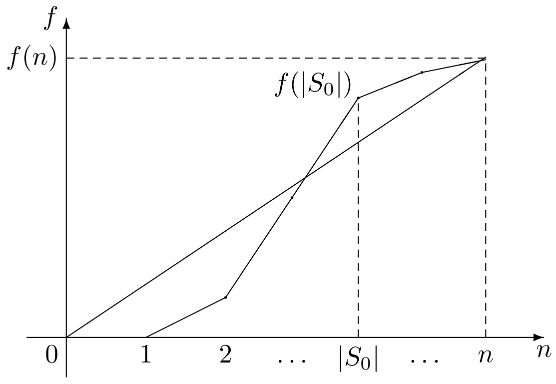
$$\alpha_1 + \alpha_2 + \alpha_3 \le 3 - (c_1 + c_2 + c_3),$$

or, since the sum of all α_i, $i = 1,2,3$, is identically equal to 1,

$$c_1 + c_2 + c_3 \le 2. \tag{3.4}$$

The last inequality is the necessary condition for the existence of a nonempty core in the game of interest. On the other hand, if (3.4) is satisfied, then there are non-negative ξ_1, ξ_2, ξ_3 such that

$$\sum_{i=1}^{3} (c_i + \xi_i) = 2,\ c_i + \xi_i \le 1,\ i = 1,2,3.$$



Plot 6.

Let $\beta_i = 1 - c_i - \xi_i$, $i = 1, 2, 3$. The numbers β_i satisfy inequalities (3.3) in such a way that the imputation $\beta = (\beta_1, \beta_2, \beta_3)$ belongs to the core of the game $(\sum_{i=1}^{3} \beta_i = 1)$; hence relation (3.4) is also sufficient for a nonempty core to exist.

Geometrically, the imputation set in the game involved is the simplex: $\alpha_1 + \alpha_2 + \alpha_3 = 1$, $\alpha_i \geq 0$, $i = 1, 2, 3$ (triangle ABC shown in Plot 7). The nonempty core is an intersection of the imputation set ($\triangle ABC$) and a convex polyhedron (parallelepiped) $0 \leq \alpha_i \leq 1 - c_i$, $i = 1, 2, 3$. It is the part of triangle ABC cut out by the lines of intersections of the planes

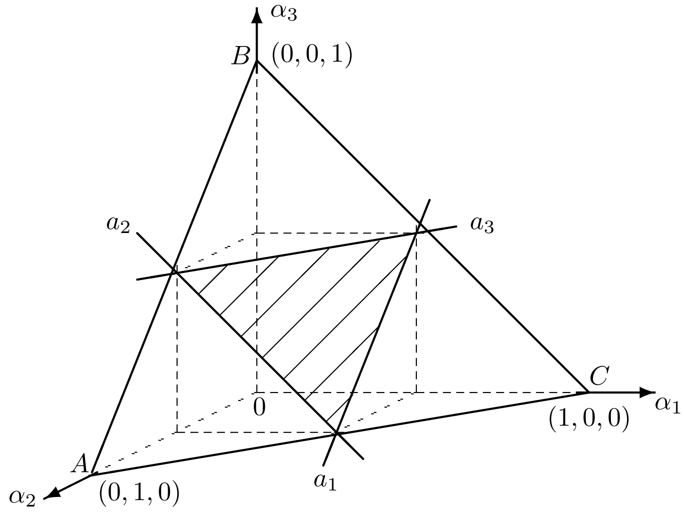
$$\alpha_i = 1 - c_i, \ i = 1, 2, 3 \tag{3.5}$$

with the plane $\triangle ABC$. We have α_i, $i = 1, 2, 3$ standing for the line formed by intersection of the planes $\alpha_i = 1 - c_i$ and $\alpha_1 + \alpha_2 + \alpha_3 = 1$. The intersection point of two lines, α_i and α_j, belongs to triangle ABC if the kth coordinate of this point, with ($k \neq i$, $k \neq j$), is non-negative; otherwise it is outside of $\triangle ABC$ (Plot 8a and 8b). Thus, the core has the form of a triangle if a joint solution to any pair of equations (3.5) and equation $\alpha_1 + \alpha_2 + \alpha_3 = 1$ is non-negative. This requirement holds for

$$c_1 + c_2 \geq 1, \ c_1 + c_3 \geq 1, \ c_2 + c_3 \geq 1. \tag{3.6}$$

The core can take one or another form, as the case requires (whereas a total of eight is possible here). For example, if none of the three inequalities (3.6) is satisfied, then the core appears to be a hexagon (Plot 8b).

Another optimality principle in cooperative games is NM-solution, which is actually a multiple optimality principle in the set of all imputations, as also is the core. Although the elements of the core are not dominated by other imputations, but we cannot say that for any previously given imputation α



Plot 7.

in the core there is its associated dominating imputation. For this reason, it seems to be wise to formulate an optimality principle by taking into account this situation.

Definition 17. *The imputation set L in the cooperative game (N, v) is called NM-solution if:*

1) $\alpha \succ \beta$ implies that either $\alpha \notin L$ or $\beta \notin L$ (interior stability);

2) for any $\alpha \notin L$ there is an imputation $\beta \in L$ such that $\beta \succ \alpha$ (exterior stability).

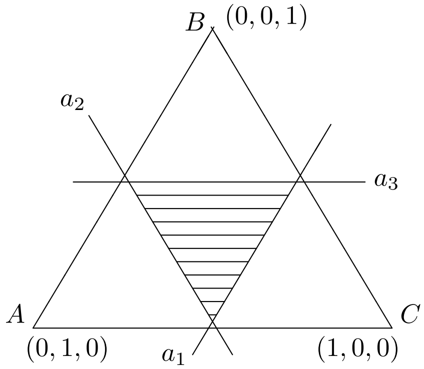
Unfortunately, the definition is not constructive and thus the notion of NM-solution cannot find practical use, and has more philosophical significance rather than a practical meaning.

There is a particular relation between the core in a cooperative game and its NM-solution. For example, if the core is nonempty and NM-solution exists, then it contains the core. Suppose the imputation α belongs to the core. In fact, if it did not belong to NM-solution L, then, by property 2, there would be an imputation α' such that $\alpha' \succ \alpha$. This, however, contradicts the fact that α belongs to the core as a set of nondominant imputations.

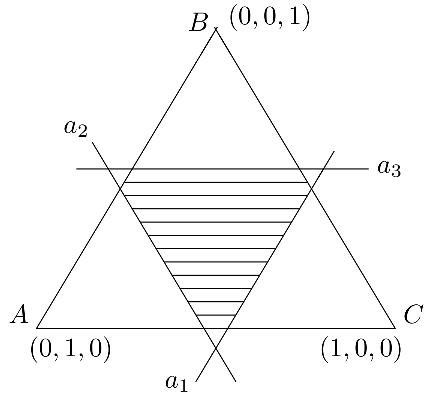
Theorem 10. *If the inequalities*

$$v(S) \leq \frac{1}{n - |S| + 1},$$

where $|S|$ is the number of players in coalition S, hold for the characteristic function of the game (N, v) in (0–1) - reduced form ($|N| = n$), then the core of this game is nonempty, and is its NM-solution.



a)



b)

Plot 8.

Proof. Take an arbitrary imputation α which is exterior to the core. Then there exists a nonempty set of those coalitions S in which it is possible to dominate α, i.e. these are the coalitions for which $\alpha(S) < v(S)$. The set $\{S\}$ is partially ordered in the inclusion, i.e. $S_1 > S_2$ if $S_2 \subset S_1$. Take in it a minimal element S_0 which apparently exists. Let k be the number of players in the coalition S_0. Evidently, $2 \le k \le n-1$. Let us construct the imputation β as follows:

$$\beta_i = \begin{cases} \alpha_i + \dfrac{v(S_0) - \alpha(S_0)}{k}, & i \in S_0, \\ \dfrac{1 - v(S_0)}{n-k}, & i \notin S_0. \end{cases}$$

Since $\beta(S_0) = v(S_0)$, $\beta_i > \alpha_i$, $i \in S_0$, then β dominates α in the coalition S_0. Show that β is contained in the core. To do this, it suffices to show that $\beta(S) \ge v(S)$ for an arbitrary S. At first, let $|S| \le k$. Note that β is not dominated for any coalition $S \subset S_0$, since $\beta_i > \alpha_i$ $(i \in S_0)$, while S_0 is a minimal coalition for which it is possible to dominate α. If, however, at least one player from S is not contained in S_0, then

$$\beta(S) \ge \frac{1 - v(S_0)}{n-k} \ge \frac{1 - \frac{1}{n-k+1}}{n-k} = \frac{1}{n-k+1} \ge \frac{1}{n-|S|+1} \ge v(S).$$

Thus, β is not dominated for any coalition containing at most k players.

Now, let $|S| > k$. If $S_0 \subset S$, then

$$\beta(S) = \frac{(|S|-k)(1 - v(S_0))}{n-k} + v(S_0) \ge \frac{|S|-k}{n-k}$$

$$\ge \frac{|S|-k+k-|S|+1}{n-k+k-|S|+1} = \frac{1}{n-|S|+1} \ge v(S).$$

However, if S does not contain S_0, then the number of players of the set S, not contained in S_0, is at least $|S|-k+1$; hence

$$\beta(S) \geq \frac{(|S|-k+1)(1-v(S_0))}{n-k} \geq \frac{|S|-k+1}{n-k+1} \geq \frac{1}{n-|S|+1} \geq v(S).$$

Thus, β is not dominated for any one of the coalitions S. Therefore, β is contained in the core. Furthermore, β dominates α. We have thus proved that the core is nonempty and satisfies property 2 which characterizes the set of NM-solutions. By definition, the core satisfies property 1 automatically. This completes the proof of the theorem.

Definition 18. *The game (N, v) in (0–1) - reduced form is called simple if for any $S \subset N$ $v(S)$ takes only one of the two values, 0 or 1. A cooperative game is called simple if its (0–1) - reduced form is simple.*

Example 12 [8]. Consider a three-person simple game in (0–1) - reduced form, in which the coalition composed of two or three players wins ($v(S) = 1$), while the one-player coalition loses ($v(\{i\}) = 0$). For this game, we consider three imputations

$$\alpha_{12} = (1/2, 1/2, 0),\ \alpha_{13} = (1/2, 0, 1/2),\ \alpha_{23} = (0, 1/2, 1/2). \tag{3.7}$$

None of the three imputations dominates each other. The imputation set (3.7) also has the property as follows: any imputation (except for three imputations α_{ij}) is dominated by one of the imputations α_{ij}. This can be verified by examining some imputation $\alpha = (\alpha_1, \alpha_2, \alpha_3)$. Since we are examining the game in (0–1) - reduced form, then $\alpha_i \geq 0$ and $\alpha_1 + \alpha_2 + \alpha_3 = 1$. Therefore, no more than two components of the vector α can be at least 1/2. If there are actually two components, then each of them is 1/2, whereas the third component is 0. But this means that α coincides with one of α_{ij}. However, if α is some other imputation, then it has no more than one component which is at least 1/2. We thus have at least two components, say, α_i and α_j ($i < j$), which are less than 1/2. But here $\alpha_{ij} \overset{ij}{\succ} \alpha$. Now three imputations (3.7) form NM-solution. But this is not the only NM-solution.

Let c be any number from the interval $[0, 1/2]$. It can be easily verified that the set

$$L_{3,c} = \{(a, 1-c-a, c) \mid 0 \leq a \leq 1-c\}$$

also is NM-solution. Indeed, this set contains imputations, on which Player 3 receives a constant c, while the players 1 and 2 divide the remaining part in all possible proportions. Internal stability follows from the fact that for any two imputations α and β from this set we have: if $\alpha_1 > \beta_1$, then $\alpha_2 < \beta_2$. But dominance for a single player coalition is not possible. To prove the external stability $L_{3,c}$, we may take any imputation $\beta \notin L_{3,c}$. This means that either $\beta_3 > c$ or $\beta_3 < c$. Let $\beta_3 > c$, e.g., $\beta_3 = c + \epsilon$. Define the imputation α as follows:

$$\alpha_1 = \beta_1 + \epsilon/2,\ \alpha_2 = \beta_2 + \epsilon/2,\ \alpha_3 = c.$$

Then $\alpha \in L_{3,c}$ and $\alpha \succ \beta$ for coalition $\{1, 2\}$. Now, let $\beta_3 < c$. It is clear that either $\beta_1 \leq 1/2$ or $\beta_2 \leq 1/2$ (otherwise their sum would be greater than 1). Let $\beta_1 \leq 1/2$. Set $\alpha = (1 - c, 0, c)$. Since $1 - c > 1/2 \geq \beta_1$, then $\alpha \succ \beta$ for coalition {1,3}. Evidently $\alpha \in L_{3,c}$. However, if $\beta_2 \leq 1/2$, then we may show in a similar manner that $\gamma \succ \beta$, where $\gamma = (0, 1 - c, c)$. Now, aside from the symmetric NM-solution, the game involved has the whole family of solutions which allow Player 3 to obtain a fixed amount c from the interval $0 \leq c \leq 1/2$. These NM-solutions are called discriminant. In the case of the set $L_{3,0}$ Player 3 is said to be completely discriminated or excluded.

From symmetry it follows that there are also two families of NM-solutions, $L_{1,c}$ and $L_{2,c}$, which discriminate Player 1 and 2, respectively.

The preceding example shows that the game may have many NM-solutions. It is not clear which of them is to be chosen. If, however, one NM-solution has been chosen, it remains unclear which of the imputations is to be chosen from this particular solution.

Although the existence of NM-solutions in the general case has not been proved, some special results have been obtained. Some of them are concerned with the existence of NM-solutions, while the others are related to the existence of NM-solutions of a particular type [9].

4 The Shapley value

The multiplicity of the previously discussed optimality principles (core and NM-solution) in cooperative games and the rigid conditions on the existence of these principles force us to a search for the principles of optimality, existence and uniqueness of which may be ensured in every cooperative game. Among such optimality principles is Shapley value. The Shapley value is defined axiomatically

Definition 19. *The carrier of the game (N, v) is called a coalition T such that $v(S) = v(S \cap T)$ for any coalition $S \subset N$.*

Conceptually, this definition states that any player, not a member of the carrier, is a "dummy", i.e. he has nothing to contribute to any one of the coalitions.

We shall consider an arbitrary permutation P of the ordered set of players $N = \{1, 2, \ldots, n\}$. This permutation has associated with the substitution π, i.e. one-to-one function $\pi : N \to N$ such that for $i \in N$ the value $\pi(i) \in N$ is an element of N to which $i \in N$ changes in a permutation P.

Definition 20. *Suppose that (N, v) is an n-person game, P is a permutation of the set N and π is its associated substitution. Denote by $(N, \pi v)$ a game (N, u) such that for any coalition $S \subset N$, $S = \{i_1, i_2, \ldots, i_s\}$*

$$u(\{\pi(i_1), \pi(i_2), \ldots, \pi(i_s)\}) = v(S).$$

The game $(N, \pi v)$ and the game (N, v) differ only in that in the latter the players exchange their roles in accordance with permutation P.

The definition permit the presentation of Shapley axiomatics. First, note that since cooperative n-person games are essentially identified with real-valued (characteristic) functions, we may deal with the sum of two or more games and the product of game by number.

We shall set up a correspondence between every cooperative game (N, v) and the vector $\varphi(v) = (\varphi_1[v], \ldots, \varphi_n[v])$ whose components are interpreted to mean the payoffs received by players under an agreement or an arbitration award. Here, this correspondence is taken to satisfy the following axioms.

Definition 21. *Shapley axioms.*

1. *If S is any carrier of the game (N, v), then*

$$\sum_{i \in S} \varphi_i[v] = v(S).$$

2. *For any substitution of π and $i \in N$*

$$\varphi_{\pi(i)}[\pi v] = \varphi_i[v].$$

3. *If (N, u) and (N, v) are any cooperative games, then*

$$\varphi_i[u + v] = \varphi_i[u] + \varphi_i[v].$$

Definition 22. *Suppose φ is the function which, by axioms 1–3, sets up a correspondence between every game (N, v) and the vector $\varphi[v]$. Then $\varphi[v]$ is called the vector of values or the Shapley value of the game (N, v).*

It turns out that these axioms suffice to define uniquely values for all n-person games.

Theorem 11. *There exists a unique function φ which is defined for all games (N, v) and satisfies axioms 1–3.*

The proof of the theorem is based on the following results.

Lemma 6. *Let the game (N, w_S) be defined for any coalition $S \subset N$ as follows:*

$$w_S(T) = \begin{cases} 0, & S \not\subset T, \\ 1, & S \subset T. \end{cases} \tag{4.1}$$

Then for the game (N, w_S) the vector $\varphi[w_S]$ is uniquely defined by axioms 1,2:

$$\varphi_i[w_S] = \begin{cases} 1/s, & i \in S, \\ 0, & i \notin S, \end{cases} \tag{4.2}$$

where $s = |S|$ is the number of players in S.

Proof. It is obvious that S is the carrier of w_S, as is any set T containing the set S. Now, by axiom 1, if $S \subset T$, then

$$\sum_{i \in T} \varphi_i[w_S] = 1.$$

But this means that $\varphi_i[w_S] = 0$ for $i \notin S$. Further, if π is any substitution which converts S to itself, then $\pi w_S = w_S$. Therefore, by axiom 2, for any $i, j \in S$ there is the equality $\varphi_i[w_S] = \varphi_j[w_S]$. Since there is a total of $s = |S|$ and their sum is 1, we have $\varphi_i[w_S] = 1/s$ if $i \in S$.

The game with the characteristic function w_S defined by (4.1) is called the *simple n-person game*. Now the lemma states that for the simple game (N, w_S) the Shapley value for the game (N, w_S) is determined in a unique manner.

Corollary 1. *If $c \geq 0$, then*

$$\varphi_i[cw_S] = \begin{cases} c/s, & i \in S, \\ 0, & i \notin S. \end{cases}$$

The proof is straightforward. Thus $\varphi[cw_S] = c\varphi[w_S]$ for $c \geq 0$.

We shall now show that if $\sum_S c_S w_S$ is a characteristic function, then

$$\varphi_i\left(\sum_S c_S w_S\right) = \sum_S \varphi_i(c_S w_S) = \sum_S c_S \varphi_i(w_S). \tag{4.3}$$

In the case of $c_S \geq 0$, the first equation in (4.3) is stated by axiom 3, while the second follows from the corollary. Further, if u, v and $u - v$ are characteristic functions, then, by axiom 3, $\varphi[u - v] = \varphi[u] - \varphi[v]$. Hence it follows that (4.3) holds for any c_S. Indeed, if $\sum_S c_S w_S$ is a characteristic function, then

$$v = \sum_S c_S w_S = \sum_{S | c_S \geq 0} c_S w_S - \sum_{S | c_S < 0} (-c_S) w_S,$$

hence

$$\varphi[v] = \varphi[\sum_{S | c_S \geq 0} c_S w_S] - \varphi[\sum_{S | c_S < 0} (-c_S) w_S]$$

$$= \sum_{S | c_S \geq 0} c_S \varphi[w_S] - \sum_{S | c_S < 0} (-c_S) \varphi[w_S] = \sum_S c_S \varphi[w_S].$$

Lemma 7. *Let (N, v) be any game. Then there are $2^n - 1$ real numbers c_S such that*

$$v = \sum_{S \subset N} c_S w_S, \tag{4.4}$$

where w_S are defined by (4.1) and summation is made over all subsets S of the set N, exclusive of an empty set. Here, representation (4.4) is unique.

Proof. Set

$$c_S = \sum_{T|T\subset S} (-1)^{s-t} v(T) \tag{4.5}$$

(here t is the number of elements in T). Show that these numbers c_S satisfy the conditions of the lemma. Indeed, if U is an arbitrary coalition, then

$$\sum_{S|S\subset N} c_S w_S(U) = \sum_{S|S\subset U} c_S$$

$$= \sum_{S|S\subset U} \left(\sum_{T|T\subset S} (-1)^{s-t} v(T) \right) = \sum_{T|T\subset U} \left[\sum_{S|T\subset S\subset U} (-1)^{s-t} \right] v(T).$$

We shall now consider the quantity which is bracketed in the last expression. For every value s between t and u there are C_{u-t}^{u-s} of sets S with s-elements such that $T \subset S \subset U$. Therefore the bracketed expression can be replaced by the following:

$$\sum_{s=t}^{u} C_{u-t}^{u-s} (-1)^{s-t} = \sum_{s=t}^{u} C_{u-t}^{s-t} (-1)^{s-t},$$

but this is a binomial decomposition of $(1-1)^{u-t}$; hence it is 0 for all $t < u$, and 1 for $t = u$. Therefore for all $U \subset N$

$$\sum_{S|S\subset N} c_S w_S(U) = v(U).$$

We shall prove the uniqueness of representation (4.4). To every characteristic function v corresponds an element in the space R^{2^n-1}. Now, let us order coalitions $T \subset U$. Then for every nonempty coalition $T \subset U$ corresponds component of the vector equal to $v(T)$. These vectors will be denoted by v as functions. It is obvious that to the simple characteristic functions w_S correspond the vectors whose components are 0 or 1. We shall prove that the simple characteristic functions (or, more specifically, their associated vectors) are linearly independent. Indeed, let

$$\sum_{S\subset N} \lambda_S w_S(T) = 0 \text{ for all } T \subset N.$$

Then for all $T = \{i\}$ we have $w_S(\{i\}) = 0$ if $S \neq \{i\}$, and $w_S(\{i\}) = 1$ if $S = \{i\}$. Hence $\lambda_{\{i\}} = 0$ for all $i \subset N$. Continue the proof by using the induction method. Let $\lambda_S = 0$ for all $S \subset T$, $S \neq T$. Show that $\lambda_T = 0$. Indeed,

$$\sum_{S\subset N} \lambda_S w_S(T) = \sum_{S\subset T} \lambda_S w_S(T) = \lambda_T = 0.$$

Now, we have $2^n - 1$ linearly independent vectors in R^{2^n-1}; therefore every vector, and hence every characteristic function v is uniquely expressed as a linear combination (4.4) of the simple characteristic functions w_S. This completes the proof of the lemma.

We shall now turn to the proof of Theorem 11. Lemma 7 shows that any game can be represented as a linear combination of games w_S and the representation (4.4) is unique. The function $\varphi[v]$ is uniquely defined by relations (4.3), (4.2).

Let (N, v) be an arbitrary game. We now obtain an expression for the vector $\varphi[v]$. By 4.3, 4.4,

$$\varphi_i[v] = \sum_{S|S\subset N} c_S \varphi_i[w_S] = \sum_{S|i\subset S\subset N} c_S(1/s),$$

but c_S are determined by (4.5). Substituting (4.5) into this expression we obtain

$$\varphi_i[v] = \sum_{S|i\subset S\subset N}(1/s)\Bigg[\sum_{T|T\subset S}(-1)^{s-t}v(T)\Bigg] = \sum_{T|T\subset N}\Bigg[\sum_{S|T\cup i\subset S\subset N}(-1)^{s-t}(1/s)v(T)\Bigg].$$

Set

$$\gamma_i(T) \equiv \sum_{S|T\cup i\subset S\subset N}(-1)^{s-t}(1/s). \tag{4.6}$$

If $i \notin T'$ and $T = T' \cup \{i\}$, then $\gamma_i(T') = -\gamma_i(T)$. In fact, all terms on the right-hand side of (4.6) in both cases are the same and only $t = t' + 1$; hence they differ only in sign. Thus we have

$$\varphi_i[v] = \sum_{T|i\in T\subset N}\gamma_i(T)[v(T) - v(T \setminus \{i\})].$$

Further, if $i \in T$, then there are exactly C_{n-t}^{s-t} coalitions S with s-elements such that $T \subset S$. This brings us to the well known definite integral

$$\gamma_i(T) = \sum_{s=t}^{n}(-1)^{s-t}C_{n-t}^{s-t}(1/s) = \sum_{s=t}^{n}(-1)^{s-t}C_{n-t}^{s-t}\int_0^1 x^{s-1}dx$$

$$= \int_0^1 \sum_{s=t}^{n}(-1)^{s-t}C_{n-t}^{s-t}x^{s-1}dx = \int_0^1 x^{t-1}\sum_{s=t}^{n}(-1)^{s-t}C_{n-t}^{s-t}x^{s-t}dx$$

$$= \int_0^1 x^{t-1}(1-x)^{n-t}dx.$$

Thus we have

$$\gamma_i(T) = (t-1)!(n-t)!/(n!)$$

and hence

$$\varphi_i[v] = \sum_{T|i\in T\subset N}\frac{(t-1)!(n-t)!}{n!}[v(T) - v(T \setminus \{i\})]. \tag{4.7}$$

Equation (4.7) determines explicitly the components of Shapley value. This expression satisfies axioms 1–3.

Note that the vector $\varphi[v]$ is an imputation. Indeed, by the superadditivity of the function v,

$$\varphi_i[v] \geq v(\{i\}) \sum_{T|i\subset T\subset N} \frac{(t-1)!(n-t)!}{n!}$$

$$= v(\{i\}) \sum_{t=1}^{n} C_{n-1}^{t-1} \frac{(t-1)!(n-t)!}{n!} = v(\{i\}).$$

Axiomatic definition apart, the Shapley value expressed by (4.7) can be interpreted conceptually as follows. Suppose the players (elements of the set N) have decided to meet in a specified place at a specified time. It would appear natural that, because of random deviations, they would arrive at various instants of time. However, it is assumed that the players' arrival orders (i.e. their permutations) have the same probability, namely $1/(n!)$. Suppose that if, on arrival, player i finds in place (only) the members of coalition $T \setminus \{i\}$, then he receives a payoff $v(T) - v(T \setminus \{i\})$, that is, the amount he contributes to that coalition. Then the component of Shapley value $\varphi_i[v]$ represents the mathematical expectation of player i's contribution to randomly selected coalition in terms of this randomization.

For a simple game, the formula for Shapley value seems to be particularly descriptive. Indeed, $v(T) - v(T \setminus \{i\})$ is always either 0 or 1, and this expression equals 1 if the coalition T wins, while 0 if the coalition $T \setminus \{i\}$ fails to win. Hence we have

$$\varphi_i[v] = \sum_T (t-1)!(n-t)!/n!,$$

where summation is extended over all such winning coalitions $T \supset i$ for which the coalition $T \setminus \{i\}$ is not the winning one.

Example 13. Game with major player [8]. The game is played by n players. One of the players is called "major". Coalition S wins 1 if it contains either a major player and at least one more player or all its $n-1$ "ordinary" players. If n is the major player, then the characteristic function of this game can be written as

$$v(S) = \begin{cases} 1, & S \supset \{i, n\},\ i \neq n, \\ 1, & S \supset \{1, \ldots, n-1\}, \\ 0, & \textit{otherwise}. \end{cases}$$

It is obvious that the conditions $v(T) = 1$ and $v(T \setminus \{n\}) = 0$ hold for any coalition $T \supset \{n\}$ if and only if $2 \leq |T| \leq n-1$. Hence

$$\varphi_n[v] = \sum_{t=2}^{n-1} C_{n-1}^{t-1} \frac{(t-1)!(n-t)!}{n!} = \frac{n-2}{n}.$$

Since the game is in (0-1) - reduced form,

$$\sum_{i=1}^{n-1} \varphi_i[v] = 1 - \varphi_n[v] = 2/n.$$

All ordinary players possess equal rights; hence, by symmetry,

$$\varphi_i[v] = \frac{2}{n(n-1)},\ i = 1, \dots, n-1.$$

Now, the "monopolistic" position of major player ensures him the payoff $(n-1)(n-2)/2$ times that of ordinary players.

Example 14. Reconsider the weighted majority voting game (example 4) for the case of a company with five shareholders $N = \{1, 2, 3, 4, 5\}$. Shareholder 1 owns 80 shares and the others hold 30 each. Let R, the return on the investment project under consideration, be 100. A decision is reached with 50% of the voting shares ($q = 0,5$). This yields to the following characteristic function:

$$v(S) = \begin{cases} 0,5\left(\sum\limits_{i\in S} w_i\right), & \text{for } \sum\limits_{i\in S} w_i \geq 0,5, \\ 0, & \text{for } \sum\limits_{i\in S} w_i < 0,5. \end{cases}$$

The following considerations will simplify computations. Firstly, note that all players but player 1 contribute in the same way to each coalition that does not contain them. Hence, they are symmetric and must obtain equal payoff. Denote this payoff by α, then $\phi_2 = \phi_3 = \phi_4 = \phi_5 = \alpha$ must hold and $\alpha = [v(N) - \phi_1]/4$. Secondly, to compute $\phi_1(v)$ note that for $\{S|\{1\} \in S\}$

$$[v(S) - v(S \setminus \{1\})] = \begin{cases} 0 & S = 1, \\ 40 + 15(S-1) & 2 \leq S \leq 5, \\ 40 & S = 5. \end{cases}$$

Using the value formula, one computes easily$\phi(v) = (\phi_1, \phi_2, \phi_3, \phi_4, \phi_5)$, where $\phi_1 = 50$; $\phi_i = 12,5$, $i = 2, 3, 4, 5$.

Example 15. Land-lord and farm labourers [8]. Suppose there are $n-1$ farm labourers (players $i = 1, \dots, n-1$) and a land-lord (player n). The land-lord engages k labourers and derives from the harvest a profit $f(k)$ ($f(k)$ increases monotonically). The farm labourers cannot derive a profit for themselves. This is described by the characteristic function

$$v(S) = \begin{cases} f(|S|-1), & \{n\} \in S, \\ 0, & otherwise. \end{cases}$$

Here, for all $T \supset \{n\}$, $|T| > 1$, $v(T) - v(T \setminus \{n\}) = f(t-1)$, where $t = |T|$ and from (4.7) follows

$$\varphi_n[v] = \sum_{t=2}^{n} C_{n-1}^{t-1} \frac{(t-1)!(n-t)!}{n!} f(t-1) = \frac{1}{n} \sum_{t=1}^{n-1} f(t).$$

By the efficiency and symmetry of all labourers, we have

$$\varphi_i[v] = \frac{1}{n-1}(f(n-1) - \frac{1}{n}\sum_{t=1}^{n-1} f(t)),\ i = 1, \dots, n-1.$$

In what follows we shall denote $\varphi_i[v]$ by Sh^i.

5 The potential of the Shapley value

Consider as before the n-person games in characteristic function form with transferable payoffs. We studied different solution concepts or different optimality principles. Some of them constitute subsets of payoff vectors or imputations (such as core and NM-solution). Finally, the Shapley value represents an optimality principle consisting from the unique payoff vector (unique imputation). In this section we follow Hart and Mas-Colell (1988) and introduce here one number which specifies the cooperative game. By using the "marginal contribution" principle we assign to each player his marginal contribution according to the numbers defined for the game. It happens that the only requirement, that the resulting payoff vector be "efficient" (i.e. that the payoffs add up to the worth of the grand coalition), determines this process uniquely.

A cooperative game with transferable payoffs is a pair (N, v), where N is a finite set of players and $v : 2^N \to R$ is the characteristic function, satisfying $v(\oslash) = 0$. A subset $S \subset N$ is called a coalition, and $v(S)$ is the worth of the coalition S. Given a game (N, v) and a coalition $S \subset N$, we write (S, v) for the subgame obtained by restricting v to (the subsets of) S; that is, the domain of the function v is restricted to 2^S.

Let $\hat{\Gamma}$ denote the set of all games. Given a function $P : \hat{\Gamma} \to R$ that associates a real number $P(N, v)$ to every game (N, v), the marginal contribution of player i in game (N, v) is defined as

$$D^i P(N, v) = P(N, v) - P(N \setminus \{i\}, v),$$

where $i \in N$. (The game $(N \setminus \{i\}, v)$ is restriction of (N, v) to $(N \setminus \{i\}, v)$.)

A function $P : \hat{\Gamma} \to R$ with $P(\oslash, v) = 0$ is called a *potential function* if it satisfies the following condition:

$$\sum_{i \in N} D^i P(N, v) = v(N) \tag{5.1}$$

for all games (N, v). Thus, a potential function is such that its marginals are always efficient; that is, they add up to the worth of the grand coalition.

Theorem 12. *There exists a unique potential function P. For every game (N, v) the resulting payoff vector $(D^i P(N, v))_{i \in N}$ of marginal contributions coincides with the Shapley value of the game. Moreover, the potential of a game (N, v) is uniquely determined by (5.1) applied only to the game and its subgames (i.e., to (S, v) for all $S \subset N$).*

Proof. Rewrite (5.1) as

$$P(N, v) = \frac{1}{|N|}\left(v(N) + \sum_{i \in N} P(N \setminus \{i\}, v)\right). \tag{5.2}$$

Starting with $P(\oslash, v) = 0$, (5.2) determines $P(N, v)$ recursively. This proves the existence and uniqueness of the potential function P and that $P(N, v)$ is uniquely determined by (5.1) (or (5.2)) applied just to (S, v) for all $S \subset N$.

It remains to show that $D^iP(N,v) = Sh^i(N,v)$ for all games (N,v) and all players $j \in N$, where P is the (unique) potential function and $Sh^i(N,v)$ denotes the Shapley value of player i in the game (N,v). We prove that all the axioms that uniquely determine the Shapley value are satisfied by D^iP. Efficiency is just (5.1); the other three axioms – dummy (null) player, symmetry, and additivity – are proved inductively using (5.2). Indeed, let i be a null player in the game (N,v) (i.e., $v(S) = v(S \setminus \{i\})$ for all S). We claim that this implies $P(N,v) = P(N \setminus \{i\}, v)$; hence $D^iP(N,v) = 0$. Assume the assertion holds for all games with less than $|N|$ players; in particular, $P(N\setminus\{j\}, v) = P(N\setminus\{j,i\}, v)$ for all $j \neq i$. Now subtract (5.2) for $N \setminus \{i\}$ from (5.2) for N, to obtain

$$|N|[P(N,v) - P(N \setminus \{i\}, v)] = [v(N) - v(N \setminus \{i\})]$$

$$+ \sum_{j \neq i} [P(N \setminus \{j\}, v) - P(N \setminus \{j,i\}, v)] = 0.$$

Next, assume players i and j are substitutes in the game (N,v). This implies that $P(N \setminus \{i\}, v) = P(N \setminus \{j\}, v)$ (use (5.2), noting that i and j are substitutes in $(N, \{k\}, v)$ for all $k \neq i, j$); thus $D^iP(N,v) = D^jP(N,v)$. Finally, another inductive argument on (5.2) shows that $P(N, v+w) = P(N,v) + P(N,w)$, implying additivity.

Present another way of viewing the potential. Given a game (N,v), the allocation of marginal contributions (i.e., $v(N) - v(N \setminus \{i\})$ to player i) is, in general, not efficient. One way to resolve this difficulty is to add a new player, say player 0, and extend the game to $N_0 = N \cup \{0\}$ in such a way that the allocation of marginal contributions in the extended game becomes efficient. Formally, let (N_0, v_0) be an extension of (N,v) (i.e., $v_0(S) = v(S)$ for all $S \subset N$). Then the requirement is

$$v_0(N_0) = \sum_{i \in N_0} [v_0(N_0) - v_0(N_0 \setminus \{i\})]$$

$$= [v_0(N_0) - v(N)] + \sum_{i \in N} [v_0(N_0) - v_0(N_0 \setminus \{i\})]. \tag{5.3}$$

This reduces to

$$v(N) = \sum_{i \in N} [v_0(N_0) - v_0(N_0 \setminus \{i\})], \tag{5.4}$$

which, yields the following restatement of the result of Theorem 12.

Corollary 2. *There exists a unique extension v_0 of v whose marginal contributions to the grand coalition are always efficient (more precisely, (5.3) is satisfied for the game and all its subgames); it is given by $v_0(S \cup \{0\}) = P(S,v)$ for all $S \subset N$, where P is the potential function.*

Note that the payoffs to the original players (in N) add up correctly to $v(N)$ (5.4); these are the Shapley values. Player 0, whose payoff is the residual $P(N,v) - v(N)$, may be regarded as a "hidden player", similarly to the "hidden factor" introduced by McKenzie in the study of production functions in order to explain the residual profit (or loss).

In (5.1) and (5.2) the potential is only given implicitly. We now present two explicit formulas. The T-*unanimity* game u_T (where T is a nonempty finite set) is defined by $u_T(S) = 1$ if $S \supset T$, and $u_T(S) = 0$ otherwise. It is well known that these games form a linear basis for Γ: Each game (N, v) has a unique representation (see Shapley[1953])

$$v = \sum_{T \subset N} \alpha_T u_T,$$

where, for all $T \subset N$,

$$\alpha_T \equiv \alpha_T(N, v) = \sum_{S \subset T} (-1)^{|T|-|S|} v(S). \tag{5.5}$$

Theorem 13. *The potential function P satisfies*

$$P(N, v) = \sum_{T \subset N} \frac{1}{|T|} \alpha_T$$

for all games (N, v), where α_T is given by (5.5).

Proof. Let $Q(N, v)$ denote the right-hand side in the preceding formula. Then $Q(\oslash, v) = 0$ and $Q(N, v) - Q(N \setminus \{i\}, v) = \sum_{i \in T} \alpha_T / |T|$, which when summed up over i shows that Q satisfies (5.1). Therefore, Q coincides with the unique potential function P.

The number $\delta_T = \alpha_T / |T|$ is called the *dividend* of each member of the coalition T and $Sh^i(N, v) = \sum_{i \in T} \delta_T$ [10].

Theorem 14. *The potential function P satisfies*

$$P(N, v) = \sum_{S \subset N} \frac{(s-1)!(n-s)!}{n!} v(S),$$

where $n = |N|$ and $s = |S|$.

Proof. The marginal contributions of the function on the right side are easily seen to yield the Shapley value.

To interpret this last formula, consider the following probabilistic model of choosing a random nonempty coalition $S \subset N$: First, choose a size $s = 1, 2, \ldots, n = |N|$ uniformly (i.e., with probability $1/n$ each). Second, choose a subset S of size s, again uniformly (i.e., each of the C_n^s subsets has the same probability). Equivalently, choose a random order of the n elements of N (with probability $1/n!$ each), choose a cutting point s ($1 \leq s \leq n$), and let S be the first s elements in that order. The probability of choosing of a set S with $|S| = s$ is

$$\pi_S = \frac{s!(n-s)!}{n \cdot n!} = \frac{s}{n} \frac{(s-1)!(n-s)!}{n!}.$$

Therefore the formula of Theorem 7, may be rewritten as

$$P(N, v) = \sum_{S \subset N} \pi_S \frac{n}{s} v(S) = E\left[\frac{|N|}{|S|} v(S)\right], \tag{5.6}$$

where E denotes expectation over S with respect to the foregoing probability model. The interpretation of (5.6) is that the potential is the *expected normalized worth*, or, equivalently, the *per capita potential* $P(N, v)/|N|$ equals the *average per capita worth* $v(S)/|S|$. This shows that the potential may be viewed as an appropriate "summary" of the characteristic function into one number (from which marginal contributions are then computed).

References

[1] Owen, G. *Game Theory*, p. 230. Saunders, Philadelphia, 1968.

[2] Roth, A. E. *Mathematics of operations research*, 2, 1977.

[3] Lemaire, J. *Cooperative game theory and its insurance applications*, p. 17–40. ASTIN Bulletin, 1991.

[4] Shubik, M. *Game Theory in the Social Sciences: Concepts and Solutions*, p. 514. Cambridge, Mass., MIT Press, 1982.

[5] Eichberger, J. *Game Theory for Economists*, p. 315. Academic Press, 1993.

[6] Moulin, H. *Game Theory for the Social Sciences*, p. 465. N.Y. Univ. Press, N.Y., 2nd edition, 1986.

[7] Moulin, H. *Theorie des jeux pour l'economie et la politique*, p. 200. Hermann, Paris, 1981.

[8] Vorobjev, N. N. *Game theory lectures for economists and system scientists*, p. 178. Springer, N.Y., 1977.

[9] Diubin, G. N. and Suzdal V. G. *Introduction to Applied Theory of Games*, p. 311. Nauka, Moscow, 1981.

[10] Harsanyi, J. C. *International Economic Review*, 4, 1963.

Positional games

November 28, 2018

1 Multistage games with perfect information

1.1. The preceding chapters dealt with games in normal form. A dynamic (i.e. continued during a period of time, not instantaneous) conflictly controlled process can be reduced to a normal form by formal introduction of the notion of a pure strategy. In the few cases when the power of a strategy space is not great and the possibility exists of numerical solutions such an approach seems to be allowable. However, in the majority of the problems connected with an optimal behavior of participants in the conflictly-controlled process the passage to normal form, i.e. the reduction of the problem to a single choice of pure strategies as elements of large dimension spaces or functional spaces, does not lead to effective ways of finding solutions, though permits illustration of one or another of the optimality principles. In a number of cases the general existence theorems for games in normal form does not allow finding or even specifying the optimal behavior in the games for which they constitute their normalizations. As is shown below, in "chess" there exists a solution in pure strategies. This result, however, cannot be obtained by direct investigation of matrix games. By investigation of differential games of pursuit and evasion it is possible in a number of cases to find explicit solutions. In such cases however, the normal form of differential game is so general that, for practical purposes, it is impossible to obtain specific results.

1.2. Mathematical dynamic models of conflict are investigated in the theory of *positional games.* The simplest class of positional games is the class of *finite stage games with perfect information.* To define a finite stage n-person game with perfect information we need a rudimentary knowledge of graph theory.

Let X be a finite set. The rule f setting up a correspondence between every element $x \in X$ and an element $f(x) \in X$ is called a single-valued map of X into X or the function defined on X and taking values in X. The set-valued map F of the set X into X is the rule which sets up a correspondence between every element $x \in X$ and a subset $F_x \subset X$ (here $F_x = \oslash$ is not ruled out). In what follows, for simplicity, the term "map" will be interpreted to mean a "set-valued map".

Let F be the map of X into X, while $A \subset X$. By the image set A will mean the set

$$FA \equiv \cup_{x \in A} F_x.$$

By definition, let $F(\oslash) = \oslash$. It can be seen that if $A_i \subset X$, $i = 1, \dots, n$, then

$$F(\cup_{i=1}^n A_i) = \cup_{i=1}^n F A_i, \quad F(\cap_{i=1}^n A_i) \subset \cap_{i=1}^n F A_i.$$

Define the maps $F^2, F^3, \dots, F^k, \dots$ as follows:

$$F_x^2 \equiv F(F_x),\ F_x^3 = F(F_x^2),\ \dots,\ F_x^k = F(F_x^{k-1}), \dots. \tag{1.1}$$

The map $\hat{F}$ of the set X into X is called a *transitive closure* of the map F, if

$$\hat{F}_x \equiv \{x\} \cup F_x \cup F_x^2 \cup \dots \cup F_x^k \cup \dots. \tag{1.2}$$

The map F^{-1} that is inverse to the map F is defined as

$$F_y^{-1} \equiv \{x | y \in F_x\},$$

i.e. this is the set of points x whose image contains the point y. The map $(F^{-1})_y^k$ is defined in much the same way as the map F_x^k, i.e.

$$(F^{-1})_y^2 = F^{-1}((F^{-1})_y), \tag{1.3}$$

$$(F^{-1})_y^3 = F^{-1}((F^{-1})_y^2), \dots, (F^{-1})_y^k = F^{-1}((F^{-1})_y^{k-1}).$$

If $B \subset X$, then let

$$F^{-1}(B) \equiv \{x | F_x \cap B \neq \oslash\}. \tag{1.4}$$

Example 1. (Chess.) Every position on a chess-board is defined by the number and composition of chess pieces for each player as well as by the arrangement of chess pieces at a given moment and the indication as to whose move it is. Suppose X is the set of positions, F_x, $x \in X$ is the set of those positions which can be realized immediately after the position x has been realized. If in the position x the number of black or white pieces is zero, then $F_x \equiv \oslash$. Now F_x^k defined by (1.1) is the set of positions which can be obtained from x in k moves, $\hat{F}_x$ is the set of all positions which can be obtained from x, $F^{-1}(A)$ $(A \subset X)$ is the set of all positions from which it is possible to make, in one move, the transition to positions of the set A (see (1.2) and (1.4)).

Depicting positions by dots and connecting by an arrow two positions x and y, $y \in F_x$, it is possible to construct the graph of a game emanating from the original position. However, because of a very large number of positions it is impossible to draw such a graph in reality.

The use of set-valued maps over finite sets makes it possible to represent the structure of many multistage games: chess, draughts, go, and other games.

Definition 1. *The pair (X, F) is called a graph if X is a finite set and F is a map of X into X.*

The graph (X, F) is denoted by G. In what follows, the elements of the set X are represented by points on a plane, and the pairs of points x and y, for which $y \in F_x$, are connected by the solid line with the arrow pointing from x to y. Then every element of the set X is called a vertex or a node of the graph, and the pair of elements (x, y), where $y \in F_x$ is called the *arc* of the graph. For the arc $p = (x, y)$ the nodes x and y are called the boundary nodes of the arc with x as the origin and y as the end point of the arc. Two arcs p and q are called contingent if they are distinct and have a boundary point in common.

The set of arcs in the graph is denoted by P. The set of arcs in the graph $G = (X, F)$ determines the map F, and vice versa, the map F determines the set P. Therefore, the graph G can be represented as $G = (X, F)$ or $G = (X, P)$.

The *path* in the graph $G = (X, F)$ is called a sequence of arcs, $p = (p_1, p_2, \ldots, p_k, \ldots)$, such that the end of the preceding arc coincides with the origin of the next one. The length of the path $p = (p_1, \ldots, p_k)$ is the number $l(p) = k$ of arcs in the sequence; in the case of an endless path p we set $l(p) \equiv \infty$.

The edge of the graph $G = (X, P)$ is called the set made up of two elements $x, y \in X$, for which either $(x, y) \in P$ or $(y, x) \in P$. The orientation is of no importance in the edge as opposed to the arc. The edges are denoted by p, q, and the set of edges by P. By the chain is meant a sequence of edges $(p_1, p_2, \ldots)$, where one of the boundary nodes of each edge p_k is also boundary for p_{k-1}, while the other is boundary for p_{k+1}.

The cycle is a finite chain starting in some node and terminating in the same node. The graph is called connected if its any two nodes can be connected by a chain.

By definition, the tree or the graph tree is a finite connected graph without cycles which has at least two nodes and has a unique node x_0 such that $\hat{F}_{x_0} = X$. The node x_0 is called the initial node of the graph G.

Example 2. Fig. 1 shows the graph or the *graph tree* with its origin at x_0. The nodes $x \in X$ or the vertices of the graph are marked by dots. The arcs are depicted as the arrowed segments emphasizing the origin and the end point of the arc.

Example 3. Generally speaking, draughts or chess cannot be represented by a graph tree if by the node of the graph is meant an arrangement of draughtsmen or chess pieces on the board at a given time and an indication of a move, since the same arrangement of pieces can be obtained in a variety of ways. However, if the node of the graph representing a structure of draughtsmen or chess pieces at a given time is taken to mean an arrangement of pieces on the board at a given time, an indication of a move and the past course of the game (all successive positions of pieces on the earlier moves), then each node is reached from the original one in a unique way (i.e. there exists the only chain passing from the original node to any given node); hence the corresponding graph of the game, contains no cycles and is the tree.

1.3. Let $z \in X$. The *subgraph* G_z of the tree graph $G = (X, F)$ is called a graph of the form (X_z, F_z), where $X_z = \hat{F}_z$ and $F_z x = F_x \cap X_z$. In Fig. 1 the dashed line encircles the subgraph starting in the node z. On the tree graph for all $x \in X_z$ the set F_x and the set $F_z x$ coincide, i.e. the map F_z is a restriction of the map F to the set X_z. Therefore, for the subgraphs of the tree graph we use the notation $G_z = (X_z, F)$.

1.4. We shall now define the *multistage game with perfect information on a finite tree graph.*

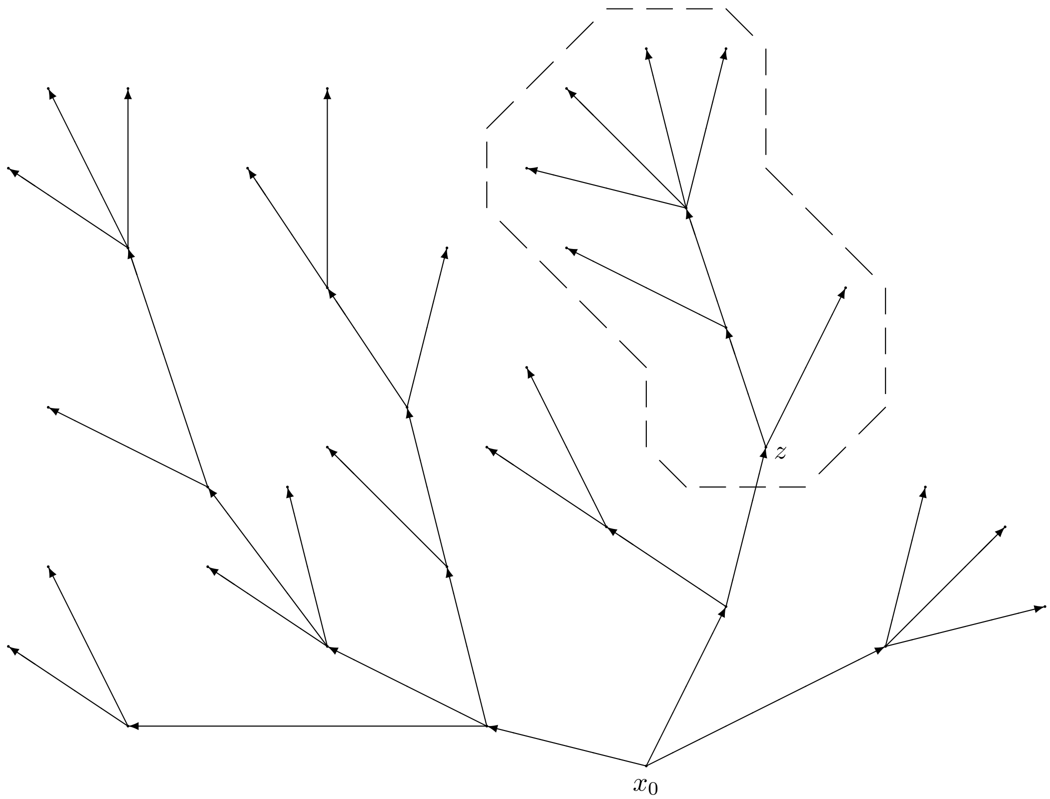


Figure 1:

Let $G = (X, F)$ be a tree graph. Consider the partition of the node set X into $n+1$ sets $X_1, \ldots, X_n, X_{n+1}$, $\cup_{i=1}^{n+1} X_i = X$, $X_k \cap X_l = \oslash$, $k \neq l$, where $F_x = \oslash$ for $x \in X_{n+1}$. The set X_i, $i = 1, \ldots, n$ is called the *set of personal positions* of the i-th player, while the set X_{n+1} is called the *set of final* or *terminal positions*. The real-valued functions $H_1(x), \ldots, H_n(x)$, $x \in X_{n+1}$ are defined on the set of final positions X_{n+1}. The function $H_i(x)$, $i = 1, \ldots, n$ is called a *payoff* to the i-th player.

The game proceeds as follows. Let there be given the set N of players designated by natural numbers $1, \ldots, i, \ldots, n$ (hereafter denoted as $N \equiv \{1, 2, \ldots, n\}$). Let $x_0 \in X_{i_1}$, then in the node (position) x_0 player i_1 "makes a move" and chooses the next node (position) $x_1 \in F_{x_0}$. If $x_1 \in X_{i_2}$, then in the node x_1 Player i_2 "makes a move" and chooses the next node (position) $x_2 \in F_{x_1}$ and so on. Thus, if the node (position) $x_{k-1} \in X_{i_k}$ is realized at the k-th step, then in this node Player i_k "makes a move" and selects the next node (position) from the set $F_{x_{k-1}}$. The game terminates as soon as the terminal node (position) $x_l \in X_{n+1}$, (i.e. the node for which $F_{x_l} = \oslash$) is reached.

Such a step-by-step selection implies a unique realization of some sequence

$x_0, \dots, x_k, \dots, x_l$ determining the path in the tree graph G which emanates from the initial position and reaches one of the final positions of the game. In what follows, such a path is called a *play* or *path of the game.* Because of the tree-like structure of the graph G, each play uniquely determines the final position x_l to be reached and, conversely, the final position x_l uniquely determines the play. In the position x_l each of the players i, $i = 1, \dots, n$, receives a payoff $H_i(x_l)$.

We assume that Player i making his choice in position $x \in X_i$ knows this position and hence, because of the tree-like structure of the graph G, can restore all previous positions. In this case, the players are said to have perfect information. Chess and draughts provide a good example of the game with perfect information, because players record their moves, and hence they know the past course of the game when making each move in turn.

Definition 2. *The single-valued map u_i which sets up a correspondence between each node (position) $x \in X_i$ and some unique node (position) $y \in F_x$ is called a strategy for player i.*

The set of all possible strategies for player i is denoted by U_i. Now the strategy of i-th player prescribes him, in any position x from his personal positions X_i, a unique choice of the next position.

The ordered set $u = (u_1, \dots, u_i, \dots, u_n)$, where $u_i \in U_i$, is called a *situation in the game*, while the Cartesian product $U = \prod_{i=1}^n U_i$ is called the *set of situations*. Each situation $u = (u_1, \dots, u_i, \dots, u_n)$ uniquely determines a play in the game, and hence payoffs of the players. Indeed, let $x_0 \in X_{i_1}$. In the situation $u = (u_1, \dots, u_i, \dots, u_n)$ the next position x_1 is then uniquely determined by the rule $u_{i_1}(x_0) = x_1$. Now let $x_1 \in X_{i_2}$. Then x_2 is uniquely determined by the rule $u_{i_2}(x_1) = x_2$. If the position $x_{k-1} \in X_{i_k}$ is realized at the k-th step, then x_k is uniquely determined by the rule $x_k = u_{i_k}(x_{k-1})$ and so on.

Suppose that the situation $u = (u_1, \dots, u_i, \dots, u_n)$ in the above sense determines a play $x_0, x_1, \dots, x_l$. Then we may introduce the notion of the *payoff function* K_i of player i by equating its value in each situation u to the value of the payoff H_i in the final position of the play $x_0, \dots, x_l$ corresponding to the situation $u = (u_1, \dots, u_n)$, that is

$$K_i(u) = K_i(u_1, \dots, u_i, \dots, u_n) \equiv H_i(x_l), \quad i = 1, \dots, n.$$

Functions K_i, $i = 1, \dots, n$, are defined on the set of situations $U = \prod_{i=1}^n U_i$. Thus, constructing the players' strategy sets U_i and defining the payoff functions K_i, $i = 1, \dots, n$, on the Cartesian product of strategy sets of players we obtain a game in normal form

$$\Gamma = (N, \{U_i\}_{i \in N}, \{K_i\}_{i \in N}),$$

where $N = \{1, \dots, i, \dots, n\}$ is the set of players, U_i is the strategy set for player i, and K_i is the payoff function for player i, $i = 1, \dots, n$.

1.5. For the purposes of further investigation of the game Γ we need to introduce the notion of a *subgame*, i.e. the game on a subgraph of the graph G in the main game.

Let $z \in X$. Consider a subgraph $G_z = (X_z, F)$ which is associated with the subgame Γ_z as follows. The players personal positions in the subgame Γ_z are determined by the rule $Y_i^z \equiv X_i \cap X_z$, $i = 1, \dots, n$, the set of final positions $Y_{n+1}^z = X_{n+1} \cap X_z$, player i's payoff $H_i^z(x)$ in the subgame is taken to be

$$H_i^z(x) \equiv H_i(x),\ x \in Y_{n+1}^z,\ i = 1, \dots, n.$$

Accordingly, player i's strategy u_i^z in the subgame Γ_z is defined to be the truncation of player i's strategy u_i in the game Γ to the set Y_i^z, i.e.

$$u_i^z(x) = u_i(x),\ x \in Y_i^z = X_i \cap X_z,\ i = 1, \dots, n.$$

The set of all strategies for player i in the subgame is denoted by U_i^z. Then each subgraph G_z is associated with the subgame in normal form

$$\Gamma_z = (N, \{U_i^z\}, \{K_i^z\}),$$

where the payoff function K_i^z, $i = 1, \dots, n$ are defined on the Cartesian product $U^z = \prod_{i=1}^n U_i^z$.

2 Absolute equilibrium (subgame-perfect)

2.1.

Definition 3. *The Nash equilibrium $u^* = (u_1^*, \dots, u_n^*)$ is called an absolute Nash equilibrium in the game Γ if for any $z \in X$ the situation $(u^*)^z = ((u_1^*)^z, \dots, (u_n^*)^z)$, where $(u_i^*)^z$ is the truncation of strategy u_i^* to the subgame Γ_z, is Nash equilibrium in this subgame.*

Then the following fundamental theorem is valid.

Theorem 1. *In any multistage game with perfect information on a finite tree graph there exists an absolute Nash equilibrium.*

Preparatory to proving this theorem we will first introduce the notion of a game length. By definition, the length of the game Γ means the length of the longest path in the corresponding graph $G = (X, F)$.

Proof will be carried out by induction along the length of the game. If the length of the game Γ is 1, then a move can be made by only one of the players who, by choosing the next node from the maximization condition of his payoff, will act in accordance with the strategy constituting an absolute Nash equilibrium.

Now, suppose the game Γ has the length k and $x_0 \in X_{i_1}$ (i.e. in the initial position x_0 player i_1 makes his move). Consider the family of subgames Γ_z, $z \in F_{x_0}$, where the length of each subgame does not exceed $k-1$. Suppose this theorem holds for all games whose length does not exceed $k-1$, and prove it for the game of length k. Since the subgame Γ_z, $z \in F_{x_0}$ has the length $k-1$ at most, under the assumption of induction the theorem holds for them and thus there exists an absolute Nash equilibrium situation. For each subgame Γ_z, $z \in F_{x_0}$ this situation will be denoted by

$$(u^*)^z = [(u_1^*)^z, \ldots, (u_n^*)^z]. \tag{2.1}$$

Using absolute equilibria in the subgames Γ_z we construct an absolute equilibrium in the game Γ. Let $u_i^*(x) = (u_i^*(x))^z$, for $x \in X_i \cap X_z$, $z \in F_{x_0}$, $i = 1, \ldots, n$, $u_{i_1}^*(x_0) = z^*$, where z^* is obtained from the condition

$$K_{i_1}^{z^*}[(u^*)^{z^*}] = \max_{z \in F_{x_0}} K_{i_1}^z[(u^*)^z]. \tag{2.2}$$

The function u_i^* is defined on player i's set of personal positions X_i, $i = 1, \ldots, n$, and for every fixed $x \in X_i$ the value $u_i^*(x) \in F_x$. Thus, u_i^*, $i = 1, \ldots, n$, is a strategy for player i in the game Γ, i.e. $u_i^* \in U_i$. By construction, the truncation $(u_i^*)^z$ of the strategy u_i^* to the set $X_i \cap X_z$ is the strategy appearing in the absolute Nash equilibrium of the game Γ_z, $z \in F_{x_0}$. Therefore, to complete the proof of the theorem, it suffices to show that the strategies u_i^*, $i = 1, \ldots, n$ constructed by formulas (2.2) constitute a Nash equilibrium in the game Γ. Let $i \neq i_1$. By the construction of the strategy $u_{i_1}^*$ after a position z^* has been chosen by player i_1 at the first step, the game Γ becomes the subgame Γ_{z^*}. Therefore, for $i \neq i_1$,

$$K_i(u^*) = K_i^{z^*}\{(u^*)^{z^*}\} \geq K_i^{z^*}\{(u^* \| u_i)^{z^*}\} = K_i(u^* \| u_i), \; u_i \in U_i, \; i = 1, \ldots, n, \tag{2.3}$$

since $(u^*)^{z^*}$ is an absolute equilibrium in the subgame Γ_{z^*}. Let $u_{i_1} \in U_{i_1}$ be an arbitrary strategy for player i_1 in the game Γ. Denote $z_0 = u_{i_1}(x_0)$. Then

$$K_{i_1}(u^*) = K_{i_1}^{z^*}\{(u^*)^{z^*}\} = \max_{z \in F_{x_0}} K_{i_1}^z\{(u^*)^z\} \geq K_{i_1}^{z_0}\{(u^*)^{z_0}\} \geq$$

$$K_{i_1}^{z_0}\{(u^* \| u_{i_1})^{z_0}\} = K_{i_1}(u^* \| u_{i_1}). \tag{2.4}$$

The assertion of this theorem now follows from (2.3), (2.4).

2.2. *Example 4.* Suppose the game Γ is played on the graph depicted in Fig. 2 and the set N is composed of two players: $N = \{1, 2\}$. Referring to Fig. 2 we determine sets of personal positions. The nodes of the set X_1 are represented by circles and those of the set X_2 by blocks. Players' payoffs are written in final positions. Designate by double indices the positions appearing in the sets X_1 and X_2, and by one index the arcs emanating from each node. The choice in the node x is equivalent to the choice of the next node $x' \in F_x$; therefore we assume that the strategies indicate in each node the index of the arc along which it is necessary to move further. For example, Player 1's strategy $u_1 = (2, 1, 2, 3, 1, 2, 1, 1)$ tells him to choose arc 2 in node 1, arc 1 in node 2, arc 2 in node 3, arc 3 in node 4, and so on. Since the set of personal positions of the first player is composed of eight nodes, the strategy for him is an eight-dimensional vector. Similarly, any strategy for Player 2 is a seven-dimensional vector. Altogether there are 864 strategies for Player 1 and 576 strategies for Player 2. Thus the corresponding normal form appears to

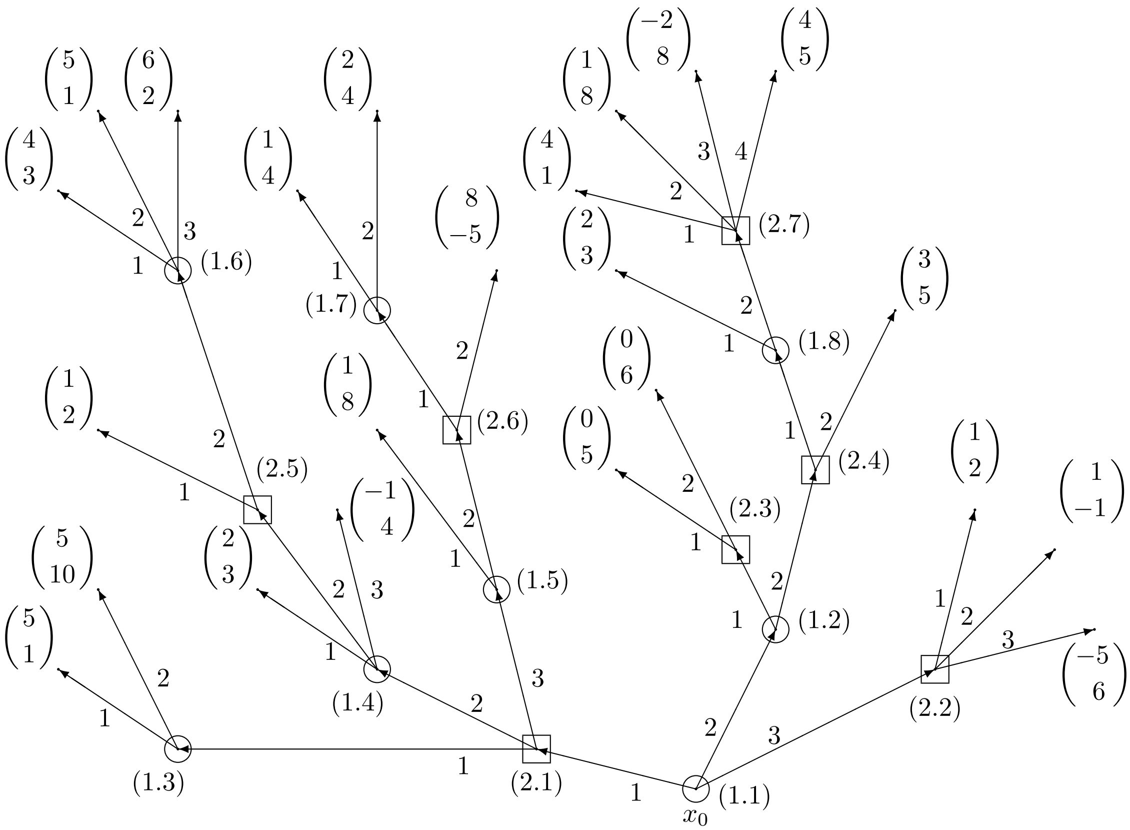


Figure 2:

be an 864×576 bimatrix game. It appears natural that the solution of such bimatrix games by the methods proposed in Chap. 3 is not only difficult, but also impossible. However, the game involved is simple enough to be solved by the backward construction of an absolute Nash equilibrium as proposed in the proof of Theorem 1 in 2.1.

Indeed, denote by $v_1(x), v_2(x)$ the payoffs in the subgame Γ_x in a fixed absolute equilibrium. First we solve subgames $\Gamma_{1.6}$, $\Gamma_{1.7}$, $\Gamma_{2.7}$. It can be easily seen that $v_1(1.6) = 6$, $v_2(1.6) = 2$, $v_1(1.7) = 2$, $v_2(1.7) = 4$, $v_1(2.7) = 1$, $v_2(2.7) = 8$. Further, solve subgames $\Gamma_{2.5}$, $\Gamma_{2.6}$, $\Gamma_{1.8}$. Subgame $\Gamma_{2.5}$ has two Nash equilibria, because Player 2 does not care which of the alternative to choose. At the same time, his choice appears to be essential for Player 1, because with Player 2's choice of left-hand arc Player 1 scores +1 or, with Player 2's choice of right-hand arc, Player 1 scores +6. We point out this feature and suppose Player 2 "favors" and chooses the right-hand arc in position (2.5). Then $v_1(2.5) = v_1(1.6) = 6$, $v_2(2.5) = v_2(1.6) = 2$, $v_1(2.6) = v_1(1.7) = 2$, $v_2(2.6) = v_2(1.7) = 4$, $v_1(1.8) = 2$, $v_2(1.8) = 3$. Further, solve games $\Gamma_{1.3}$, $\Gamma_{1.4}$, $\Gamma_{2.3}$, $\Gamma_{1.5}$, $\Gamma_{2.4}$. Subgame $\Gamma_{1.3}$ has two Nash equilibria, because Player 1 does not care which of the alternative to choose. At the same time, his choice, appears to be essential for Player 2, because with Player 1's choice of the left-hand alternative he scores +1, whereas with the choice of the right-hand alternative he scores +10. Suppose Player 1 "favors" and chooses in position (1.3) the right-hand alternative. Then $v_1(1.3) = 5$, $v_2(1.3) = 10$, $v_1(1.4) = v_1(2.5) = 6$, $v_2(1.4) = v_2(2.5) = 2$, $v_1(1.5) = v_1(2.6) = 2$, $v_2(1.5) = v_2(2.6) = 4$, $v_1(2.3) = 0$, $v_2(2.3) = 6$, $v_1(2.4) = 3$, $v_2(2.4) = 5$. Further, solve games $\Gamma_{2.1}$, $\Gamma_{1.2}$, $\Gamma_{2.2}$; $v_1(2.1) = v_1(1.3) = 5$, $v_2(2.1) = v_2(1.3) = 10$, $v_1(1.2) = v_1(2.4) = 3$, $v_2(1.2) = v_2(2.4) = 5$, $v_1(2.2) = -5$, $v_2(2.2) = 6$. Now solve the game $\Gamma = \Gamma_{1.1}$. Here $v_1(1.1) = v_1(2.1) = 5$, $v_2(1.1) = v_2(2.1) = 10$.

As a result we have an absolute Nash equilibrium (u_1^*, u_2^*), where

$$u_1^* = (1, 2, 2, 2, 2, 3, 2, 1),\ u_2^* = (1, 3, 2, 2, 2, 1, 2). \tag{2.5}$$

In the situation (u_1^*, u_2^*) the game follows the path (1.1), (2.1), (1.3). It is apparent from the construction that the strategies u_i^*, $i = 1, 2$, are "favorable" in that the player i making his move and being equally interested in the choice of the subsequent alternatives, chooses that alternative which is favorable for player $3 - i$.

The game Γ has absolute equilibria in which the payoffs to players are different. To construct such equilibria, it suffices to replace the players' "favorableness" condition by the inverse condition, i.e. the "unfavorableness" condition. Denote by $\overline{v}_1(x)$, $\overline{v}_2(x)$ the payoffs to players in subgame Γ_x when players use an "unfavorable" equilibrium. Then we have: $v_1(1.6) = \overline{v}_1(1.6) = 6$, $v_2(1.6) = \overline{v}_2(1.6) = 2$, $v_1(1.7) = \overline{v}_1(1.7) = 2$, $v_2(1.7) = \overline{v}_2(1.7) = 4$, $v_1(2.7) = \overline{v}_1(2.7) = -2$, $v_2(2.7) = \overline{v}_2(2.7) = 8$. As noted before, subgame $\Gamma_{2.5}$ has two Nash equilibria. Contrary to the preceding case, we assume that Player 2 "does not favor" and chooses the node which ensures a maximum payoff to him and a minimum payoff to Player 1. Then $\overline{v}_1(2.5) = 1$, $\overline{v}_2(2.5) = 2$, $\overline{v}_1(2.6) = v_1(1.7) = 2$, $\overline{v}_2(2.6) = v_2(1.7) = 4$, $\overline{v}_1(1.8) = v_1(1.8) = 2$, $\overline{v}_2(1.8) = v_2(1.8) = 3$. Further, we seek a solution to the games $\Gamma_{1.3}$, $\Gamma_{1.4}$, $\Gamma_{1.5}$, $\Gamma_{2.3}$, $\Gamma_{2.4}$. Subgame $\Gamma_{1.3}$ has two Nash equilibria. As in the preceding case, we choose "unfavorable" actions for Player 1. Then we have $\overline{v}_1(1.3) = v_1(1.3) = 5$, $\overline{v}_2(1.3) = 1$, $\overline{v}_1(1.4) = 2$, $\overline{v}_2(1.4) = 3$, $\overline{v}_1(1.5) = \overline{v}_1(2.6) = v_1(1.5) = 2$, $\overline{v}_2(1.5) = \overline{v}_2(2.6) = v_2(2.6) = 4$, $\overline{v}_1(2.3) = v_1(2.3) = 0$, $\overline{v}_2(2.3) = v_2(2.3) = 6$, $\overline{v}_1(2.4) = v_1(2.4) = 3$, $\overline{v}_2(2.4) = v_2(2.4) = 5$. Further, we solve games $\Gamma_{2.1}$, $\Gamma_{1.2}$, $\Gamma_{2.2}$. We have: $\overline{v}_1(2.1) = \overline{v}_1(1.5) = 2$, $\overline{v}_2(2.1) = \overline{v}_2(1.5) = 4$, $\overline{v}_1(1.2) = \overline{v}_1(2.4) = 3$, $\overline{v}_2(1.2) = \overline{v}_2(2.4) = 5$, $\overline{v}_2(2.2) = v_2(2.2) = 6$, $\overline{v}_1(2.2) = v_1(2.2) = -5$. Now solve the game $\Gamma = \Gamma_{1.1}$. Here $\overline{v}_1(1.1) = \overline{v}_1(1.2) = 3$, $\overline{v}_2(1.1) = \overline{v}_2(1.2) = 5$.

We have thus obtained a new Nash equilibrium

$$\overline{u}_1^*(\cdot) = (2, 2, 1, 1, 2, 3, 2, 1),\ \overline{u}_2^*(\cdot) = (3, 3, 2, 2, 1, 1, 3). \tag{2.6}$$

Payoff to both players in situation (2.6) are less than those in situation (2.5). Just as situation (2.5), situation (2.6) is an absolute equilibrium.

2.3. It is apparent that in parallel with "favorable" and "unfavorable" absolute Nash equilibria there exists the whole family of intermediate absolute equilibria. Of interest is the question concerning the absence of two distinct absolute equilibria differing by payoffs to players.

Theorem 2. *[1]. Let the players' payoffs $H_i(x)$, $i = 1, \dots, n$, in the game Γ be such that if there exists an i_0 and there are x, y such that $H_{i_0}(x) = H_{i_0}(y)$, then $H_i(x) = H_i(y)$ for all $i \in N$. Then in the game Γ, the players' payoffs coincide in all absolute equilibria.*

Proof. Consider the family of subgames Γ_x of the game Γ and prove the theorem by induction over their length $l(x)$. Let $l(x) = 1$ and suppose player i_1 makes a move in a unique nonterminal position x. Then in the equilibrium he makes his choice from the condition

$$H_{i_1}(\overline{x}) = \max_{x' \in F_x} H_{i_1}(x').$$

If the point $\overline{x}$ is unique, then so is the payoff vector in the equilibrium which is here equal to $H(\overline{x}) = \{H_1(\overline{x}), \dots, H_n(\overline{x})\}$. If there exists a point $\overline{\overline{x}} \neq \overline{x}$ such that $H_{i_1}(\overline{\overline{x}}) = H_{i_1}(\overline{x})$, then there is one more equilibrium with payoffs $H(\overline{\overline{x}}) = \{H_1(\overline{\overline{x}}), \dots, H_{i_1}(\overline{\overline{x}}), \dots, H_n(\overline{\overline{x}})\}$. From the condition of the theorem, however, it follows that if $H_{i_1}(\overline{\overline{x}}) = H_{i_1}(\overline{x})$, then $H_i(\overline{\overline{x}}) = H_i(\overline{x})$ for all $i \in N$.

Let $v(x) = \{v_i(x)\}$ be the payoff vector in the equilibrium in a single-stage subgame Γ_x which, as is shown above, is determined in a unique way. Show that if the equality $v_{i_0}(x') = v_{i_0}(x'')$ holds for some i_0 (x', x'' are such that the lengths of the subgames $\Gamma_{x'}$, $\Gamma_{x''}$ are 1), then $v_i(x') = v_i(x'')$ for all $i \in N$. Indeed, let $x' \in X_{i_1}$, $x'' \in X_{i_2}$, then

$$v_{i_1}(x') = H_{i_1}(\overline{x}') = \max_{y \in F_{x'}} H_{i_1}(y),$$

$$v_{i_2}(x'') = H_{i_2}(\overline{x}'') = \max_{y \in F_{x''}} H_{i_2}(y)$$

and $v_i(x') = H_i(\overline{x}')$, $v_i(x'') = H_i(\overline{x}'')$ for all $i \in N$. From the equality $v_{i_0}(x') = v_{i_0}(x'')$ it follows that $H_{i_0}(\overline{x}') = H_{i_0}(\overline{x}'')$. But, under the condition of the theorem, $H_i(\overline{x}') = H_i(\overline{x}'')$ for all $i \in N$. Hence $v_i(x') = v_i(x'')$ for all $i \in N$.

We now assume that in all subgames Γ_x of length $l(x) \leq k - 1$ the payoff vector in equilibria is determined uniquely and if for some two subgames $\Gamma_{x'}, \Gamma_{x''}$ whose length does not exceed $k - 1$, $v_{i_0}(x') = v_{i_0}(x'')$ for some i_0, then $v_i(x') = v_i(x'')$ for all $i \in N$.

Suppose the game Γ_{x_0} is of length k and player i_1 makes his move in the initial position x_0. By the induction hypothesis, for all $z \in F_{x_0}$ in the game Γ_z the payoffs in Nash equilibria are determined uniquely. Let the payoff vector in Nash equilibria in the game Γ_z be $\{v_i(z)\}$. Then as follows from (2.2), in the node x_0 player i chooses the next node $\overline{z} \in F_{x_0}$ from the condition

$$v_{i_1}(\overline{z}) = \max_{z \in F_{x_0}} v_{i_1}(z). \tag{2.7}$$

If the point $\overline{z}$ determined by (2.7) is unique, then the vector with components $v_i(x_0) = v_i(\overline{z})$, $i = 1, \ldots, n$, is a unique payoff vector in Nash equilibria in the game Γ_{x_0}. If, however, there exist two nodes $\overline{z}$, $\overline{\overline{z}}$ for which $v_{i_1}(\overline{z}) = v_{i_1}(\overline{\overline{z}})$, then, by the induction hypothesis, since the lengths of subgames $\Gamma_{\overline{z}}$ and Γ_z do not exceed $k-1$, the equality $v_{i_1}(\overline{z}) = v_{i_1}(z)$ implies the equality $v_i(\overline{z}) = v_i(z)$ for all $i \in N$. Thus, in this case the payoffs in equilibria $v_i(x_0)$, $i \in N$ are also determined uniquely.

2.4. *Example 5.* We have seen in the previous example that "favorableness" of the players give them higher payoffs in the corresponding Nash equilibria, than the "unfavorable" behavior. But it is not always the case. Sometimes the "unfavorable" Nash equilibrium gives higher payoffs to all the players than "favorable" one. We shall illustrate this rather nontrivial fact on example. Consider the two-person game on the Fig. 3. The nodes from the personal

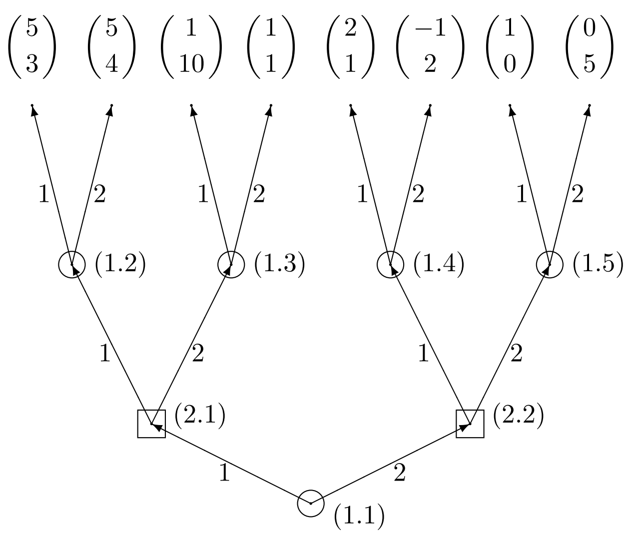


Figure 3:

positions X_1 are represented by circles and those from X_2 by blocks, with players payoffs written in final position. On the figure positions from the sets X_i ($i = 1, 2$) are numbered by double indexes (i, j) where i is the index of the player and j the index of the node x in the set X_i. One can easily see that the "favorable" equilibrium has the form $((2, 2, 1, 1, 1), (2, 1))$ with payoffs $(2, 1)$. The "unfavorable" equilibrium has the form $((1, 1, 2, 1, 1), (1, 1))$ with payoffs $(5, 3)$.

2.5. [2]. Consider the n-person game with perfect information, where each player $i \leq n$ can either end the game by playing D or play A and give the move to player $i + 1$ (see Fig. 4).

If player i selects D, each player gets $1/i$, if all players select A each gets 2. The backward induction algorithm for computing the subgame perfect (ab-

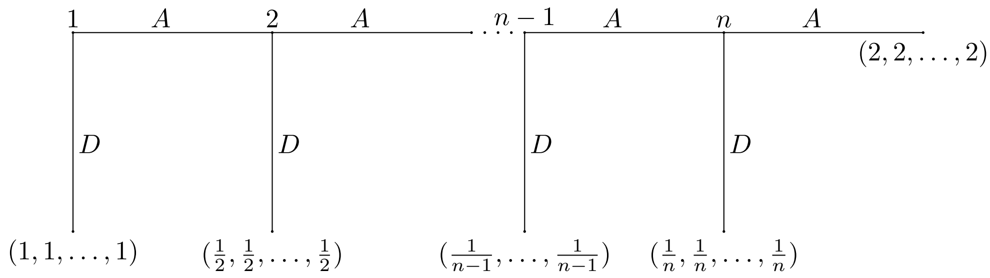


Figure 4:

solute) equilibria predicts that all players should play A. Thus the situation $(A, A, \ldots, A)$ is a subgame perfect Nash equilibrium. (Note that in the game under consideration each player is moving only once and has two alternatives, wich are also his strategies.) But there are also other equilibria. One class of Nash equilibria has the form $(D, A, A, D, \ldots)$, where the first player selects D and at least one of the others selects D. The payoffs in the first case are $(2, 2, \ldots, 2)$ and in the second $(1, 1, \ldots, 1)$. On the basis of robustness argument it seems that the equilibrium $(A, A, \ldots, A)$ is inefficient if n is very large. The equilibrium $(D, A, A, D, \ldots)$ is such because the player 4 uses the punishment strategy to enforce player 1 to play D. This equilibrium is not subgame perfect, because it is not an equilibrium in any subgame starting from the positions $2, 3$.

2.6. ***Cuban missile crises.*** Consider now the example which can in a very simplified version be modelled by a game on the tree with perfect information. Namely the Cuban missile crises between the United States under John Kennedy and Soviet Union under Nikita Khrushchev in 1963. (see [3]).

Khrushchev got information from his secret service that USA is planning a nuclear air over USSR planning to destroy 60 main USSR cities. To prevent this attack he starts the game by deciding whether or not to place intermediate range ballistic missiles in Cuba. If he places the missiles, his opponent player, Kennedy, will have three options: not react, blockade Cuba or eliminate the missiles by special airstrike. If Kennedy chooses the aggressive action of a blockade or an airstrike, Khrushchev may acquiesce or he may go by way of escalation with possible nuclear war at the end.

Consider the game tree on Figure 5.

In this game Khrushchev – the first player – moves in circled vertexes and Kennedy – the second player – in squared vertexes. Payoffs are written in the final vertexes and on the first place is the payoff of Khrushchev.

Interpret the payoffs. If Player 1 (Khrushchev) decides not to place missiles his payoff is (-5), since there will be the probability of nuclear strike against USSR, in this case Player 2 (Kennedy) will get (5) since at that time USSR did not have the opportunity to strike back with nuclear bombs on long distances. If Player 1 chooses on first stage place missiles, then the second player after some time will be informed and will have the possibility of taking one of following

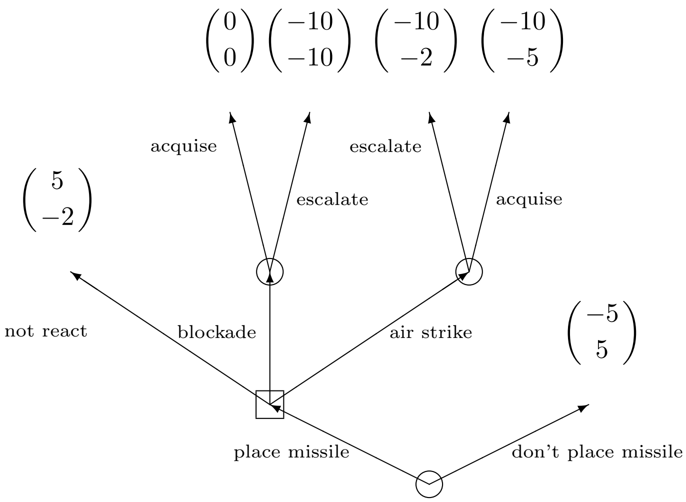


Figure 5:

alternatives: not react (n), blockade Cuba (b), air strike to eliminate the missiles (a). If Player 2 chooses (n) then the game is over and the payoff of the Player 1 will be (5) and payoff of the Player 2 will be (-2). Since in this case the nuclear war will have a very small probability but strategic position of USSR would be much better than of USA. If Player 2 decides to blockade, then Player 1 has two alternatives: to acquire (a) or to escalate (e). In the case Player 1 chooses (a) the game is draw with payoffs (0, 0), which means that USSR will not keep missiles in Cuba and USA will forget the idea of nuclear air strike against USSR after Cuba lesson. If Player 1 decides to escalate the nuclear war is possible and the payoffs will be (-10, -10) the same for both, since compared with the case after 1 player alternative do not place missiles the nuclear war will make symmetric damage on both countries. Suppose now that Player 2 chooses air strike (alternative a), then if Player 1 chooses (a) the payoffs will be (-10, -2) since in this case Player 2 will have incentive to start nuclear war (as he wanted before the crises) but maybe some missiles in Cuba will remain untouched by air strike and few USA cities may be destroyed by nuclear attack (-2). In the case if Player 1 choused (e), the payoff will be (-10, -5) with less damage for USA since air strike can decrease the power of USSR missiles in Cuba (but not too much). Nash equilibrium can be found by backward induction and has payoff (0, 0), which really happens.

2.7. Indifferent equilibrium. As we have seen from 2.4 Example 5 a subgame perfect equilibrium may appear non-unique in an extensive form game. This happens when the payoffs of some players coincides in terminal positions. Then the behavior of the player depends on his attitude to his opponents and the behavior of the player type naturally arises. In two person cases one way destinguish with between two types of players "favorable" and "unfavorable". The resulting two different subgame perfect Nash equilibrium where demonstrated in Example 5.

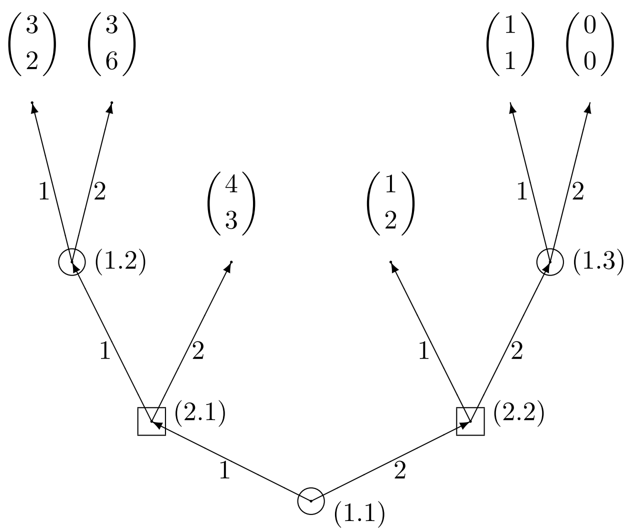


Figure 6:

There is another approach to avoid ambiguity in the behavior of players when any continuation of the game gives the same payoffs. Such an approach was proposed in [4]. It realizes the following idea: in a given personal position the player moving in this position randomizes the alternatives yielding the same payoffs with equal probabilities. It can be easily proved that such behavior will also form a subgame perfect Nash equilibrium (not only in the case the randomization is made with equal probabilities, but also if it is made with arbitrary probability distribution over the alternatives yielding the same payoff).

For instance let us evaluate an indifferent equilibrium in the game from Example 5. In this example the alternatives 1 and 2 in position (1.2) will be chosen with probabilities $(\frac{1}{2},\ \frac{1}{2})$, and in the same manner the alternatives 1 and 2 in position (1.3). In this case the Nash equilibrium (indifferent) will give the same payoffs (2,1) as favorable equilibrium.

Consider now another example.

Here Player I moves in positions $\{(1.1),\ (1.2),\ (1.3)\}$, and Player II in positions $\{(2.1), (2.2)\}$. In this game in "favorable" equilibrium Player I chooses alternative 2 in position (1.2), and Nash equilibrium will give the payoffs (3, 6). In unfavorable equilibrium Player I chooses in position (1.2) alternative 1, and Nash equilibrium will give the payoffs (4, 3). In indifferent equilibrium Player I will choose alternatives 1, 2 with probabilities $(\frac{1}{2},\ \frac{1}{2})$ in position (1.2). The payoffs in indifferent equilibrium will be (3, 4). In this example as well as in Example 5 there are infinite many subgame perfect Nash equilibrium, since

Player I in position (1.2) can choose the alternatives 1 and 2 with any probability $p = (p_1, p_2)$, $p_1 \geq 0$, $p_2 \geq 0$, $p_1 + p_2 = 1$, and all this mixed behavior in position (1.2) will be part of some subgame perfect Nash equilibrium.

3 Penalty strategies

3.1. In 2.1 we proved the existence of absolute (Nash) equilibria in multistage games with perfect information on a finite graph tree. However, the investigation of particular games of this class may reveal the whole family of equilibria whose truncations are not necessarily equilibria in all subgames of the original game. Among such equilibria are equilibria in penalty strategies. We shall demonstrate this with the examples below.

Example 6. Suppose the game Γ proceeds on the graph depicted in Fig. 7. The set $N = \{1, 2\}$ is made up of two players. In Fig. 7, as an Example 5, the circles represent the nodes making up the set X_1 and the blocks represent the set X_2. The nodes of the graph are designated by double indices and the arcs by single indices.

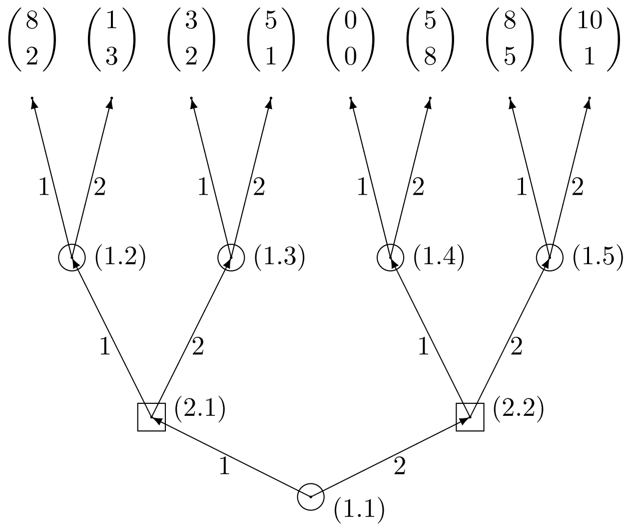


Figure 7:

It can be easily seen that the situation $u_1^* = (1, 1, 2, 2, 2)$, $u_2^* = (1, 1)$ is an absolute equilibrium in the game Γ. In this case, the payoffs to players are 8 and 2 units, respectively. Now consider the situation $\overline{u}_1 = (2, 1, 2, 1, 2)$, $\overline{u}_2 = (2, 2)$. In this situation the payoffs to players respectively are 10 and 1, and thus Player 1 receives a greater amount than in the situation (u_1^*, u_2^*). The situation $(\overline{u}_1, \overline{u}_2)$ is equilibrium in the game Γ but not absolute equilibrium. In fact, in the subgame $\Gamma_{1.4}$ the truncation of the strategy $\overline{u}_1$ tells Player 1 to choose the left-hand arc, which is not optimal for him in position 1.4. Such an action taken by Player 1 in position 1.4 can be interpreted as a "penalty" threat to Player 2 if he avoids Player 1's desirable choice of arc 2 in position 2.2, thereby depriving Player 1 of the maximum payoff of 10 units. But this "penalty" threat is unlikely to be treated as valid, because the penalizer (Player 1) may lose in this case 5 units (acting nonoptimally).

3.2. We shall now give a strict definition of penalty strategies. For simplicity, we shall restrict ourselves to the case of a nonzero-sum two-person game. Let there be a multistage nonzero sum two-person game

$$\Gamma = \langle U_1, U_2, K_1, K_2 \rangle.$$

The game Γ will be associated with two zero-sum games Γ_1 and Γ_2 as follows. The game Γ_1 is a zero-sum game constructed in terms of the game Γ, where Player 2 plays against Player 1, i.e. $K_2 = -K_1$. The game Γ_2 is a zero-sum game constructed in terms of the game Γ, where Player 1 plays against Player 2, i.e. $K_1 = -K_2$. The graphs of the games $\Gamma_1, \Gamma_2, \Gamma$ and the sets therein coincide. Denote by (u_{11}^*, u_{21}^*) and (u_{12}^*, u_{22}^*) absolute equilibria in the games Γ_1, Γ_2 respectively. Let Γ_{1x}, Γ_{2x} be subgames of the games Γ_1, Γ_2; $v_1(x), v_2(x)$ are the values of these subgames. Then the situations $\{(u_{11}^*)^x, (u_{21}^*)^x\}$ and $\{(u_{12}^*)^x, (u_{22}^*)^x\}$ are equilibria in the games Γ_{1x}, Γ_{2x}, respectively, and $v_1(x) = K_1^x((u_{11}^*)^x, (u_{21}^*)^x)$, $v_2(x) = K_2^x((u_{12}^*)^x, (u_{22}^*)^x)$.

Consider an arbitrary pair (u_1, u_2) of strategies in the game Γ. Of course, this pair is the same in the games Γ_1, Γ_2. Let $Z = (x_0 = z_0, z_1, \ldots, z_l)$ be the path to be realized in the situation (u_1, u_2).

Definition 4. *The strategy $\tilde{u}_1(\cdot)$ is called a penalty strategy of Player 1 if*

$$\tilde{u}_1(z_k) = z_{k+1} \text{ for } z_k \in Z \cap X_1, \tag{3.1}$$

$$\tilde{u}_1(y) = u_{12}^*(y) \text{ for } y \in X_1,\ y \notin Z.$$

The strategy $\tilde{u}_2(\cdot)$ is called a penalty strategy for Player 2 if

$$\tilde{u}_2(z_k) = z_{k+1} \text{ for } z_k \in Z \cap X_2, \tag{3.2}$$

$$\tilde{u}_2(y) = u_{21}^*(y) \text{ for } y \in X_2,\ y \notin Z.$$

3.3. From the definition of penalty strategies we immediately obtain the following properties:

1^0. $K_1(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot)) = H_1(z_l)$, $K_2(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot)) = H_2(z_l)$.

2^0. Suppose one of the players, say, Player 1, uses strategy $u_1(\cdot)$ for which the position $z_k \in Z \cap X_1$ is the first in the path Z, where $u_1(\cdot)$ dictates the choice of the next position z'_{k+1} that is different from the choice dictated by the strategy $\tilde{u}_1(\cdot)$, i.e. $z'_{k+1} \neq z_{k+1}$. Hence from the definition of the penalty strategy $\tilde{u}_2(\cdot)$ it follows that

$$K_1(u_1(\cdot), \tilde{u}_2(\cdot)) \leq v_1(z_k). \tag{3.3}$$

Similarly, if Player 2 uses strategy $u_2(\cdot)$ for which the position $z_k \in Z \cap X_2$ is the first in the path Z, where $u_2(\cdot)$ dictates the choice of the next position z'_{k+1}

that is different from the choice dictated by $\tilde{u}_2(\cdot)$ i.e. $z'_{k+1} \neq z_{k+1}$, then from the definition of the penalty strategy $\tilde{u}_1(\cdot)$ it follows that

$$K_2(\tilde{u}_1(\cdot), u_2(\cdot)) \leq v_2(z_k). \tag{3.4}$$

Hence, in particular, we obtain the following theorem.

Theorem 3. *Let $(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot))$ be a situation in penalty strategies. For the situation $(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot))$ to be equilibrium, it is sufficient that for all $k = 0, 1, \dots, l-1$ there be the inequalities*

$$K_1(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot)) \geq v_1(z_k), \tag{3.5}$$

$$K_2(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot)) \geq v_2(z_k),$$

where $z_0, z_1, \dots, z_l$ is the path realized in the situation $(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot))$.

3.4. Suppose that $u^*_{11}(\cdot)$ and $u^*_{22}(\cdot)$ are optimal strategies for Players 1 and 2, respectively, in the auxiliary zero-sum games Γ_1 and Γ_2 and $\overline{Z} = \{\overline{z}_0, \overline{z}_1, \dots, \overline{z}_l\}$ is the path corresponding to the situation $(u^*_{11}(\cdot), u^*_{22}(\cdot))$. Also, suppose the penalty strategies $\tilde{\tilde{u}}_1(\cdot)$ and $\tilde{\tilde{u}}_2(\cdot)$ are such that $\tilde{\tilde{u}}_1(\overline{z}_k) = u^*_{11}(z_k)$ for $\overline{z}_k \in \overline{Z} \cap X_1$ and $\tilde{\tilde{u}}_2(\overline{z}_k) = u^*_{21}(\overline{z}_k)$ for $\overline{z}_k \in \overline{Z} \cap X_2$. Then the situation $(\tilde{\tilde{u}}_1(\cdot), \tilde{\tilde{u}}_2(\cdot))$ forms a Nash equilibrium in penalty strategies. To prove this assertion it suffices to show that

$$K_1(u^*_{11}(\cdot), u^*_{22}(\cdot)) = K_1(\tilde{\tilde{u}}_1(\cdot), \tilde{\tilde{u}}_2(\cdot)) \geq v_1(\overline{z}_k), \tag{3.6}$$

$$K_2(u^*_{11}(\cdot), u^*_{22}(\cdot)) = K_2(\tilde{\tilde{u}}_1(\cdot), \tilde{\tilde{u}}_2(\cdot)) \geq v_2(\overline{z}_k),\ k = 0, 1, \dots, l-1,$$

and use the Theorem in 3.3. Inequalities (3.6) follow from the optimality of strategies $u^*_{11}(\cdot)$ and $u^*_{22}(\cdot)$ in the games Γ_1 and Γ_2, respectively. The proof is offered as an exercise. We have thus obtained the following theorem.

Theorem 4. *In the game Γ there always exists an equilibrium in penalty strategies. In the special case described above, the payoffs in this situation are equal to $K_i(u^*_{11}(\cdot), u^*_{22}(\cdot))$, where $u^*_{11}(\cdot)$ and $u^*_{22}(\cdot)$ are optimal strategies for Player 1 and 2 in the auxiliary zero-sum games Γ_1 and Γ_2, respectively.*

The meaning of penalty strategies is that a player causes his partner to follow the particular path in the game (the particular choices) by constantly threatening to shift to a strategy that is optimal in a zero-sum game against the partner. Although the set of equilibria in the class of penalty strategies is sufficiently representative, these strategies should not be regarded as very "good", because by penalizing the partner the player can penalize himself to a greater extent.

4 Multistage games with incomplete information

4.1. In Secs. 1–3 we considered multistage games with perfect information defined in terms of a finite tree graph $G = (X, F)$ in which each of the players exactly knows at his move the position or the tree node where he stays. That is why we were able to introduce the notion of player i's strategy as a single-valued function $u_i(x)$ defined on the set of personal positions X_i with its values in the set F_x. If, however, we wish to study a multistage game in which the players making their choices have no exact knowledge of positions in which they make their moves or may merely speculate that this position belongs to some subset A of personal positions X_i, then the realization of player's strategy as a function of position $x \in X_i$ turns out to be impossible. In this manner the wish to complicate the information structure of a game inevitably involves changes in the notion of a strategy. In order to provide exact formulations, we should first formalize the notion of information in the game. Here the notion of an information set plays an important role. This will be illustrated with some simple, already classical examples from texts on game theory [6].

Example 7. Zero-sum game. Player 1 selects at the first move a number from the set $\{1, 2\}$. The second move is made by Player 2. He is informed about Player 1's choice and selects a number from the set $\{1, 2\}$. The third move is again to be made by Player 1. He knows Player 2's choice, remembers his own choice and selects a number from the set $\{1, 2\}$. At this point the game terminates and Player 1 receives a payoff H (Player 2 receives a payoff $(-H)$, i.e. the game is zero-sum), where the function H is defined as follows:

$$H(1,1,1) = -3,\ H(2,1,1) = 4,$$

$$H(1,1,2) = -2,\ H(2,1,2) = 1,$$

$$H(1,2,1) = 2,\ H(2,2,1) = 1,$$

$$H(1,2,2) = -5,\ H(2,2,2) = 5, \tag{4.1}$$

The graph $G = (X, F)$ of the game is depicted in Fig. 8. The circles in the graph represent positions in which Player 1 makes a move, whereas the blocks represent positions in which Player 2 makes a move.

If the set X_1 is denoted by X, the set X_2 by Y and the elements of these sets by $x \in X$, $y \in Y$, respectively, then Player 1's strategy $u_1(\cdot)$ is given by the five-dimensional vector $u_1(\cdot) = \{u_1(x_1), u_1(x_2), u_1(x_3), u_1(x_4), u_1(x_5)\}$ prescribing the choice of one of the two numbers $\{1, 2\}$ in each position of the set X. Similarly, Player 2's strategy $u_2(\cdot)$ is a two-dimensional vector $u_2(\cdot) = \{u_2(y_1), u_2(y_2)\}$ prescribing the choice of one of the two numbers $\{1, 2\}$ in each of the positions of the set Y. Now, in this game Player 1 has 32 strategies and Player 2 has 4 strategies. The corresponding normal form of the game has a 32×4 matrix which (this follows from the Theorem in 2.1) has an equilibrium in pure strategies. It can be seen that the value of this game is 4. Player 1 has four optimal pure strategies: (2,1,1,1,2), (2,1,2,1,2), (2,2,1,1,2), (2,2,2,1,2). Player 2 has two optimal strategies: (1,1), (2,1).

Example 8. We shall slightly modify the information conditions of Example 7. The game is zero-sum. The first move is made by Player 1. He selects a number from the set $\{1, 2\}$. The second move is made by Player 2. He is informed about Player 1's choice and selects a number from the set $\{1, 2\}$. The

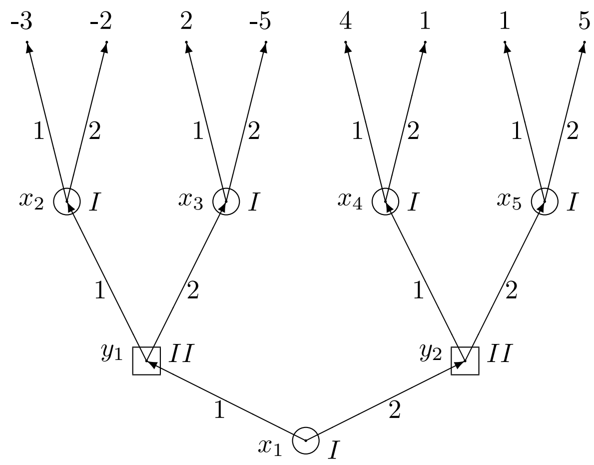


Figure 8:

third move is made by Player 1. Without knowledge of Player 2's choice and with no memory of his own choice he chooses a number of the set $\{1, 2\}$. At this point the game terminates and the payoff is determined in the same way as Example 7.

The graph of the game, $G = (X, F)$, remains unaffected. In the nodes x_2, x_3, x_4, x_5 (at the third move in the game) Player 1 cannot identify exactly the node in which he actually stays. With the knowledge of the priority of his move (the third move), he can be sure that he is not in the node x_1. In the graph G the nodes x_2, x_3, x_4, x_5 are traced by dashed line (Fig. 9).

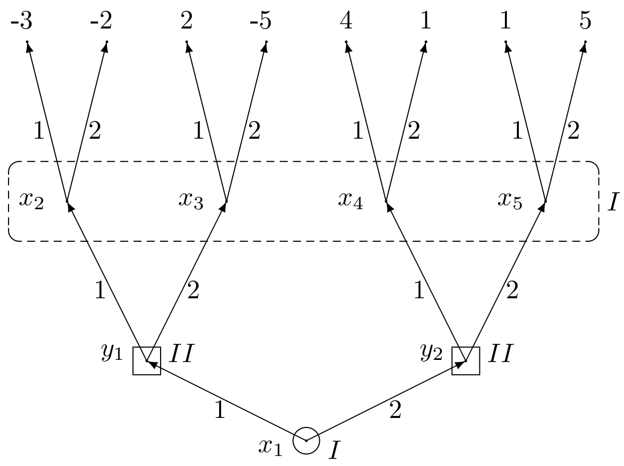


Figure 9:

The node x_1 is enclosed in a circle, which may be interpreted for Player 1 as an exact knowledge of this node when he stayed in it. The nodes y_1, y_2 are enclosed in blocks, which also means that Player 2 staying in one of them at his move can distinguish this node from the other. Combining the nodes x_2, x_3, x_4, x_5 into one set, we shall illustrate the fact that they are indistinguishable for Player 1.

The sets into which the nodes are collected in this way are called *information sets*.

We shall now describe strategies. The information state of Player 2 remains unchanged; therefore the set of his strategies is the same as in Example 7, i.e. it consists of four vectors (1,1), (1,2), (2,1), (2,2). The information state of Player 1 changed. At the third step of the game he knows only the number of this step, but does not know the position in which he stays. Therefore, he cannot realize the choice of the next node (or the choice of a number from the set $\{1, 2\}$) depending on the position in which he stays at the third step. For this reason irrespective of the actually realized position he has to choose at the third step one of the two numbers $\{1, 2\}$. Thus the strategy for him is a pair of numbers (i, j), $i, j \in \{1, 2\}$, where the number i is chosen in position x_1 while the number j at the third step is the same in all positions x_2, x_3, x_4, x_5. Now the choice of a number j turns out to be a function of the set and can be written as $u\{x_2, x_3, x_4, x_5\} = j$. In this game both players have four strategies and the matrix of the game is

$$
\begin{array}{c c}
 & \begin{array}{cccc} (1.1) & (1.2) & (2.1) & (2.2) \end{array} \\
\begin{array}{c} (1.1) \\ (1.2) \\ (2.1) \\ (2.2) \end{array} &
\begin{bmatrix} -3 & -3 & 2 & 2 \\ -2 & -2 & -5 & -5 \\ 4 & 1 & 4 & 1 \\ 1 & 5 & 1 & 5 \end{bmatrix}
\end{array}.
$$

This game has no equilibium in pure stategies. The value of the game is 19/7, an optimal mixed strategy for Player 1 is the vector $(0, 0, 4/7, 3/7)$, and an optimal mixed stategy for Player 2 is $(4/7, 3/7, 0, 0)$. The guaranteed payoff to Player 1 is reduced as compared to the one in Example 7. This is due to the degradation of his information state.

It is interesting to note that the game in Example 8 has a 4×4 matrix, whereas the game in Example 7 has a 32×4 matrix. The deterioration of available information thus reduces the size of the payoff matrix and hence facilitates the solution of the game itself. But this contradicts the wide belief that the deterioration of information results in complication of decision-making.

Modifying information conditions we may obtain other variants of the game described in Example 7.

Example 9. Player 1 chooses at the first move a number from the set $\{1, 2\}$. The second move is made by Player 2, who, without knowing Player 1's choice, chooses a number from the set $\{1, 2\}$. Further, the third move is made by Player 1. Being informed about Player 2's choice and with the memory of his own choice on the first step he chooses a number from the set $\{1, 2\}$. The payoff is determined in the same way as in Example 7 (Fig. 8).

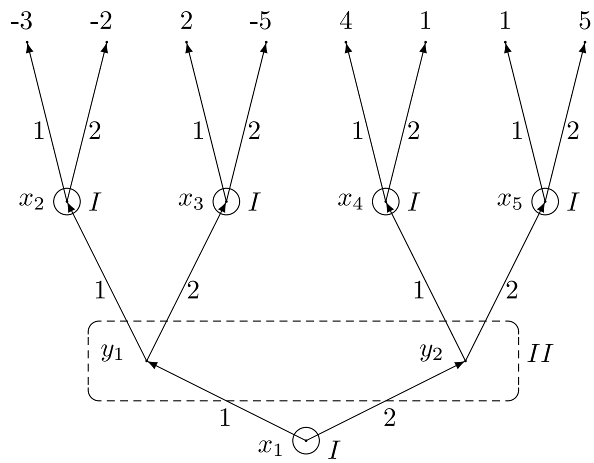


Figure 10:

Since on the third move the player knows the position in which he stays, the positions of the third level are enclosed in circles and the two nodes, in which Player 2 makes his move, are traced by the dashed line and are included in one information set.

Example 10. Player 1 chooses a number from the set $\{1, 2\}$ on the first move. The second move is made by Player 2 without being informed about Player 1's choice. Further, on the third move Player 1 chooses a number from the set $\{1, 2\}$ without knowing Player 2's choice and with no memory of his own choice at the first step. The payoff is determined in the same way as in Example 7 (Fig. 10).

Here the strategy of Player 1 consists of a pair of numbers (i, j), the i-th choice is at the first step, and j-th choice is at the third step; the strategy of Player 2 is a choice of number j at the second step of the game. Now, Player 1 has four strategies and Player 2 has two strategies. The game in normal form has a 4×2 matrix:

$$\begin{array}{c c} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} (1.1) \\ (1.2) \\ (2.1) \\ (2.2) \end{array} & \begin{bmatrix} -3 & 2 \\ -2 & -5 \\ 4 & 1 \\ 1 & 5 \end{bmatrix} \end{array}.$$

The value of the game is 19/7, an optimal mixed strategy for Player 1 is $(0, 0, 4/7, 3/7)$, whereas an optimal strategy for Player 2 is $(4/7, 3/7)$.

In this game the value is found to be the same as in Example 8, i.e. it turns out that the deterioration of information conditions for Player 2 did not improve the state of Player 1. This condition is random in nature and is accountable to special features of the payoff function.

Example 11. In the previous example the players fail to distinguish among

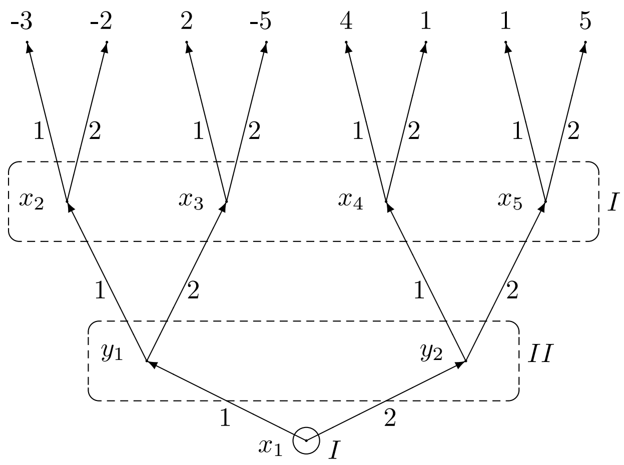


Figure 11:

positions placed at the same level of the game tree, but they do know the move to be made. It is possible to construct the game in which the players may reveal their ignorance to a greater extent.

Let us consider a zero-sum two-person game in which Player 1 is one person, whereas Player 2 is the team of two persons, A and B. All three persons are placed in different rooms and cannot communicate with each other. At the start of the game a mediator comes to Player 1 and suggests that he should choose a number from the set $\{1, 2\}$. If Player 1 chooses 1, the mediator suggests that A should be the first to make his choice. However, if Player 1 chooses 2, the mediator suggests that B should be the first to make his choice. Once these three numbers have been chosen, Player 1 wins an amount $K(x, y, z)$, where x, y, z are the choices made by Player 1 and members of Team 2, A and B, respectively. The payoff function $K(x, y, z)$ is defined as follows:

$$K(1,1,1) = 1,\ K(1,1,2) = 3,$$

$$K(1,2,1) = 7,\ K(1,2,2) = 9,$$

$$K(2,1,1) = 5,\ K(2,1,2) = 1,$$

$$K(2,2,1) = 6,\ K(2,2,2) = 7.$$

From the rules of the game it follows that when a member of the team, A or B, is suggested that he should make his choice he does not know whether he makes his choice at the second or at the third step of the game. The structure of the game is shown in Fig. 12.

Now, the information sets of Player 2 contain the nodes belonging to different levels, this means that Player 2 does not know on which step (second or third) he makes a mave. Here Player 1 has two strategies, whereas Player 2 has four

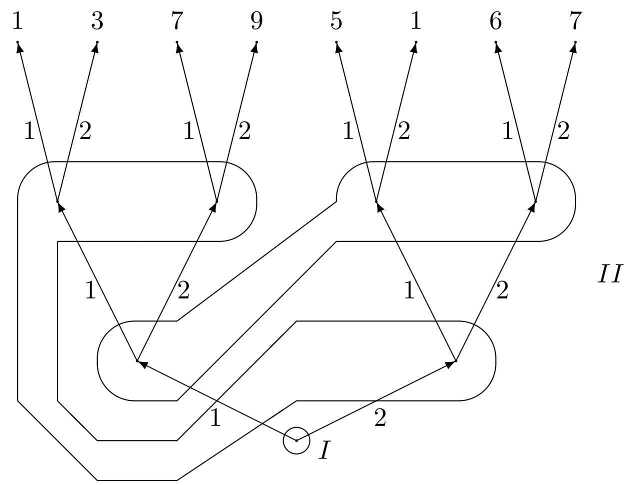


Figure 12:

strategies composed of all possible choices by members of the team, A and B, i.e. strategies for him are the pairs $(1,1),(1,2),(2,1),(2,2)$.

In order to understand how the elements of the payoff matrix are determined, we consider a situation $(2,(2,1))$. Since Player 1 has chosen 2, the mediator goes to B who, in accordance with strategy $(2,1)$, chooses 1. Then the mediator goes to A who chooses 2. Thus the payoff in situation $(2,(2,1))$ is $K(2,1,2)=1$. The payoff matrix for the game in normal form is

$$\begin{array}{c c}
 & \begin{array}{cccc} (1.1) & (1.2) & (2.1) & (2.2) \end{array} \\
\begin{array}{c} 1 \\ 2 \end{array} & \begin{bmatrix} 1 & 3 & 7 & 9 \\ 5 & 6 & 1 & 7 \end{bmatrix}
\end{array}.$$

The value of the game is $17/5$ and optimal mixed strategies for players 1 and 2 respectively are $(2/5,3/5)$, $(3/5,0,2/5,0)$.

Note that in multistage games with perfect information (see Theorem in 2.1) there exists a Nash equilibrium in the class of pure strategies, while in multistage zero-sum games there exists an equilibrium in pure strategies. Yet all the games with incomplete information discussed in Examples 8–11 have no equilibrium in pure strategies.

4.2. We shall now give a formal definition of a multistage *game in extensive form.*

Definition 5. *[7]. The n-person game in extensive form is defined by*

1) Specifying the tree graph $G=(X,F)$ with the initial vertex x_0 referred to as the initial position of the game.

2) Partition the sets of all vertices X into $n+1$ sets $X_1,X_2,\ldots,X_n,X_{n+1}$, where the set X_i is called the set of personal positions of the i-th player, $i=1,\ldots,n$, and the set $X_{n+1}=\{x: F_x=\oslash\}$ is called the set of final positions.

3) Specifying the vector function $K(x) = (K_1(x), \ldots, K_n(x))$ on the set of final positions $x \in X_{n+1}$; the function $K_i(x)$ is called the payoff to the i-th player.

4) Subpartition of the set X_i, $i = 1, \ldots, n$ into nonoverlapping subsets X_i^j referred to as information sets of the i-th player. In this case, for any position of one and the same information set the set of its subsequent vertices should contain one and the same number of vertices, i.e. for any $x, y \in X_i^j |F_x| = |F_y|$ ($|F_x|$ is the number of elements of the set F_x), and no vertex of the information set should follow another vertex of this set, i.e. if $x \in X_i^j$, then there is no other vertex $y \in X_i^j$ such that $y \in \hat{F}_x$ (see 1.2).

The definition of a multistage game with perfect information (see 1.4) is distinguished from the one given here only by condition 4, where additional partitions of players' personal positions X_i into information sets are introduced. As may be seen from the above examples, the conceptual meaning of such a partition is that when player i makes his move in position $x \in X_i$ in terms of incomplete information he does not know the position x itself, but knows that this position is in a certain set $X_i^j \subset X_i$ ($x \in X_i^j$). Some restrictions are imposed by condition 4 on the players' information sets. The requirement $|F_x| = |F_y|$ for any two vertices of the same information set are introduced to make vertices $x, y \in X_i^j$ indistinguishable. In fact, with $|F_x| \neq |F_y|$ Player i could distinguish among the vertices $x, y \in X_i^j$ by the number of arcs emanating therefrom. If one information set could have two vertices x, y such that $y \in \hat{F}_x$ this would mean that a play of the game can intersect twice an information set, but this in turn is equivalent to the fact that player j has no memory of the index of his move in this play which can hardly be conceived in the actual play of the game.

5 Behavior strategy

We shall continue examination of the game in extensive form and show that in the complete memory case for all players it has an equilibrium in behavior strategies.

5.1. For the purposes of further discussion we need to introduce some additional notions.

Definition 6. *The arcs incidental with x, i.e. $\{(x, y) : y \in F_x\}$, are called alternatives at the vertex $x \in X$.*

If $|F_x| = k$, at the vertex x there are k alternatives. We assume that if at the vertex x there are k alternatives, then they are designated by integers $1, \ldots, k$ with the vertex x bypassed in a clockwise sense. The first alternative at the vertex x_0 is indicated in an arbitrary way. If some vertex $x \neq x_0$ is bypassed in a clockwise sense, then an alternative which follows a single arc (F_x^{-1}, x) entering into x (Fig. 13) is called the first alternative at x.

Suppose that in the game Γ all alternatives are enumerated as above. Let A_k be the set of all vertices $x \in X$ having exactly k alternatives, i.e. $A_k = \{x :$

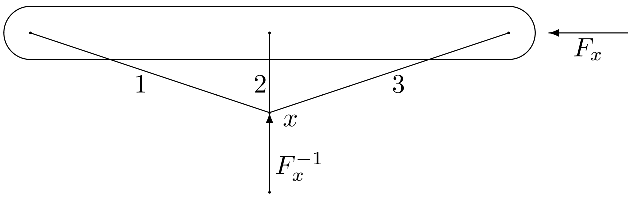


Figure 13:

$|F_x| = k\}$. Let $I_i = \{X_i^j : X_i^j \subset X_i\}$ be the set of all information sets for player i. By definition the *pure strategy* of player i means the function u_i mapping I_i into the set of positive numbers so that $u_i(X_i^j) \leq k$ if $X_i^j \subset A_k$. We say that the strategy u_i chooses alternative l in position $x \in X_i^j$ if $u_i(X_i^j) = l$, where l is the number of the alternative.

As in 1.4, we may show that to each situation $u(\cdot) = (u_1(\cdot), \ldots, u_n(\cdot))$ uniquely corresponds a play (path) ω, and hence the payoff in the final position of this play (path).

Let $x \in X_{n+1}$ be a final position and ω is the only path (F is the tree) leading from x_0 to x. The condition that the position y belongs to the path ω will be written as $y \in \omega$ or $y < x$.

Definition 7. *Position $x \in X$ is called possible for $u_i(\cdot)$ if there exists a situation $u(\cdot)$ containing $u_i(\cdot)$ such that the path ω containing position x is realized in situation $u(\cdot)$, i.e. $x \in \omega$. The information set X_i^j is called relevant for $u_i(\cdot)$ if some position $x \in X_i^j$ is possible for $u_i(\cdot)$.*

The set of positions, possible for $u_i(\cdot)$, is denoted by $Possu_i(\cdot)$, while the collection of information sets that are relevant for $u_i(\cdot)$ are denoted by $Relu_i(\cdot)$.

Lemma 1. *Position $x \in X$ is possible for $u_i(\cdot)$ if and only if $u_i(\cdot)$ chooses alternatives lying on the segment of the path ω_x from x_0 to x in all its information sets intersecting ω_x.*

Proof. Let $x \in Possu_i(\cdot)$. Then there exists a situation $u(\cdot)$ containing $u_i(\cdot)$ such that the path ω realized in this situation passes through x, which exactly means that in all its information sets intersecting the segment of the path ω_x the strategy $u_i(\cdot)$ chooses alternatives (arcs) belonging to ω_x.

Now let $u_i(\cdot)$ choose all alternatives for player i in ω_x. In order to prove the possibility of x for $u_i(\cdot)$ we need to construct a situation $u(\cdot)$ containing $u_i(\cdot)$ in which the path would pass through x. For player $k \neq i$ we construct a strategy $u_k(\cdot)$ which, in the information sets X_k^j intersecting the segment of the path ω_x, chooses alternatives (arcs) lying on this path and is arbitrary otherwise. Since each information set only intersects once the path ω, this can always be done. In the resulting situation $u(\cdot)$ the path ω necessarily passes through x; hence we have shown that $x \in Possu_i(\cdot)$.

5.2.

Definition 8. *The probability distribution over the set of pure strategies of player i which prescribes to every pure strategy $u_i(\cdot)$ the probability $q_{u_i}(\cdot)$ (for simplicity we write q_{u_i}) is called a mixed strategy μ_i for player i.*

The situation $\mu = (\mu_1, \dots, \mu_n)$ in mixed strategies determines the probability distribution over all plays (paths) ω (hence, in final positions X_{n+1} as well) by the formula

$$P_\mu(\omega) = \sum_u q_{u_1} \dots q_{u_n} P_u(\omega),$$

where $P_u(\omega) = 1$ if the play (path) ω is realized in situation $u(\cdot)$ and $P_u(\omega) = 0$ otherwise.

Lemma 2. *Denote by $P_\mu(x)$ the probability that the position x is realized in situation μ. Then we have*

$$P_\mu(x) = \sum_{\{u(\cdot): x \in Possu_i(\cdot), i=1,\dots,n\}} q_{u_1} \dots q_{u_n} = \prod_{i=1}^{n} \sum_{\{u_i : x \in Possu_i\}} q_{u_i}. \tag{5.1}$$

The proof of this statement immediately follows from Lemma in 5.1. The mathematical expectation of the payoff $E_i(\mu)$ for player i in situation μ is

$$E_i(\mu) = \sum_{x \in X_{n+1}} K_i(x) P_\mu(x), \tag{5.2}$$

where $P_\mu(x)$ is computed by formula (5.1).

Definition 9. *Position $x \in X$ is possible for μ_i if there exists a mixed strategy situation μ containing μ_i such that $P_\mu(x) > 0$. The information set X_i^j for player i is called relevant for μ_i if some $x \in X_i^j$ is possible for μ_i.*

The set of positions, possible for μ_i, is denoted by $Poss\mu_i$ and the collection information sets essential for μ_i is denoted by $Rel\mu_i$

5.3. Examination of multistage games with perfect information (see 3.3) shows that strategies can be chosen at each step in a suitable position of the game, while in the solution of specific problems it is not necessary (and it is not feasible) previously determine a strategy, i.e. a complete set of the recommended behavior in all positions (informations sets), since such a rule (see Example in 2.2) "suffers from strong redudancy". The question now arises of whether a similar simplification is feasible in the games with incomplete information. In other words, is it possible to form a strategy as it arise at a suitable information set rather than to construct the strategy as a certain previously fixed rule for selection in all information sets? It turns out that in the general case it is not feasible. However, there exists a class of games in extensive form where such a simplification is feasible. Let us introduce the notion of a behavior strategy.

Definition 10. *By definition the behavior strategy β_i for player i means the rule which places each information set $X_i^j \subset A_k$ for player i in correspondence* *with a system of k numbers $b(X_i^j, \nu) \geq 0,\ \nu = 1, \ldots, k$ such that*

$$\sum_{\nu=1}^{k} b(X_i^j, \nu) = 1,$$

where $A_k = \{x : |F_x| = k\}$.

The numbers $b(X_i^j, \nu)$ can be interpreted as the probabilities of choosing alternative ν in the information set $X_i^j \subset A_k$ each position of which contains exactly k alternatives.

Any behavior strategy set $\beta = (\beta_1, \ldots, \beta_n)$ for n players determines the probability distribution over the plays (paths) of the game and in final positions as follows:

$$P_\beta(\omega) = \prod_{X_i^j \cap \omega \neq \oslash, \nu \in \omega} b(X_i^j, \nu). \tag{5.3}$$

Here the product is taken over all X_i^j, ν such that $X_i^j \cup \omega \neq \oslash$ and the choice in the point $X_i^j \cap \omega$ of an alternative numbered as ν leads to a position belonging to the path ω.

In what follows it is convenient to interpret the notion of a "path" not only as a set of its component positions, but also as a set of suitable alternatives (arcs).

The expected payoff $E_i(\beta)$ in the behavior strategy situation $\beta = (\beta_1, \ldots, \beta_n)$ is defined to be the expectation

$$E_i(\beta) = \sum_{x \in X_{n+1}} K_i(x) P_\beta(\omega_x),\ i = 1, \ldots, n,$$

where ω_x is the play (path) terminating in position $x \in X_{n+1}$.

5.4. For every mixed strategy μ_i there can be a particular behavior strategy β_i.

Definition 11. *The behavior strategy β_i corresponding to player i's mixed strategy $\mu_i = \{q_{u_i}\}$ is the behavior strategy defined as follows.*

If $X_i^j \in Rel\mu_i$, then

$$b(X_i^j, \nu) = \frac{\sum_{\{u_i : X_i^j \in Relu_i,\ u_i(X_i^j) = \nu\}} q_{u_i}}{\sum_{\{u_i : X_i^j \in Relu_i\}} q_{u_i}}. \tag{5.4}$$

If $X_i^j \notin Rel\mu_i$, then on the set X_i^j the strategy β_i can be defined as distinct from (5.4) in an arbitrary way. (In the case $X_i^j \notin Rel\mu_i$ the denominator in (5.4) goes to zero.) For definiteness, let

$$b(X_i^j, \nu) = \sum_{\{u_i : u_i(X_i^j) = \nu\}} q_{u_i}. \tag{5.5}$$

We shall present the following result without proof.

Lemma 3. *Let β_i be a behavior strategy for player i and $\mu_i = \{q_{u_i}\}$ be a mixed strategy determined by*

$$q_{u_i} = \prod_{X_i^j} b(X_i^j, u_i(X_i^j)).$$

Then β_i is the behavior strategy corresponding to μ_i.

5.5.

Definition 12. *[7]. The game Γ is called a game with perfect recall for the ith player if for any $u_i(\cdot), X_i^j, x$ from the conditions $X_i^j \in Relu_i$ and $x \in X_i^j$ it follows that $x \in Possu_i$.*

From the definition it follows that in the perfect recall game for the ith player any position from the information set relevant for $u_i(\cdot)$ is also possible for $u_i(\cdot)$. The term "perfect recall" underlines the fact that, appearing in any one of his information sets the ith player can exactly reconstruct which of the alternatives (i.e. numbers) he has chosen on all his previous moves (by one-to-one correspondence) and remembers everything he has known about his opponents. The perfect recall game for all players becomes the game with perfect information if all its information sets contain one vertex each.

5.6.

Lemma 4. *Let Γ be a perfect recall game for all players with ω as a play in Γ. Suppose $x \in X_i^j$ is the final position of the path ω in which player i makes his move, and suppose he chooses in x an arc ν. Let*

$$T_i(\omega) = \{u_i : X_i^j \in Relu_i,\ u_i(X_i^j) = \nu\}.$$

If ω has no positions from X_i, then we denote by $T_i(\omega)$ the set of all pure strategies for player i. Then the play ω is realized only in those situations $u(\cdot) = (u_1(\cdot), \dots, u_n(\cdot))$ for which $u_i \in T_i(\omega)$.

Proof. Sufficiency. It suffices to show that if $u_i \in T_i(\omega)$, then the strategy u_i chooses all the arcs (alternatives) for player i appearing in the play ω (if player i has a move in ω). However, if $u_i \in T_i(\omega)$ then $X_i^j \in Relu_i$, and since the game Γ has perfect recall, $x \in Possu_i$ $(x \in \omega)$. Thus, the strategy u_i chooses all the alternatives for player i appearing in the play ω.

Necessity. Suppose the play ω is realized in situation $u(\cdot)$, where $u_i \notin T_i(\omega)$ for some i. Since $X_i^j \in Relu_i$, this means that $u_i(X_i^j) \neq \nu$. But then the path ω is not realized. This contradiction completes the proof of the lemma.

5.7.

Lemma 5. *Let Γ be a perfect recall game for all players. Suppose ν is an alternative (arc) in a play ω that is incidental to $x \in X_i^j$, where $x \in \omega$, and the next position for player i (if any) on the path ω is $y \in X_i^k$. Consider the sets S and T, where*

$$S = \{u_i : X_i^j \in Relu_i,\ u_i(X_i^j) = \nu\},$$
$$T = \{u_i : X_i^k \in Relu_i\}.$$

Then $S = T$.

Proof. Let $u_i \in S$. Then $X_i^j \in Relu_i$, and since Γ has perfect recall $x \in Possu_i$. By Lemma 5.1, it follows that the strategy u_i chooses all the arcs incidental to player i's positions on the path from x_0 to x, though $u_i(X_i^j) = \nu$. Thus, u_i chooses all the arcs incidental to Player i's positions on the path from x_0 to y, i.e. $y \in Possu_i$, $X_i^k \in Relu_i$ and $u_i \in T$.

Let $u_i \in T$. Then $X_i^k \in Relu_i$, and since Γ has perfect recall $y \in Possu_i$. But this means that $x \in Possu_i$ and $u_i(X_i^j) = \nu$, i.e. $u_i \in S$. This completes the proof of the lemma.

5.8.

Theorem 5. *Let β be a situation in behavior strategies corresponding to a situation in mixed strategies μ in the game Γ (in which all positions have at least two alternatives). Then for*

$$E_i(\beta) = E_i(\mu),\ \ i = 1, \dots, n,$$

it is necessary and sufficient that Γ be a perfect recall game for all players.

Proof. Sufficiency. Let Γ be a perfect recall game for all players. Fix an arbitrary μ. It suffices to show that $P_\beta(\omega) = P_\mu(\omega)$ for all plays ω. If in ω there exists a position for player i belonging to the information set that is irrelevant for μ_i, then there is $X_i^j \in Rel\mu_i$, $X_i^j \cap \omega \neq \oslash$ such that the equality $b(X_i^j, \nu) = 0$ where $\nu \in \omega$ holds for the behavior strategy β_i corresponding to μ_i. Hence we have $P_\beta(\omega) = 0$. The validity of relationship $P_\mu(\omega) = 0$ in this case is obvious.

We now assume that all the information sets for the i-th player through which the play ω passes, are relevant for μ_i, $i = 1, 2, \dots, n$. Suppose player i in the play ω makes his succeeding moves in the positions belonging to the sets $X_i^1, \dots, X_i^s$ and chooses in the set X_i^j an alternative ν_j, $i = 1, \dots, s$. Then, by formula (5.4) and Lemma 5.7, we have

$$\prod_{j=1}^{s} b(X_i^j, \nu_j) = \sum_{u_i \in T_i(\omega)} q_{u_i}.$$

Indeed, since in the play ω player i makes his first move from the set X_i^1, it is relevant for all $u_i(\cdot)$, therefore the denominator in formula (5.4) for $b(X_i^1, \nu_1)$ is equal to 1. Further, by Lemma 5.7, the numerator $b(X_i^j, \nu_j)$ in formulas (5.4) is equal to the denominator $b(X_i^{j+1}, \nu_{j+1})$, $i = 1, \dots, s$. By formula (5.3), we finally get

$$P_\beta(\omega) = \prod_{i=1}^{n} \sum_{u_i \in T_i(\omega)} q_{u_i},$$

where $T_i(\omega)$ is determined in Lemma 4.6.

By Lemma in 5.6

$$P_\mu(\omega) = \sum_{u(\cdot)} q_{u_1} \dots q_{u_n} P_u(\omega) = \sum_{u: u_i \in T_i(\omega),\ i=1,\dots,n} q_{u_1} \dots q_{u_n},$$

i.e. $P_\mu(\omega) = P_\beta(\omega)$. This proves the sufficiency part of the theorem.

Necessity. Suppose Γ is not a perfect recall game for all players. Then there exist player i, a strategy u_i, an information set $X_i^j \in Relu_i$ and two positions $x, y \in X_i^j$ such that $x \in Possu_i$, $y \notin Possu_i$. Let u_i' be a strategy for player i for which $y \in Possu_i'$ and ω is the corresponding play passing through y in situation u'. Denote by μ_i a mixed strategy for player i which prescribes with probability $1/2$ the choice of strategy u_i or u_i'. Then $P_{u'\|\mu_i}(y) = P_{u'\|\mu_i}(\omega) = 1/2$ (here $u'\|\mu_i$ is a situation in which the pure strategy u_i' is replaced by the mixed strategy μ_i). From the condition $y \notin Possu_i$ it follows that the path $\overline{\omega}$ realized in situation $u'\|u_i$ does not pass through y. This means that there exists X_i^k such that $X_i^k \cap \omega = X_i^k \cap \overline{\omega} \neq \oslash$ and $u_i(X_i^k) \neq u_i'(X_i^k)$. Hence, in particular, it follows that $X_i^k \in Relu_i$, $X_i^k \in Relu_i'$. Let β_i be the behavior strategy corresponding to μ_i. Then $b(X_i^k, u_i'(X_i^k)) = 1/2$. We may assume without loss of generality that $u_i(X_i^j) \neq u_i'(X_i^j)$. Then $b(X_i^j, u_i'(X_i^j)) = 1/2$. Denote by β a situation in behavior strategies corresponding to a mixed strategy situation $u'\|\mu_i$. Then $P_\beta(\omega) \leq 1/4$, whereas $P_{u'\|\mu_i}(\omega) = 1/2$. This completes the proof of the theorem.

From Theorem 5.8, in particular, it follows that in order to find an equilibrium in the games with perfect recall it is sufficient to restrict ourselves to the class of behavior strategies.

5.9. *Battle of the Sexes with Incomplete information* [5]. This example is a version of Battle of the Sexes game. In this setting the second player (the woman) may go with him (man) even when their choices do not coincide. But the man (player 1) does not exactly know whether this will happens or not. Thus he did not know what game he is playing, game

$$A = \begin{array}{c} \\ a_1 \\ a_2 \end{array} \begin{array}{c} \begin{array}{cc} b_1 & \quad b_2 \end{array} \\ \begin{bmatrix} (4,2) & (0,0) \\ (0,0) & (2,4) \end{bmatrix} \end{array},$$

or game

$$B = \begin{array}{c} \\ a_1 \\ a_2 \end{array} \begin{array}{c} \begin{array}{cc} b_1 & \quad b_2 \end{array} \\ \begin{bmatrix} (4,2) & (3,1) \\ (2,3) & (2,4) \end{bmatrix} \end{array}.$$

Payoff in solution (a_1, b_2) in game B can be interpreted for Player 1 in the following way: Player 1 gets 3 since his partner is coming with less pleasure, similar payoff 1 shows this loss of pleasure for woman visiting Soccer game against her general tastes. Similarly one can explain the payoffs in game B in situation (a_2, b_1).

Suppose that games A and B can occur with probabilities $\frac{1}{2}$, $\frac{1}{2}$. And this probabilities are known to both players and also player 2 knows with probability 1 which game A or B she is playing.

Then this game can be presented as game in extensive form with incomplete information on the Figure 14.

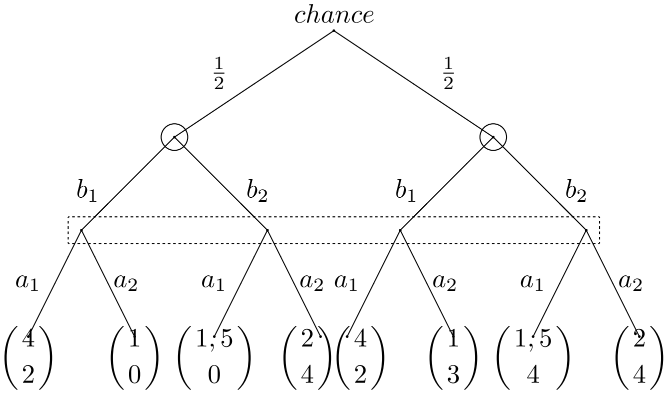


Figure 14:

The normal form of this game is

$$\begin{array}{cc} & \begin{array}{cccc} (b_1, b_1) & (b_1, b_2) & (b_2, b_1) & (b_2, b_2) \end{array} \\ \begin{array}{c} a_1 \\ a_2 \end{array} & \left[\begin{array}{cccc} (4;2) & (3,5;1,5) & (2;1) & (1,5;0,5) \\ (1;1,5) & (1;1,5) & (2,5;3,5) & (2,4) \end{array} \right]. \end{array}$$

In this game we get also two Nash equilibrium $[a_1; (b_1, b_1)]$ and $[a_2; (b_2, b_2)]$ with payoffs (4, 2) and (2, 4).

References

[1] Rochet, J. C. *Selection of unique equilibrium payoff for extension games with perfect information.* Mimeo, Universite de Paris, ix, 1980.

[2] Fudenberg, D. and J. Tirole. *Game theory*, p. 580. The MIT Press, Cambridge, 1992.

[3] Dixit, A. K. and S. Skeath. *Games of Strategy.* Norton, 1999.

[4] Petrosjan, L. A., N. A. Zenkevich and E. V. Shevkoplyas. *Game Theory*, p. 424. BCV-press, Saint-Petersburg, 2012.

[5] Peters, H. *Game Theory*, p. 364. Springer, New York, 2008.

[6] McKinsey, J. C. *Introduction to the Theory of Games*, p. 371. McGraw-Hill, N.Y., 1952.

[7] Kuhn, H. In *Annals of Mathematics Studies*. Princ. Univ. Press, Princeton, 1953.

[8] Diubin, G. N. and Suzdal V. G. *Introduction to Applied Theory of Games*, p. 311. Nauka, Moscow, 1981.

[9] Owen, G. *Game Theory*, p. 230. Saunders, Philadelphia, 1968.

[10] van Damme, E. *Stability and Perfection of Nash equilibria*, p. 215. Springer–Verlag, Berlin, 1991.

[11] Petrosjan L. A. and Sedakov A. A.. *Vestnik of the St. Petersburg State University, Ser.10*, 4, 2014.

[12] Marschak T. and Selten R. *General Equilibrium with Price-Making Firms*. Berlin: Springer-Verlag, 1974.

Positional games. Part 2.

December 17, 2018

1 Cooperative multistage games with perfect information

1.1. In what follows as basic model we shall consider the game in extensive form with perfect information.

Suppose that finite oriented treelike graph G with the root x_0 is given.

For simplicity we shall use the following notations. Let x be some vertex (position). We denote by G(x) a subtree of G with root in x. We denote by $Z(x)$ immediate successors of x. As before the vertices y, directly following after x, are called alternatives in x ($y \in Z(x)$). The player who makes a decision in x (who selects the next alterative position in x), will be denoted by $i(x)$. The choice of player $i(x)$ in position x will be denoted by $\bar{x} \in Z(x)$.

Definition 1. *A game in extensive form with perfect information $\Gamma(x_0)$ is a graph tree $G(x_0)$, with the following additional properties:*

1. *The set of vertices (positions) is split up into $n+1$ subsets $X_1, X_2, \ldots, X_{n+1}$, which form a partition of the set of all vertices of the graph tree $G(x_0)$. The vertices (positions) $x \in X_i$ are called players i personal positions, $i = 1, \cdots, n$; vertices (positions) $x \in X_{n+1}$ are called terminal positions.*

2. *For each vertex $x \notin X_{n+1}$ and $y \in Z(x)$ define an arc (x,y) on the graph $G(x_0)$. On each arc (x, y) n real numbers (payoffs of players on this arc) $h_i(x, y)$, $i = 1, \ldots, n$, $h_i \geq 0$ are defined, and also terminal payoffs $g_i(x) \geq 0$, for $x \in X_{n+1}$, $i = 1, \ldots, n$.*

Definition 2. *A strategy of player i is a mapping $U_i(\cdot)$, which associate to each position $x \in X_i$ a unique alternative $y \in Z(x)$.*

Denote by $H_i(x; u_1(\cdot), \ldots, u_n(\cdot))$ the payoff function of player $i \in N$ in the subgame $\Gamma(x)$ starting from the position x.

$$H_i(x; u_1(\cdot), \ldots, u_n(\cdot)) = \sum_{k=0}^{l-1} h_i(x_k, x_{k+1}) + g_i(x_l), \quad h_i \geq 0, \quad g_i \geq 0 \qquad (1.1)$$

where $x_l \in X_{n+1}$ is the last vertex (position) in the path $\tilde{x} = (x_0, x_1, \dots, x_l)$ realized in subgame $\Gamma(x)$, and $x_0 = x$, when n-tuple of strategies $(u_1(\cdot), \dots, u_n(\cdot))$ is played.

Denote by $\bar{u}(\cdot) = (\bar{u}_1(\cdot), \dots, \bar{u}_n(\cdot))$ the n-tuple of strategies and the trajectory (path) $\bar{x} = (\bar{x}_0, \bar{x}_1, \dots, \bar{x}_l)$, $\bar{x}_l \in P_{n+1}$ such that

$$\max_{u_1(\cdot),\dots,u_n(\cdot)} \sum_{i=1}^{n} H_i(x_0; u_1(\cdot), \dots, u_n(\cdot)) =$$

$$= \sum_{i=1}^{n} H_i(x_0; \bar{u}_1(\cdot), \dots, \bar{u}_n(\cdot)) = \sum_{i=1}^{n} \left(\sum_{k=0}^{l-1} h_i(\bar{x}_k, \bar{x}_{k+1}) + g_i(\bar{x}_l) \right). \tag{1.2}$$

The path $\bar{x} = (\bar{x}_0, \bar{x}_1, \dots, \bar{x}_l)$ satisfying (1.2) we shall call “optimal cooperative trajectory”.

Define in $\Gamma(x_0)$ characteristic function in a classical way

$$V(x_0; N) = \sum_{i=1}^{n} \left(\sum_{k=0}^{l-1} h_i(\bar{x}_k, \bar{x}_{k+1}) + g_i(\bar{x}_l) \right),$$

$$V(x_0; \emptyset) = 0,$$

$$V(x_0; S) = Val\Gamma_{S, N\setminus S}(x_0),$$

where $Val\Gamma_{S,N\setminus S}(x_0)$ is a value of zero-sum game played between coalition S acting as first player and coalition $N \setminus S$ acting as player 2, with payoff of player S equal to

$$\sum_{i \in S} H_i(x_0; u_1(\cdot), \dots, u_n(\cdot)).$$

If the characteristic function is defined then we can define the set of imputations in the game $\Gamma(x_0)$

$$C(x_0) = \left\{ \xi = (\xi_1, \dots, \xi_n) : \quad \xi_i \geq V(x_0; \{i\}), \quad \sum_{i \in N} \xi_i = V(x_0; N) \right\},$$

the core $M(x_0) \subset C(x_0)$

$$M(x_0) = \left\{ \xi = (\xi_1, \dots, \xi_n) : \quad \sum_{i \in S} \xi_i \geq V(x_0; S), \quad S \subset N \right\} \subset C(x_0),$$

NM solution, Shapley value and other optimality principles of classical game theory. In what follows we shall denote by $M(x_0) \subset C(x_0)$ anyone of this optimality principles.

Suppose at the beginning of the game players agree to use the optimality principle $M(x_0) \subset C(x_0)$ as the basis for the selection of the ”optimal” imputation $\bar{\xi} \in M(x_0)$.

This means that playing cooperatively by choosing the strategy maximizing the common payoff each one of them is waiting to get the payoff $\bar{\xi}_i$ from the optimal imputation $\bar{\xi} \in M(x_0)$ after the end of the game (after the maximal common payoff $V(x_0; N)$ is really earned by the players).

But when the game Γ actually develops along the "optimal" trajectory $\bar{x} = (\bar{x}_0, \bar{x}_1, \dots, \bar{x}_l)$ at each vertex $\bar{x}_k$ the players find themselves in the new multistage game with perfect information $\Gamma_{\bar{x}_k}$, $k = 0, \dots, l$, which is the subgame of the original game Γ starting from $\bar{x}_k$ with payoffs

$$H_i(\bar{x}_k; u_1(\cdot), \dots, u_n(\cdot)) = \sum_{j=k}^{l-1} h_i(x_j, x_{j+1}) + g_i(x_l), \quad i = 1, \dots, n.$$

It is important to mention that for the problem (1.2) the Bellman optimality principle holds and the part $\bar{x}^k = (\bar{x}_k, \dots, \bar{x}_j, \dots, \bar{x}_l)$ of the trajectory $\bar{x}$, starting from $\bar{x}_k$ maximizes the sum of the payoffs in the subgame $\Gamma_{\bar{x}_k}$, i.e.

$$\max_{x_k, \dots, x_j, \dots, x_l} \sum_{i=1}^{n} \left[\sum_{j=k}^{l-1} h_i(x_j, x_{j+1}) + g_i(x_l) \right] = \sum_{i=1}^{n} \left[\sum_{j=k}^{l-1} h_i(\bar{x}_j, \bar{x}_{j+1}) + g_i(\bar{x}_l) \right],$$

which means that the trajectory $\bar{x}^k = (\bar{x}_k, \dots, \bar{x}_j, \dots, \bar{x}_l)$ is also "optimal" in the subgame $\Gamma_{\bar{x}_k}$.

Before entering the subgame $\Gamma_{\bar{x}_k}$ each of the players i have already earned the amount

$$H_i^{\bar{x}_k} = \sum_{j=0}^{k-1} h_i(\bar{x}_j, \bar{x}_{j+1}).$$

At the same time at the beginning of the game $\Gamma = \Gamma(x_0)$ the player i was oriented to get the payoff $\bar{\xi}_i$ – the ith component of the "optimal" imputation $\bar{\xi} \in M(x_0) \subset C(x_0)$. From this it follows that in the subgame $\Gamma_{\bar{x}_k}$ he is expected to get the payoff equal to

$$\bar{\xi}_i - H_i^{\bar{x}_k} = \bar{\xi}_i^{\bar{x}_k}, \quad i = 1, \dots, n$$

and then the question arises whether the new vector $\bar{\xi}^{\bar{x}_k} = (\bar{\xi}_1^{\bar{x}_k}, \dots, \bar{\xi}_i^{\bar{x}_k}, \dots, \bar{\xi}_n^{\bar{x}_k})$ remains to be optimal in the same sense in the subgame $\Gamma_{\bar{x}_k}$ as the vector $\bar{\xi}$ was in the game $\Gamma(\bar{x}_0)$. If this will not be the case, it will mean that the players in the subgame $\Gamma_{\bar{x}_k}$ will not orient themselves on the same optimality principle as in the game $\Gamma(\bar{x}_0)$ which may enforce them to go out from the cooperation by changing the chosen cooperative strategies $\bar{u}_i(\cdot)$, $i = 1, \dots, n$ and thus changing the optimal trajectory $\bar{x}$ in the subgame $\Gamma(\bar{x}_k)$. Try now formalize this reasoning.

Introduce in the subgame $\Gamma(\bar{x}_k)$, $k = 1, \dots, l$, the characteristic function $V(\bar{x}_k; S)$, $S \subset N$ in the same manner as it was done in the game $\Gamma = \Gamma(x_0)$.

Based on the characteristic function $V(\bar{x}_k; S)$ we can introduce the set of imputations

$$C(\bar{x}_k) = \left\{ \xi = (\xi_1, \ldots, \xi_n) : \quad \xi_i \geq V(\bar{x}_k; \{i\}), \quad \sum_{i \in N} \xi_i = V(\bar{x}_k; N) \right\},$$

the core $M(\bar{x}_k) \subset C(\bar{x}_k)$

$$M(\bar{x}_k) = \left\{ \xi = (\xi_1, \ldots, \xi_n) : \quad \sum_{i \in S} \xi_i \geq V(\bar{x}_k; S), \quad S \subset N \right\} \subset C(\bar{x}_k),$$

NM solution, Shapley value and other optimality principles of classical game theory. Denote by $M(\bar{x}_k) \subset C(\bar{x}_k)$ the optimality principle $M \subset C$ (which was selected by players in the game $\Gamma(x_0)$) considered in the subgame $\Gamma(\bar{x}_k)$.

If we suppose that the players in the game $\Gamma(x_0)$ when moving along the optimal trajectory $(\bar{x}_0, \bar{x}_1, \ldots, \bar{x}_l)$ follow the same ideology of optimal behaviour then the vector $\bar{\xi}^{\bar{x}_k} = \bar{\xi} - H^{\bar{x}_k}$ must belong to the set $M(\bar{x}_k)$ – the corresponding optimality principle in the cooperative game $\Gamma(\bar{x}_k)$, $k = 0, \ldots, l$.

It is clearly seen that it is very difficult to find games and corresponding optimality principles for which this condition is satisfied. Try to illustrate this on the following example.

Suppose that in the game Γ $\quad h_i(x_k, x_{k+1}) = 0, k = 0, \ldots, l-1$, $g_i(x_l) \neq 0$ (the game Γ is the game with terminal payoff). Then the last condition would mean that

$$\bar{\xi} = \bar{\xi}^{\bar{x}_k} \in M(\bar{x}_k), \quad k = 0, \ldots, l,$$

which gives us

$$\bar{\xi} \in \bigcap_{k=0}^{l} M(\bar{x}_k). \tag{1.3}$$

For $k = l$ we shall have that

$$\bar{\xi} \in M(\bar{x}_l).$$

But $M(\bar{x}_l) = C(\bar{x}_l) = \{g_i(\bar{x}_l)\}$. And this condition have to be valid for all imputations of the set $M(\bar{x}_0)$ and for all optimality principles $M(x_0) \subset C(x_0)$, which means that in the cooperative game with terminal payoffs the only reasonable optimality principle will be

$$\bar{\xi} = \{g_i(\bar{x}_l)\},$$

the payoff vector obtained at the end point of the cooperative trajectory in the game $\Gamma(x_0)$. At the same time the simplest examples show that the intersection (1.3) except the “dummy” cases, is void for the games with terminal payoffs.

How to overcome this difficulty. The plausible way of finding the outcome is to introduce a special rule of payments (stage salary) on each stage of the game in such a way that the payments on each stage will not exceed the common amount earned by the players on this stage and the payments received by the players starting from the stage k(in the subgame $\Gamma(\bar{x}_k)$) will belong to the same optimality principle as the imputation $\bar{\xi}$ on which players agree in the game $\Gamma(x_0)$ at the beginning of the game. Whether it is possible or not we shall consider now.

Introduce the notion of the imputation distribution procedure (IDP).

Definition 3. *Suppose that $\xi = \{\xi_1, \dots, \xi_i, \dots, \xi_n\} \in M(x_0)$. Any matrix $\beta = \{\beta_{ik}\}$, $i = 1, \dots, n$, $k = 0, \dots, n$ such that*

$$\xi_i = \sum_{k=0}^{l} \beta_{ik}, \tag{1.4}$$

is called the imputation distribution procedure (IDP).

Denote $\beta_k = (\beta_{1k}, \dots, \beta_{nk})$, $\beta(k) = \sum\limits_{m=0}^{k-1} \beta_m$. The interpretation of IDP β is: β_{ik} is the payment to player i on the stage k of the game Γ_{x_0}, i.e. on the first stage of the subgame $\Gamma(\bar{x}_k)$. From the definition (1.4) it follows that in the game $\Gamma(x_0)$ each player i gets the amount ξ_i, $i = 1, \dots, n$, which he expects to get as the ith component of the optimal imputation $\xi_i \in M(x_0)$ in the game $\Gamma(x_0)$.

The interpretation of $\beta_i(k)$ is: $\beta_i(k)$ is the amount received by player i on the first k stages of the game Γ_{x_0}.

Definition 4. *The optimality principle $M(x_0)$ is called time-consistent if for every $\xi \in M(x_0)$ there exists IDP β such that*

$$\xi^k = \xi - \beta(k) \in M(\bar{x}_k), \quad k = 0, 1, \dots, l. \tag{1.5}$$

Definition 5. *The optimality principle $M(x_0)$ is called strongly time-consistent if for every $\xi \in M(x_0)$ there exists IDP β such that*

$$\beta(k) \oplus M(\bar{x}_k) \subset M(x_0), \quad k = 0, 1, \dots, l.$$

Here $a \oplus A = \{a + a' : \ a' \in A, a \in R^n, A \subset R^n\}$.

The time-consistency of the optimality principle $M(x_0)$ implies that for each imputation $\xi \in M$ there exits such IDP β that if the payments on each arc $(\bar{x}_k, \bar{x}_{k+1})$ on the optimal trajectory $\bar{x}$ will be made to the players according to IDP β, in every subgame $\Gamma(\bar{x}_k)$ the players may expect to receive the payments $\bar{\xi}^k$ which are optimal in the subgame $\Gamma(\bar{x}_k)$ in the same sense as it was in the game $\Gamma(x_0)$.

The strongly time-consistency means that if the payments are made according to IDP β then after earning on the stage k amount $\beta(k)$ the players (if they are oriented in the subgame $\Gamma(\bar{x}_k)$ on the same optimality principle as in $\Gamma(x_0)$) start with reconsidering of the imputation in this subgame (but optimal) they will get as a result in the game $\Gamma(x_0)$ the payments according to some imputation, optimal in the previous sense, i.e. the imputation belonging to the set $M(x_0)$.

For any optimality principle $M(x_0) \subset C(x_0)$ and for every $\bar{\xi} \in M(x_0)$ we can define β_{ik} by the following formulas

$$\begin{aligned} \bar{\xi}_i^{\bar{x}_k} - \bar{\xi}_i^{\bar{x}_{k+1}} = \beta_{ik}, \quad i = 1, \dots, n, \quad k = 0, \dots, l-1, \\ \bar{\xi}_i^{\bar{x}_l} = \beta_{il}. \end{aligned} \tag{1.6}$$

From the definition it follows that

$$\sum_{k=0}^{l} \beta_{ik} = \sum_{k=0}^{l-1} (\bar{\xi}_i^{\bar{x}_k} - \bar{\xi}_i^{\bar{x}_{k+1}}) + \bar{\xi}_i^{\bar{x}_l} = \bar{\xi}_i^{\bar{x}_0} = \bar{\xi}_i.$$

And at the same time

$$\bar{\xi} - \beta(k) = \bar{\xi}^{\bar{x}_k} \in M(\bar{x}_k), k = 0, \dots, l.$$

The last inclusion would mean the time consistency $M(x_0)$.

Unfortunately the elements β_{ik} may take in many cases negative values, which may stimulate questions about the use of this payment mechanism in real life situations. Because this means that players in some cases have to pay to support time–consistency. We understand that this argument can be waved since the total amount the player gets in the game is equal to the component ξ_i of the optimal imputation, and he can borrow the money to cover the side payment β_{ik} on stage k.

But we have another approach which enables us to use only nonnegative IDP's, and get us result not only time-consistent, but strongly time-consistent solution. For this reason some integral transformation of characteristic function is needed.

Example 16. Time inconsistency of the Shapley Value. Consider 3 person the cooperative game. The following coalitions are possible $\{1,2,3\}$, $\{1,2\}$, $\{1,3\}$, $\{2,3\}$, $\{1\}$, $\{2\}$, $\{3\}$. The characteristic function has the form $v(\{1,2,3\}) = 6$, $v(\{1,2\}) = 2$, $v(\{1,3\}) = 2$, $v(\{2,3\}) = 2$, $v(\{1\}) = 1$, $v(\{2\}) = \frac{1}{2}$, $v(\{3\}) = \frac{1}{2}$. Computing the Shapley Value we get

$$Sh(x_0) : Sh_1 = \frac{26}{12},\ Sh_2 = \frac{23}{12},\ Sh_3 = \frac{23}{12}.$$

Suppose the game develops along the optimal cooperative trajectory, which corresponds to the choices (A, A, A), and coincides with the path $\bar{x} = (x_0, x_1, x_2, x_3)$. As we have seen $v(x_0; \{1,2,3\}) = 6$, $v(x_0; \{1,2\})= v(x_0; \{1,3\}) =v(x_0; \{2,3\}) = 2$, $v(x_0; \{1\}) = 1$, $v(x_0; \{2\}) =v(x_0; \{3\}) = \frac{1}{2}$. Consider now the subgame starting on cooperative trajectory from vertex $\bar{x}_1$. It can be easily seen that $v(\bar{x}_1; \{1,2,3\}) = 6$, $v(\bar{x}_1; \{1,2\}) = 1$, $v(\bar{x}_1; \{1,3\}) = 1$, $v(\bar{x}_1; \{2,3\}) = 4$,

$v(\bar{x}_1;\{1\}) = \frac{1}{3}$, $v(\bar{x}_1;\{2\}) = \frac{1}{2}$, $v(\bar{x}_1;\{3\}) = \frac{1}{2}$. And the Shapley Value in the subgame $\Gamma(\bar{x}_1)$ is equal to

$$Sh(\bar{x}_1) = \left(\frac{34}{36}, \frac{91}{36}, \frac{91}{36}\right),$$

and we see that $Sh(x_0) \neq Sh(\bar{x}_1)$.

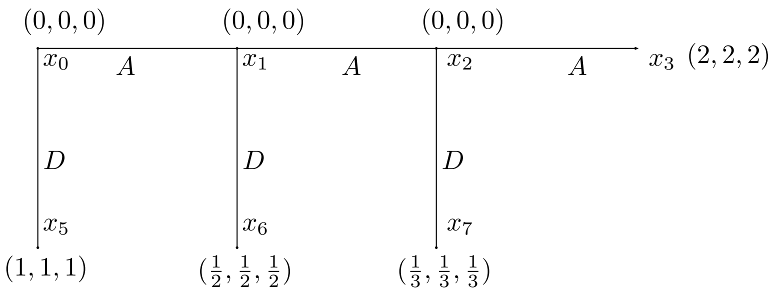


Figure 1:

Consider now the subgame starting on cooperative trajectory from vertex $\bar{x}_2$. It can be easily seen that $v(\bar{x}_2;\{1,2,3\}) = 4$, $v(\bar{x}_2;\{1,2\}) = 1$, $v(\bar{x}_2;\{1,3\}) = 1$, $v(\bar{x}_2;\{2,3\}) = 4$, $v(\bar{x}_2;\{1\}) = \frac{1}{3}$, $v(\bar{x}_2;\{2\}) = \frac{1}{2}$, $v(\bar{x}_2;\{3\}) = \frac{1}{2}$. And the Shapley Value in the subgame $\Gamma(\bar{x}_2)$ is equal to

$$Sh(\bar{x}_1) = \left(\frac{21}{18}, \frac{21}{18}, \frac{21}{18}\right),$$

and we see that $Sh(\bar{x}_0) \neq Sh(\bar{x}_1) \neq Sh(\bar{x}_2)$. It is obvious that $Sh(\bar{x}_3) = (3, 3, 3)$.

IPD for Shapley Value in this game can be easily calculated

$$Sh(0) = \left(\frac{44}{36}, -\frac{22}{36}, -\frac{22}{36}\right) + Sh(1),$$

$$Sh(1) = \left(-\frac{8}{36}, \frac{49}{36}, -\frac{41}{36}\right) + Sh(2),$$

$$Sh(2) = \left(-\frac{15}{18}, -\frac{15}{18}, \frac{30}{18}\right) + Sh(3),$$

$$Sh(3) = (3,\ 3,\ 3).$$

The strongly time consistency condition is more obligatory. We cannot even derive the formula like (1.6).

1.2. Now introduce the following functions

$$\beta_i^0 = \frac{\xi_i^0 \sum_{i=1}^n h_i(\overline{x}_0, \overline{x}_1)}{V(N; x_0)},$$

where $\xi^0 \in C(x_0)$

$$\beta_i^1 = \frac{\xi_i^1 \sum_{i=1}^n h_i(\overline{x}_1, \overline{x}_2)}{V(N, \overline{x}_1)}, \quad \xi^1 \in C(\overline{x}_1);$$
$$\dots$$
$$\beta_i^k = \frac{\xi_i^k \sum_{i=1}^n h_i(\overline{x}_k, \overline{x}_{k+1})}{V(N, \overline{x}_k)}, \quad \xi^k \in C(\overline{x}_k);$$
$$\dots$$
$$\beta_i^{l-1} = \frac{\xi_i^l \sum_{i=1}^n h_i(\overline{x}_{l-1}, \overline{x}_l)}{V(N, \overline{x}_l)}, \quad \xi^l \in C(\overline{x}_l). \tag{1.7}$$
$$\beta_i^l = q_i(\overline{x}_l).$$

Define the IDP $\overline{\beta}^k = \{\beta_i^k,\ i = 1, \dots, n\}$, $k = 0, \dots, l$. It is easily seen that $\overline{\beta}^k \geq 0$. Consider the formula (1.7). For different imputations $\xi^k \in C(\overline{z}_{k-1})$ we get different values of β_i^k and, hence, different values of $\overline{\beta}$. Let B^k be the set of all possible $\overline{\beta}^k$ for all $\xi^k \in C(\overline{z}_k)$, $k = 1, \dots, l$.

Consider the set:

$$\tilde{C}(z_0) = \{\overline{\xi} : \overline{\xi} = \sum_{k=0}^{l} \overline{\beta}^k,\ \overline{\beta}^k \in B^k\}$$

and the sets $\tilde{C}(\overline{z}_k) = \{\overline{\xi}^k : \overline{\xi}^k = \sum_{m=k}^l \overline{\beta}^m, \overline{\beta}^m \in B^m\}$.

The set $\tilde{C}(z_0)$ is called the regularized OP $\overline{C}(z_0)$ and, correspondingly, $\tilde{C}(\overline{z}_k)$ is a regularized OP $\overline{C}(\overline{z}_k)$.

We consider $\tilde{C}(z_0)$ as a new optimality principle in the game $\Gamma(z_0)$.

Theorem 1. *If the IDP β is defined as a $\overline{\beta}$, $k = 1, \dots, l$, then always:*

$$\beta(k) \oplus \tilde{C}(\overline{z}_k) \subset \tilde{C}(\overline{z}_0),$$

i.e. the OP $\tilde{C}(z_0)$ is strongly time consistent.

Proof. Suppose
$$\overline{\xi} \in \beta(k) \oplus \tilde{C}(\overline{z}_k),$$

then $\overline{\xi} = \beta(k) + \sum_{m=k}^l \beta^m$, for some $\beta^m \in B^m$, $m = k, \dots, l$.

But $\beta(k) = \sum_{m=0}^{k-1} \beta'^m$ for some $\beta'^m \in B^m$, $m = 0, \dots, k-1$.

Consider
$$(\beta'')^m = \begin{cases} \beta'^m, & \text{m=0,\dots,k-1,} \\ \beta^m, & \text{m=k,\dots,l,} \end{cases}$$

then $(\beta'')^m \in B^m$ and $\overline{\xi} = \sum_{m=0}^l (\beta'')^m$ and thus $\overline{\xi} \in \tilde{C}(z_0)$. The theorem is proved.

The defined IDP has the advantage (compared with β defined by (1.6)):

$$\beta_i^k \geq 0, \quad \sum_{i=1}^{n} \beta_i^k = 1, \quad k = 0, \ldots, l,$$

and thus

$$\sum_{i=1}^{n} \beta_i(\Theta) = \sum_{i=1}^{n} \sum_{k=0}^{\Theta-1} h_i(\bar{z}_{k+1}), \tag{1.8}$$

which is the actual amount to be divided between the players on the first $\Theta + 1$ stages and which is as it is seen by the formula (1.4) exactly equal to the amount earned by them on this stages.

Regularized game Γ_α. For every $\alpha \in M(x_0)$ define the noncooperative game $\Gamma_\alpha(x_0)$, which differs from the game $\Gamma(x_0)$ only by payoffs defined along optimal cooperative path $\bar{x} = (\bar{x}_0, \ldots, \bar{x}_l)$. Let $\alpha \in M(x_0)$. Define the imputation distribution procedure (IDP) as function $\beta_k = (\beta_1(k), \ldots, \beta_n(k))$, $k = 0, 1, \ldots, m$ such that

$$\alpha_i = \sum_{k=0}^{l} \beta_i(k). \tag{1.9}$$

Consider current subgame $\Gamma(\bar{x}_k)$ along the optimal path $\bar{x}$ and current imputation sets $C(\bar{x}_k)$. Let $\alpha^k \in C(\bar{x}_k)$.

Define by $H_i^\alpha(x_0; u_1(\cdot), \ldots, u_n(\cdot))$ the payoff function in the game $\Gamma_\alpha(x_0)$ and by $\bar{x} = \{\bar{x}_0, \ldots, \bar{x}_l\}$ the cooperative path.

Suppose $x = (x_1, x_2, \ldots, x_l)$ is the path resulting from the initial state x_0, when the situation $(u_1(\cdot), \ldots, u_n(\cdot))$ is used, and suppose that m is the maximal index for which $x_k = \bar{x}_k$ (the maximal number of stages in which the path coincides with cooperative path $\bar{x}$). Then

$$H_i^\alpha(x_0; u_1(\cdot), \ldots, u_n(\cdot)) = \sum_{k=0}^{m-1} \beta_{ik} + \sum_{k=m}^{l-1} h_i(x_k, x_k + 1) + g_i(x_l)$$

and

$$H_i^\alpha(x_0; \bar{u}_1(\cdot), \ldots, \bar{u}_n(\cdot)) = \alpha_i.$$

By the definition of the payoff function in the game $\Gamma_\alpha(x_0)$ we get that the payoffs along the optimal cooperative trajectory are equal to the components of the imputation $\alpha = (\alpha_1, \ldots, \alpha_n)$.

Definition 6. *The game $\Gamma_\alpha(x_0)$ is called regularization of the game $\Gamma(x_0)$ (α-regularization) if the IDP β is defined in such a way that*

$$\alpha_i^k = \sum_{j=k}^{l} \beta_i(j)$$

or $\beta_i(j) = \alpha_i^j - \alpha_i^{j+1}$, $i \in N$, $k = 0, 1, \ldots, l-1$, $\beta_i(l) = \alpha_i^l$, $\alpha_i^0 = \alpha_i$.

Theorem 2. *In the regularization of the game $\Gamma_\alpha(x_0)$ there exist a Nash equilibrium with payoffs $\alpha = (\alpha_1, \dots, \alpha_n)$.*

Let $y \in X_i$, $y \notin \bar{x}$, and let $\bar{x}_r \in \bar{x}$ is such vertex on cooperative path $\bar{x}$, that $(\Gamma_y^{-1})^m = \bar{x}_r$ and m is minimal. This means that y belongs to the trajectory $x' = (\bar{x}_0, ..., \bar{x}_r, x_{r+1}, ..., x_{l'})$ which has exactly r vertexes on cooperative trajectory $(y = x_{r+m})$. And suppose that $y \in X_j$. Consider now the subgame Γ_y, and on basis of this subgame construct a zero-sum game $\Gamma_y(k; N \setminus \{k\})$ with player $k \in N$ as first player and the coalition $N \setminus \{k\}$ as second player since the game is zero-sum, the aim of coalition $N \setminus \{k\}$ is to minimize the payoff of player k in the game $\Gamma_y(k; N \setminus \{k\})$.

Denote the optimal pure strategy of player $N \setminus \{k\}$ (the game is with perfect information and there exists a saddle point in pure strategies) by $\{\tilde{u}_l^{N\setminus\{k\}}(\cdot), l \in N \setminus \{k\}\}$.

Definition 7. *The strategy $\tilde{u}_i(\cdot)$ is called penalty strategy of player i if*

$$\tilde{u}_i(\bar{x}_s) = \bar{x}_{s+1} \text{ for } \bar{x}_x \in \bar{x},$$

$$\tilde{u}_i(y) = \tilde{u}_i^{N\setminus\{k\}}(y) \text{ for } i \neq k,\ y \notin \bar{x}, \tag{1.10}$$

$$\tilde{u}_k(y) \quad \text{arbitrary for } i = k,\ y \notin \bar{x},$$

We have to prove that

$$H_i^\alpha(x_0, \tilde{u}_i(\cdot)) = H_i^\alpha(x_0; \tilde{u}_1(\cdot), ..., \tilde{u}_{i-1}(\cdot), \tilde{u}_i(\cdot), \tilde{u}_{i+1}(\cdot), ..., \tilde{u}_n(\cdot)) \geq$$

$$\geq H_i^\alpha(x_0; \tilde{u}_1(\cdot), ..., \tilde{u}_{i-1}(\cdot), u_i(\cdot), \tilde{u}_{i+1}(\cdot), ..., \tilde{u}_n(\cdot)) =$$

$$= H_i^\alpha(x_0, \tilde{u}_i(\cdot)||u_i(\cdot)), \tag{1.11}$$

for all $u_i(\cdot)$, $i \in N$.

Here two cases are possible.

Case 1. $u_i(\bar{x}_k) = \tilde{u}_i(\bar{x}_k)$, for all $x_k \in \bar{x}$, then in this case $H_i^\alpha(x_0, \tilde{u}_i(\cdot))$= $H_i^\alpha(x_0, \tilde{u}_i(\cdot)||u_i(\cdot))$, $i = 1, ..., n$, and (1.11) holds.

Case 2. There exists $\bar{x}_k \in \bar{x} \cap X_i$, such that $u_i(\bar{x}_k) \neq \tilde{u}_i(\bar{x}_k) = \bar{x}_{k+1}$, suppose that k is the minimal integer for which $u_i(\bar{x}_k) \neq \tilde{u}_i(\bar{x}_k)$, then we shall have

$$H_i^\alpha(x_0; \tilde{u}_1(\cdot), ..., \tilde{u}_{i-1}(\cdot), u_i(\cdot), \tilde{u}_{i+1}(\cdot), ..., \tilde{u}_n(\cdot)) = \tag{1.12}$$

$$= H_i^\alpha(x_0, \tilde{u}_i(\cdot)||u_i(\cdot)) = \sum_{j=0}^{k-1} \beta_i(j) + \sum_{j=k}^{l-1} h_i(x_j, x_{j+1}) + g_i(x_l),$$

where $x_k = \bar{x}_k$.

But since the strategies $\tilde{u}_j(\cdot)$ (see 1.10) are constructed in such a way that in any subgame Γ_x, where $x \in Z(\bar{x}_k)$ players $j \neq i$ use behavior prescribed by strategies $\tilde{u}_j^{N\setminus\{i\}}(\cdot)$ optimal in zero-sum game played by coalition $N \setminus \{i\}$

against player i, player i cannot get in this subgame more than the value of this subgame which is equal to $v(\bar{x}_k;\{i\})$. And we have

$$\sum_{j=k}^{l-1} h_i(x_j, x_{j+1}) + g_i(x_l) \leq v(\bar{x}_k;\{i\}),$$

and finally

$$= H_i^{\alpha}(x_0, \tilde{u}_i(\cdot)||u_i(\cdot)) = \sum_{j=0}^{k-1} \beta_i(j) + \sum_{j=k}^{l-1} h_i(x_j, x_{j+1}) + g_i(x_l) \leq$$

$$\leq \sum_{j=0}^{k-1} \beta_i(j) + v(\bar{x}_k;\{i\}). \tag{1.13}$$

But

$$= H_i^{\alpha}(x_0, \tilde{u}_i(\cdot)) = \alpha_i = \sum_{j=0}^{l} \beta_i(j) = \sum_{j=0}^{k-1} \beta_i(j) + \sum_{j=k}^{l} \beta_i(j) \geq$$

$$\geq \sum_{j=0}^{k-1} \beta_i(j) + v(\bar{x}_k;\{i\}),\ i = 1, ..., n,\ k = 0, ..., l.$$

This follows from the condition that along the cooperative path we have

$$\alpha_i^k \geq V(\bar{x}_k;\{i\}), \quad i \in N, \quad k = 0, 1, \ldots, l.$$

since $\alpha^k = (\alpha_1^k, \ldots, \alpha_n^k) \in C(\bar{x}_k)$ is an imputation in $\Gamma(\bar{x}_k)$ (note that here $V(\bar{x}_k;\{i\})$ is computed in the subgame $\Gamma(\bar{x}_k)$ but not $\Gamma_{\alpha}(\bar{x}_k)$). In the same time

$$\alpha_i^k = \sum_{j=k}^{l} \beta_i(j)$$

and we get

$$\sum_{j=k}^{l} \beta_i(j) \geq V(\bar{x}_k;\{i\}), \quad i \in N, \quad k = 0, 1, \ldots, l. \tag{1.14}$$

And we have from (1.13), (1.14)

$$\alpha_i = H_i^{\alpha}(x_0, \tilde{u}_i(\cdot)) \geq \sum_{j=0}^{k-1} \beta_i(j) + v(\bar{x}_k;\{i\}) \geq$$

$$\geq H_i^{\alpha}(x_0, \tilde{u}_i(\cdot)||u_i(\cdot)),\ i \in N.$$

The theorem is proved.

Thus we constructed the Nash equilibrium in penalty strategies with payoffs $\alpha = (\alpha_1, \ldots, \alpha_n)$ and resulting cooperative path $\bar{x} = (\bar{x}_0, \ldots, \bar{x}_l)$.

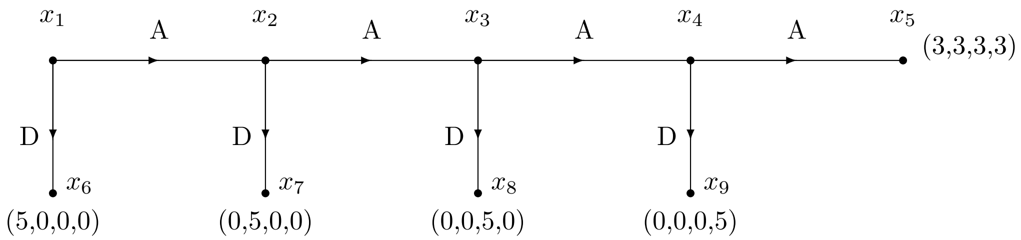


Figure 2: Game Γ_1

Example 17. In this example as an imputation we shall consider the Shapley value. Using the proposed regularization of the game we shall see that there exists a Nash equilibrium with payoffs equal to the components of Shapley value.

In the game Γ_1 (see Fig. 2), $N = \{1,2,3,4\}$, $P_1 = \{x_1\}$, $P_2 = \{x_2\}$, $P_3 = \{x_3\}$, $P_4 = \{x_4\}$, $P_5 = \{x_5, x_6, x_7, x_8, x_9\}$. $h(x_5) = (3,3,3,3)$, $h(x_6) = (5,0,0,0)$, $h(x_7) = (0,5,0,0)$, $h(x_8) = (0,0,5,0)$, $h(x_9) = (0,0,0,5)$. The cooperative path is

$$\bar{x} = \{x_1, x_2, x_3, x_4, x_5\}.$$

It can be easily seen that (D,D,D,D) is Nash equilibrium, but (A,A,A,A) is not Nash equilibrium.

Characteristic function of the game Γ_1 (C.f. of Γ_1)
$V_1(1,2,3,4) = 12$, $V_1(1,2,3) = 5$, $V_1(1,3,4) = 5$, $V_1(2,3,4) = 0$, $V_1(1,2,4) = 5$, $V_1(1,2) = 5$, $V_1(1,3) = 5$, $V_1(1,4) = 5$, $V_1(2,3) = 0$, $V_1(2,4) = 0$, $V_1(3,4) = 0$, $V_1(1) = 5$, $V_1(2) = 0$, $V_1(3) = 0$, $V_1(4) = 0$.

Shapley value for Γ_1 is

$$Sh^1 = (\frac{27}{4}, \frac{7}{4}, \frac{7}{4}, \frac{7}{4}).$$

Consider now the subgame $\Gamma_2 = \Gamma(x_2)$ (see Fig. 2).

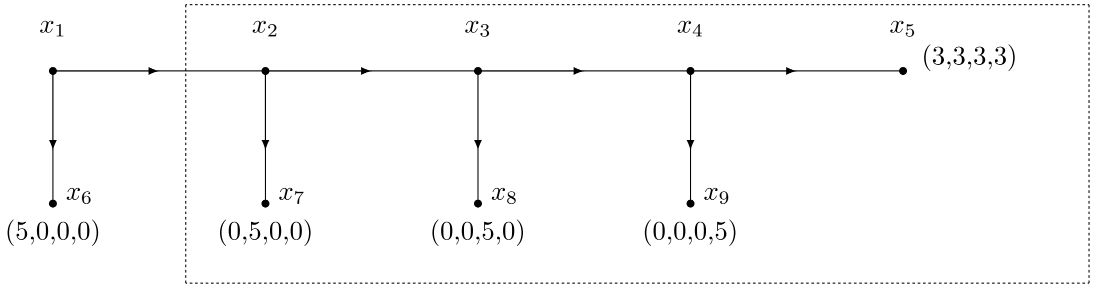


Figure 3: Subgame Γ_2

C.f. of Γ_2

$V_2(1,2,3,4) = 12$, $V_2(1,2,3) = 5$, $V_2(1,3,4) = 5$, $V_2(2,3,4) = 9$, $V_2(1,2) = 5$, $V_2(1,3) = 0$, $V_2(1,4) = 0$, $V_2(2,3) = 5$, $V_2(2,4) = 5$, $V_2(3,4) = 0$, $V_2(1) = 0$, $V_2(2) = 5$, $V_2(3) = 0$, $V_3(4) = 0$.

Shapley value for Γ_2 is

$$Sh^2 = (\frac{19}{12}, \frac{65}{12}, \frac{30}{12}, \frac{30}{12}),$$

and $Sh^1 \neq Sh^2$.

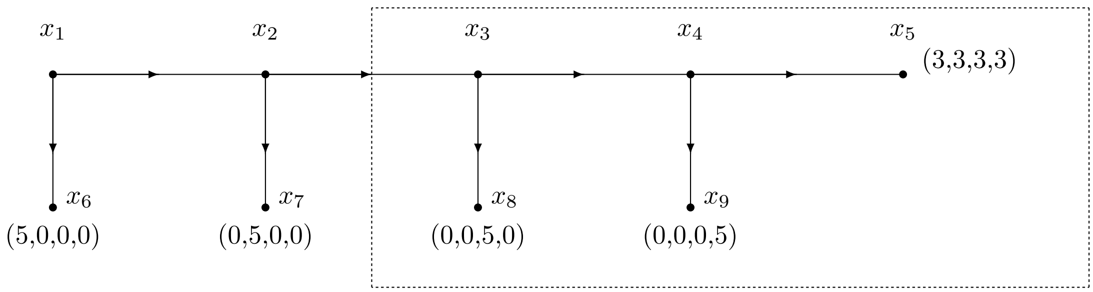


Figure 4: Subgame $\Gamma_3 = \Gamma(x_3)$

C.f. of Γ_3
$V_3(1,2,3,4) = 12$, $V_3(1,2,3) = 5$, $V_3(1,3,4) = 9$, $V_3(2,3,4) = 9$, $V_3(1,2,4) = 0$, $V_3(1,2) = 0$, $V_3(1,3) = 5$, $V_3(1,4) = 0$, $V_3(2,3) = 5$, $V_3(2,4) = 0$, $V_3(3,4) = 6$, $V_3(1) = 0$, $V_3(2) = 0$, $V_3(3) = 5$, $V_2(4) = 0$.

Shapley value for Γ_3 is

$$Sh^3 = (1, 1, \frac{90}{12}, \frac{30}{12}),$$

and $Sh^2 \neq Sh^3$.

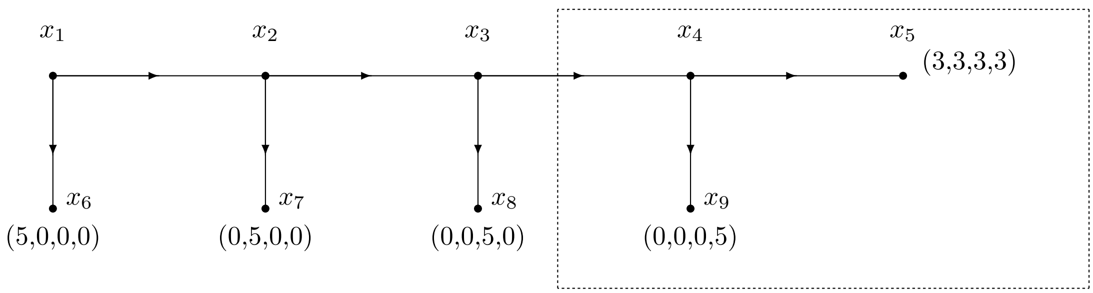


Figure 5: Subgame $\Gamma_4 = \Gamma(x_4)$

C.f. of Γ_4
$V_4(1,2,3,4) = 12$, $V_4(1,2,3) = 0$, $V_4(1,3,4) = 9$, $V_4(2,3,4) = 9$, $V_4(1,2,4) = 9$, $V_4(1,2) = 0$, $V_4(1,3) = 0$, $V_4(1,4) = 5$, $V_4(2,3) = 0$, $V_4(2,4) = 5$, $V_4(3,4) = 5$, $V_4(1) = 0$, $V_4(2) = 0$, $V_4(3) = 0$, $V_4(4) = 5$.

Shapley value for Γ_4 is

$$Sh^4 = (\frac{17}{12}, \frac{17}{12}, \frac{17}{12}, \frac{93}{12}),$$

and $Sh^4 \neq Sh^3$.

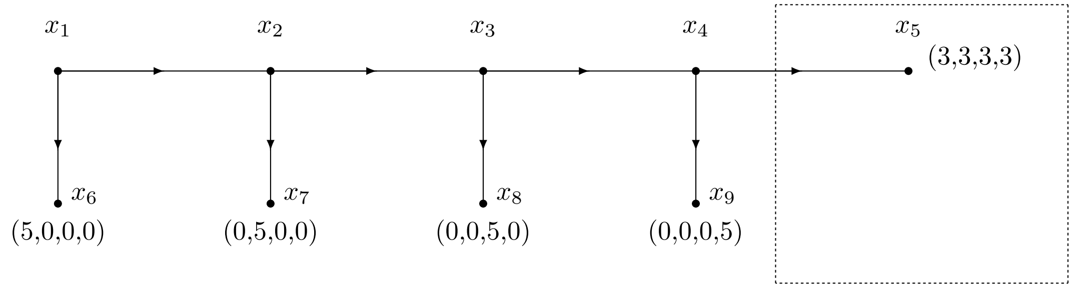


Figure 6: Subgame Γ_5

C.f. of Γ_5
$V_5(1,2,3,4) = 12$, $V_5(1,2,3) = V_5(1,3,4) = V_5(2,3,4) = V_5(1,2,4) = 9$, $V_5(1,2) = V_5(1,3) = V_5(1,4) = V_5(2,3) = V_5(2,4) = V_5(3,4) = 6$, $V_5(1) = V_5(2) = V_5(3) = V_5(4) = 3$.

Shapley value for Γ_4 is

$$Sh^5 = (3,3,3,3),$$

and $Sh^5 \neq Sh^4$.

Compute now the IDP (imputation distribution procedure)

$$Sh^1 = \beta_1 + Sh^2, \quad Sh^2 = \beta_2 + Sh^3, \dots, Sh^4 = \beta_4 + Sh^5,$$

$$\beta_1 = (Sh^1 - Sh^2), \beta_2 = (Sh^2 - Sh^3), \beta_3 = (Sh^3 - Sh^4), \beta_4 = (Sh^4 - Sh^5), \beta_5 = Sh^5.$$

$$\sum_{k=1}^{5} \beta_k = Sh^1, \sum_{k=2}^{5} \beta_k = Sh^2, \sum_{k=3}^{5} \beta_k = Sh^3,$$

$$\sum_{k=4}^{5} \beta_k = Sh^4, \sum_{k=5}^{5} \beta_k = Sh^5,$$

$$\beta_1 = (\frac{62}{12}, -\frac{44}{12}, -\frac{9}{12}, -\frac{9}{12}),$$

$$\beta_2 = (\frac{7}{12}, \frac{53}{12}, -\frac{60}{12}, 0),$$

$$\beta_3 = (-\frac{5}{12}, -\frac{5}{12}, \frac{73}{12}, -\frac{63}{12}),$$

$$\beta_4 = (-\frac{19}{12}, -\frac{19}{12}, -\frac{19}{12}, \frac{57}{12}),$$

$$\beta_5 = (3,3,3,3).$$

Regularization Γ_α of the game Γ_1 (when α is the Shapley value in Γ_1) and Nash equilibrium strategically supported cooperation

$$\begin{pmatrix} 62/12 \\ -44/12 \\ -9/12 \\ -9/12 \end{pmatrix} \quad \begin{pmatrix} 7/12 \\ 53/12 \\ -60/12 \\ 0 \end{pmatrix} \quad \begin{pmatrix} -5/12 \\ -5/12 \\ 73/12 \\ -63/12 \end{pmatrix} \quad \begin{pmatrix} -19/12 \\ -19/12 \\ -19/12 \\ 57/12 \end{pmatrix}$$

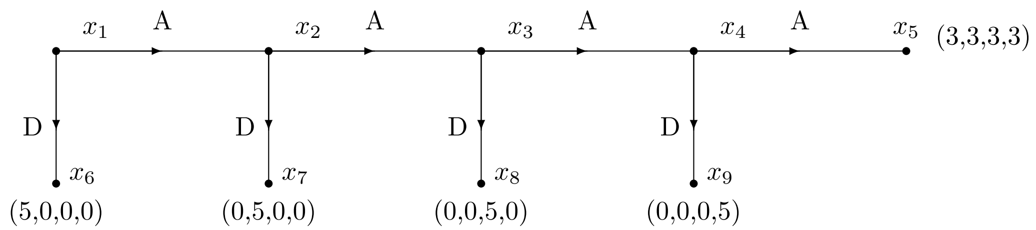


Figure 7: Game Γ_α

Here the payoffs β_1, β_2, β_3, β_4 are defined on arcs $(1,2)$, $(2,3)$, $(3,4)$, $(4,5)$ correspondingly.

We can see that the inequalities (1.14) hold in this game Γ_α.

$$\sum_{j=1}^{4} \beta_1(j) = \frac{62}{15} + \frac{7}{12} - \frac{5}{12} - \frac{19}{12} + 3 > 5 = V(\bar{x}_1; \{1\}),$$

$$\sum_{j=2}^{4} \beta_2(j) = \frac{53}{2} - \frac{5}{12} - \frac{19}{12} + 3 > 5 = V(\bar{x}_2; \{2\}),$$

$$\sum_{j=2}^{4} \beta_3(j) = \frac{73}{12} - \frac{19}{12} + 3 > 5 = V(\bar{x}_3; \{3\}),$$

$$\sum_{j=3}^{4} \beta_4(j) = \frac{57}{2} + 3 > 5 = V(\bar{x}_4; \{4\}).$$

This means that in Nash equilibrium (A,A,A,A) the payoffs in G_α are $(\frac{27}{4}, \frac{7}{4}, \frac{7}{4}, \frac{7}{4})$ exactly equal to Shapley value. Thus the computed Nash equilibrium supports the cooperative outcome (payoffs) in original game which are redistributed according to IDP guaranteeing the time-consistency of the Shapley value.

2 One-way flow two-stage network games

2.1. As we have seen in the theory of dynamic cooperative games, time-consistency of a solution is the key problem. Namely, having agreed on the particular solution before the game starts, players have to get the payoff prescribed by this solution at the end of the game. Such a problem is quite common for dynamic cooperative games, since during the game in the case of time-inconsistency, players may break initial agreement by their actions. Time-consistency of the cooperative solution based on a special payment scheme stimulates players to follow agreed upon cooperative behavior.

In two-stage network games were considered, in which players form a network at the first stage, and then at the second stage they choose admissible controls. In particular, it was proved that in the cooperative setting, the cooperative solution – the Shapley value [**?**]– is time-inconsistent.

2.2. *The model.* Let $N = \{1, \dots, n\}$ be a finite set of players, and g be a given network—the set of pairs $(i, j) \in N \times N$ where $(i, j) \in g$ means that there is a direct link connecting players i and j and such a link generates communication of player i with player j. The network under consideration is one-way flow network, that is any link (i, j) is direct link, i.e. $(i, j) \neq (j, i)$.

Consider a two-stage model. At the first stage, players choose their partners—players with whom they want to form links. Once the partners are chosen, a communication structure, i.e. a network is formed. At the second stage, players choose admissible control variables, which together with the formed network, affect their payoffs. Consider the model in details.

2.3. *First stage: network formation.* A network is formed as a result of simultaneous choices of all players. Let $M_i \subseteq N \setminus \{i\}$ be the set of players to whom player $i \in N$ can offer a link, and $a_i \in \{0, \dots, n-1\}$ be the maximal number of links which player i can offer. At the first stage, behavior of player $i \in N$ is a profile $g_i = (g_{i1}, \dots, g_{in})$ where

$$g_{ij} = \begin{cases} 1, & \text{if player } i \text{ offers a link to player } j \in M_i, \\ 0, & \text{otherwise,} \end{cases} \tag{2.1}$$

subject to the constraint:

$$\sum_{j \in N} g_{ij} \leq a_i. \tag{2.2}$$

From (2.1) we get $g_{ii} = 0$, $i \in N$, which excludes loops from the network, whereas the condition (2.2) shows that the number of "offers" is limited. Note that if $M_i = N \setminus \{i\}$, player i can offer a link to any player, whereas if $a_i = n-1$, he can offer any number of links.

Denote the set of all possible behaviors of player $i \in N$ at the first stage satisfying (2.1), (2.2) by G_i. The set $\prod_{i \in N} G_i$ is the set of behavior profiles at the first stage. Supposing that players choose their behaviors $g_i \in G_i$, $i \in N$, simultaneously and independently from each other, the behavior profile $(g_1, \dots, g_n)$ is formed. A resulting network g consists of directed links (i, j) s.t. $g_{ij} = 1$.

Define the closure of network g as an undirected network $\bar{g}$ where $\bar{g}_{ij} = \max\{g_{ij}, g_{ji}\}$. Denote neighbors of player i in the network g by $N_i(g) = \{j \in N \setminus \{i\} : (i,j) \in g\}$, whereas neighbors of player i in the closure $\bar{g}$ are denoted by $N_i(\bar{g}) = \{j \in N \setminus \{i\} : (i,j) \in \bar{g}\}$.

Example 18. Consider a four player case. Let $N = \{1,2,3,4\}$ and players choose the following behaviors: $g_1 = (0,0,0,1)$, $g_2 = (0,0,1,0)$, $g_3 = (0,1,0,1)$, $g_4 = (1,0,0,0)$. The network g consists of five links $g = \{(1,4),(2,3),(3,2),(3,4),(4,1)\}$, whereas its closure $\bar{g}$ consists of three undirected links $\bar{g} = \{(1,4),(2,3),(3,4)\}$. Note that, for instance, $N_4(g) = \{1\}$ while $N_4(\bar{g}) = \{1,3\}$.

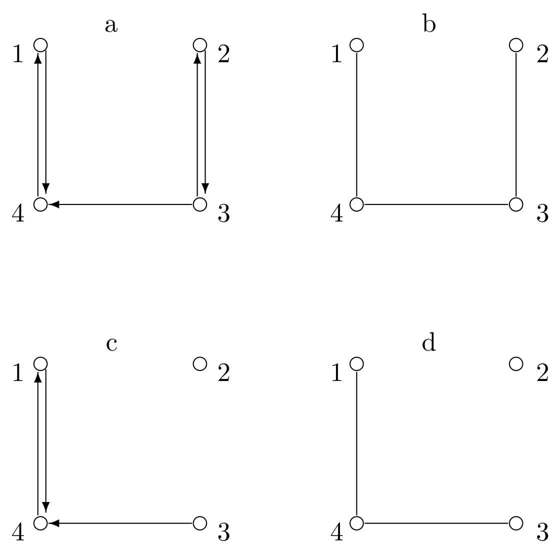


Figure 8:

2.4. *Second stage: controls.* To allow players to break formed links at the first stage (we introduce this possibility to punish neighbors from the complementary coalition in the case of zero-sum game which can appear at coalition formation stage), we define an n-dimensional profile $d_i(g)$ as follows:

$$d_{ij}(g) = \begin{cases} 1, \text{if player } i \text{ keeps the link formed at the first stage} \\ \text{with player } j \in N_i(g) \text{ in network } g, \\ 0, \text{otherwise.} \end{cases} \tag{2.3}$$

Denote all profiles $d_i(g)$ satisfying (2.3) by $D_i(g)$, $i \in N$. At the second stage players simultaneously and independently choose $d_i(g)$, $i \in N$, thus the profile $(d_1(g), \ldots, d_n(g))$ changes network g forming a new network which is denoted by g^d.

Example 19. Suppose that players choose their profiles g_i, $i \in N$ as in Example 1 forming the network $g = \{(1,4),(2,3),(3,2),(3,4),(4,1)\}$. Let

$d_1(g) = (0,0,0,1)$, $d_2(g) = (0,0,0,0)$, $d_3(g) = (0,0,0,1)$, $d_4(g) = (1,0,0,0)$, i.e. Player 1 keeps the link with Player 4, Player 2 breaks the link with Player 3, Player 3 breaks the link with Player 2, and Player 4 keeps the link with Player 1. Then we have a new network $g^d = \{(1,4),(3,4),(4,1)\}$. The closure $\bar{g}^d$ consists of two undirected links $\bar{g}^d = \{(1,4),(3,4)\}$ (see figure 8, c and d).

Also at the second stage player $i \in N$ chooses control u_i from a given set U_i. Then behavior of player $i \in N$ at the second stage is a pair $(d_i(g), u_i)$.

Payoff function K_i of player i depends on network g^d, his control u_i and controls u_j, $j \in N_i(\bar{g}^d)$ of his neighbors in the closure $\bar{g}^d$. More formally,

$$K_i(u_i, u_{N_i(\bar{g}^d)}) : U_i \times \prod_{j \in N_i(\bar{g}^d)} U_j \mapsto R^+, \quad i \in N,$$

is a real-valued function where notation $u_{N_i(\bar{g}^d)}$ means the profile of chosen controls u_j of all player $j \in N_i(\bar{g}^d)$ in network $\bar{g}^d$.

2.5. *Cooperation in one-way flow two-stage network games.* In this section we study the cooperative case: we allow players to coordinate their actions and choose behaviors jointly. Players being rational, choose their behaviors $g_i \in G_i$, $(d_i(g^d), u_i) \in D_i(g) \times U_i$, $i \in N$, to maximize the joint payoff, the value:

$$\sum_{i \in N} K_i(u_i, u_{N_i(\bar{g}^d)}). \tag{2.4}$$

It can be easily seen that to maximize the total sum (2.4) of players' payoffs (supposing that maximum in (2.4) exists), it is sufficient to form the network at the first stage without changing it at the second stage, i.e. $d_i(g) \equiv g_i$, for all $i \in N$ and

$$\max_{\substack{(g_i, d_i(g), u_i) \in G_i \times D_i(g) \times U_i, \\ i \in N}} \sum_{i \in N} K_i(u_i, u_{N_i(\bar{g}^d)}) =$$

$$\max_{\substack{(g_i, u_i) \in G_i \times U_i, \\ i \in N}} \sum_{i \in N} K_i(u_i, u_{N_i(\bar{g})}).$$

The profile (g_i^*, u_i^*), $i \in N$, maximizing (1.4) we call the cooperative profile. Behavior profile $(g_1^*, \dots, g_n^*)$ forms the network g^* and

$$\sum_{i \in N} K_i(u_i^*, u^*_{N_i(\bar{g}^*)}) =$$

$$\max_{\substack{(g_i, u_i) \in G_i \times U_i, \\ i \in N}} \sum_{i \in N} K_i(u_i, u_{N_i(\bar{g})}).$$

To allocate the maximal sum of players' payoffs according to some solution concept, one needs to construct a cooperative TU-game (N, V). The characteristic function V is defined in the sense of von Neumann and Morgenstern as:

$$V(N) = \sum_{i \in N} K_i(u_i^*, u^*_{N_i(\bar{g}^*)}),$$

$$V(S) = \max_{\substack{(g_i, d_i(g), u_i) \in G_i \times D_i(g) \times U_i, \\ i \in S}} \min_{\substack{(g_j, d_j(g), u_j) \in G_j \times D_j(g) \times U_j, \\ j \in N \setminus S}} \sum_{i \in S} K_i(u_i, u_{N_i(\bar{g}^d)}),$$

$$V(\emptyset) = 0,$$

where the network g is formed by profile $(g_1, \ldots, g_n)$ and the network g^d is formed by profile $(d_1(g), \ldots, d_n(g))$.

Consider a non-empty coalition $S \subset N$. Denote a network, formed by profiles g_i, $i \in N$, s.t. $g_j = (0, \ldots, 0)$ for all $j \in N \setminus S$, by g_S. Let $\bar{g}_S$ be the closure of g_S. For any controls u_i, $i \in S$ let controls $\tilde{u}_j(u_S)$, $j \in N \setminus S$, where $u_S = \{u_i\}$, $i \in S$, solve the following optimization problem

$$\sum_{i \in S} K_i \left(u_i, u_{N_i(\bar{g}_S) \cap S}, \tilde{u}_{(N \setminus S) \cap N_i(\bar{g}_S)}(u_S)\right) =$$

$$= \min_{u_j, j \in (N \setminus S) \cap N_i(\bar{g}_S)} \sum_{i \in S} K_i \left(u_i, u_{N_i(\bar{g}_S) \cap S}, u_{(N \setminus S) \cap N_i(\bar{g}_S)}\right).$$

Here $u_{N_i(\bar{g}_S) \cap S}$ is the profile of controls chosen by all neighbors of player i from coalition S in the network $\bar{g}_S$, and $\tilde{u}_{(N \setminus S) \cap N_i(\bar{g}_S)}(u_S)$ is a profile of controls chosen by all players from coalition $N \setminus S$ who are neighbors of player i in the network $\bar{g}_S$.

Theorem 3. *Suppose that functions K_i, $i \in N$, are non-negative and satisfy the following property: for any two networks g and g' s.t. $g' \subseteq g$, controls $(u_i, u_{N_i(\bar{g})}) \in U_i \times \prod_{j \in N_i(\bar{g})} U_j$ and player i, the inequality $K_i(u_i, u_{N_i(\bar{g})}) \geq K_i(u_i, u_{N_i(\bar{g}')})$ holds. Then for all $S \subset N$ we have*

$$V(S) = \max_{\substack{(g_i, u_i) \in G_i \times U_i, \\ i \in S}} \sum_{i \in S} K_i \left(u_i, u_{N_i(\bar{g}_S) \cap S}, \tilde{u}_{(N \setminus S) \cap N_i(\bar{g}_S)}(u_S)\right).$$

Proof. Consider the maxmin value for coalition $S \subset N$:

$$V(S) = \max_{\substack{(g_i, d_i(g), u_i) \in G_i \times D_i(g) \times U_i, \\ i \in S}} \min_{\substack{(g_j, d_j(g), u_j) \in G_j \times D_j(g) \times U_j, \\ j \in N \setminus S}} \sum_{i \in S} K_i(u_i, u_{N_i(\bar{g}^d)}).$$

Since the presence of a link $(j, i) \in g$, $i \in S$, $j \in N \setminus S$, increases payoff of coalition S according to the property formulated in the statement of the Proposition 1, therefore, player $j \in N \setminus S$, as a neighbor of i, changes his component $d_{ji}(g)$ from 1 to 0 in the profile $d_j(g)$, i.e. removes link (j, i) to minimize the payoff of coalition S. Thus, to minimize the value $\sum_{i \in S} K_i(u_i, u_{N_i(\bar{g}^d)})$ players from $N \setminus S$ remove all links with players from S and use controls $\tilde{u}_j(u_S)$, $j \in N \setminus S$. Note that it is not important for coalition S how players from its complement $N \setminus S$ connect with each other. Therefore, without loss of generality assuming that $d_j(g) = (0, \ldots, 0)$ for all $j \in N \setminus S$, we obtain

$$V(S) = \max_{\substack{(g_i, d_i(g), u_i) \in G_i \times D_i(g) \times U_i, \\ i \in S}} \sum_{i \in S} K_i \left(u_i, u_{N_i(\bar{g}_S) \cap S}, \tilde{u}_{(N \setminus S) \cap N_i(\bar{g}_S)}(u_S)\right).$$

To maximize the sum $\sum_{i\in S} K_i(u_i, u_{N_i(\bar{g}_S)\cap S}, \tilde{u}_{(N\setminus S)\cap N_i(\bar{g}_S)(u_S)})$, it is sufficient for players from coalition S to form the network at the first stage without changing it at the second stage, i.e. $d_i(g) \equiv g_i$, for all $i \in S$. Then we get

$$V(S) = \max_{\substack{(g_i,u_i)\in G_i\times U_i,\\ i\in S}} \sum_{i\in S} K_i(u_i, u_{N_i(\bar{g}_S)\cap S}, \tilde{u}_{(N\setminus S)\cap N_i(\bar{g}_S)}(u_S))$$

which proves the statement.

An imputation in the cooperative two-stage network game is a profile $\xi(V) = (\xi_1(V), \ldots, \xi_n(V))$ s.t. $\sum_{i\in N} \xi_i(V) = V(N)$ and $\xi_i(V) \geq V(\{i\})$ for all $i \in N$. We denote the set of all imputations in the game (N, V) by $I(V)$. A solution concept (or simply solution) of TU-game (N, V) is a rule that uniquely assigns a subset of $I(V)$ to the game (N, V). For example, if the solution concept is the Shapley value $\phi(V) = (\phi_1(V), \ldots, \phi_n(V))$, its components can be calculated as

$$\phi_i(V) = \sum_{S\subseteq N, i\in S} \frac{(|N| - |S|)!(|S| - 1)!}{|N|!} [V(S) - V(S \setminus \{i\})] \text{ for all } i \in N.$$

2.6. *Time-consistency problem.* In this section we study time-consistency of the Shapley value $\phi(V)$. We already found behavior profiles (g_i^*, u_i^*), $i \in N$, of players which maximize the sum (2.4) allowing players to get the value $V(N)$. Allocating $V(N)$ according to the Shapley value, we obtain the solution $\phi(V) = (\phi_1(V), \ldots, \phi_n(V))$. In other words, in the cooperative two-stage network game player $i \in N$ should receive the amount of $\phi_i(V)$ as his payoff.

After the first stage (after forming network g^*) players may recalculate the solution according to the same solution concept. To find the new, recalculated solution, one needs to consider the subgame (one-stage game) starting from the second stage, provided that players chose behavior profile $(g_1^*, \ldots, g_n^*)$ at the first stage, and therefore formed network g^*. Consider this subgame. The characteristic function for the subgame will depend on parameter—the network g^*—formed at the first stage, and we denote this function as $v(g^*, S)$ to stress the dependence from the network. The characteristic function $v(g^*, S)$ in the sense of von Neumann and Morgenstern is defined as follows:

$$v(g^*, N) = \sum_{i\in N} K_i(u_i^*, u^*_{N_i(\bar{g}^*)}) = V(N),$$

$$v(g^*, S) = \max_{\substack{(d_i(g^*),u_i)\in D_i(g^*)\times U_i,\\ i\in S}} \min_{\substack{(d_j(g^*),u_j)\in D_j(g^*)\times U_j,\\ j\in N\setminus S}} \sum_{i\in S} K_i(u_i, u_{N_i(\bar{g}^{d*})}),$$

$$v(g^*, \emptyset) = 0.$$

The following proposition can be proved similarly to Proposition 1.

Theorem 4. *If functions K_i, $i \in N$, are non-negative and satisfy the property stated in Proposition 1, the value $v(g^*, S)$ can be calculated by formula*

$$v(g^*, S) = \max_{\substack{u_i\in U_i,\\ i\in S}} \sum_{i\in S} K_i(u_i, u_{N_i(\bar{g}_S^*)\cap S}, \tilde{\tilde{u}}_{(N\setminus S)\cap N_i(\bar{g}_S^*)}(u_S)),$$

where $\tilde{\tilde{u}}_j(u_S)$, $j \in N \setminus S$, *solve the following optimization problem:*

$$\sum_{i \in S} K_i \left(u_i, u_{N_i(\bar{g}^*_S) \cap S}, \tilde{\tilde{u}}_{(N \setminus S) \cap N_i(\bar{g}^*_S)}(u_S) \right) =$$

$$= \min_{u_j, j \in (N \setminus S) \cap N_i(\bar{g}^*_S)} \sum_{i \in S} K_i \left(u_i, u_{N_i(\bar{g}^*_S) \cap S}, u_{(N \setminus S) \cap N_i(\bar{g}^*_S)} \right)$$

and $\bar{g}^*_S$ *is the closure of network* g^*_S*, formed by profiles* g^*_i*,* $i \in N$*, s.t.* $g^*_j = (0, \ldots, 0)$ *for all* $j \in N \setminus S$*.*

In the subgame, an imputation $\xi(g^*, v) = (\xi_1(g^*, v), \ldots, \xi_n(g^*, v))$ satisfies two conditions: $\sum_{i \in N} \xi_i(g^*, v) = v(g^*, N)$ and $\xi_i(g^*, v) \geq v(g^*, \{i\})$, $i \in N$. Recalculate players' payoffs in the subgame using the same solution concept—the Shapley value $\phi(g^*, v) = (\phi_1(g^*, v), \ldots, \phi_n(g^*, v))$, where its components can be computed as

$$\phi_i(g^*, v) = \sum_{S \subseteq N, i \in S} \frac{(|N| - |S|)!(|S| - 1)!}{|N|!} \left[v(g^*, S) - v(g^*, S \setminus \{i\}) \right],$$

for all $i \in N$.

Definition 8. *The Shapley value* $\phi(V)$ *is time-consistent if:*

$$\phi_i(V) = \phi_i(g^*, v), \quad i \in N. \tag{2.5}$$

The equality (2.5) means that if we use the imputation $\xi(V) = \phi(V)$ at the first stage, and then at the second stage recalculate players' payoffs according to the same solution concept $\xi(g^*, v) = \phi(g^*, v)$, i.e. calculate a new imputation $\xi(g^*, v) = \phi(g^*, v)$, subject to formed network g^*, players' payoffs prescribed by this imputation will not change. Since in most games the condition (1.5) is not satisfied (in our setting characteristic functions $V(S)$ and $v(g^*, S)$ are different), the time-consistency problem arises: player $i \in N$, who initially expected his payoff to be equal to $\phi_i(V)$, can receive different payoff $\phi_i(g^*, v)$. To avoid such situation in the game, we propose a stage payments mechanism—imputation distribution procedure for the Shapley value $\phi(V)$.

Definition 9. *Imputation distribution procedure for* $\phi(V)$ *in the cooperative two-stage network game is a matrix*

$$\beta = \begin{pmatrix} \beta_{11} & \beta_{12} \\ \vdots & \vdots \\ \beta_{n1} & \beta_{n2} \end{pmatrix},$$

s.t. $\phi_i(V) = \beta_{i1} + \beta_{i2}$ *for all* $i \in N$*.*

The value β_{ik} is a payment to player i at stage $k = 1, 2$. Therefore, the following payment scheme is applied: player $i \in N$ at the first stage of the game receives the payment β_{i1}, at the second stage of the game he receives the payment β_{i2} in order to his total payment received on both stages $\beta_{i1} + \beta_{i2}$ would be equal to the component of the Shapley value $\phi_i(V)$, which he initially wanted to get in the game as the payoff.

Definition 10. *Imputation distribution procedure β for the Shapley value $\phi(V)$ is time-consistent if for all $i \in N$*

$$\phi_i(V) - \beta_{i1} = \phi_i(g^*, v).$$

It is obvious that time-consistent imputation distribution procedure for $\phi(V)$ in the cooperative two-stage network game can be defined as follows:

$$\begin{aligned} \beta_{i1} &= \phi_i(V) - \phi_i(g^*, v), \\ \beta_{i2} &= \phi_i(g^*, v), \quad i \in N. \end{aligned} \tag{2.6}$$

2.7. *Numerical Example.* To illustrate the theoretical results obtained in the previous sections, consider a three-person game as an example. Let $N = \{1, 2, 3\}$ be the set of players. We suppose that Player 1 can establish a link with Player 3, Player 2 can establish links with Player 1 and Player 3, and, finally, Player 3 can establish links with Player 1 and Player 2. Therefore, we have: $M_1 = \{3\}$, $M_2 = \{1, 3\}$, $M_3 = \{1, 2\}$. Moreover, we suppose that each player can offer a limited number of links: Player 1 and Player 3 can offer only one link, while Player 2 can offer two links that is $a_1 = a_3 = 1$, $a_2 = 2$. Thus, at the first stage the sets of behaviors of players are: $G_1 = \{(0, 0, 0); (0, 0, 1)\}$, $G_2 = \{(0, 0, 0); (1, 0, 0); (0, 0, 1); (1, 0, 1)\}$; $G_3 = \{(0, 0, 0); (1, 0, 0); (0, 1, 0)\}$.

Let u_i be behavior of player $i \in N$ at the second stage, and $U_i = [0, \infty)$ for all $i \in N$. The payoff function of player i depends on players connected with player i as well as on players with whom player i established links and have the following form (the expression of the payoff function is justified in above mentioned papers and is used in network models of public goods):

$$K_i(g, u) = \ln\left(1 + u_i + \sum_{j \in N_i(\bar{g})} u_j\right) - c_i u_i - k|N_i(g)|,$$

where parameters $c_1 = 0.2$, $c_2 = 0.25$, $c_3 = 0.4$, $k = 0.75$, and network g is formed by the profile (g_1, g_2, g_3), and $u = (u_1, u_2, u_3)$. To find cooperative behavior, one needs to maximize the total payoff, i.e. to solve the optimization problem:

$$\max_{\substack{(g_i, u_i) \in G_i \times U_i, \\ i=1,2,3}} \sum_{i=1}^{3} K_i(g, u) = \max_{\substack{(g_i, u_i) \in G_i \times U_i, \\ i=1,2,3}} \sum_{i=1}^{3} \left[\ln\left(1 + u_i + \sum_{j \in N_i(\bar{g})} u_j\right) - c_i u_i - k|N_i(\bar{g})|\right].$$

From Table 1 we conclude that $V(N) = 3.8242$ which is attained at two different profiles:

$$\begin{array}{lll} g_1^* = (0,0,1), \quad u_1^* = 14, & & g_1^* = (0,0,0), \quad u_1^* = 14, \\ g_2^* = (1,0,0), \quad u_2^* = 0, & \text{and} & g_2^* = (1,0,0), \quad u_2^* = 0, \\ g_3^* = (0,0,0), \quad u_3^* = 0, & & g_3^* = (1,0,0), \quad u_3^* = 0, \end{array}$$

which form networks $\{(1, 3), (2, 1)\}$ and $\{(2, 1), (3, 1)\}$ at the first stage respectively.

Network g	$\sum_{i=1}^{3} K_i$	Network g	$\sum_{i=1}^{3} K_i$
$\emptyset$	1.7620	$\{(1,3),(3,2)\}$	2.2870
$\{(1,3)\}$	2.6915	$\{(2,1)\}$	2.3715
$\{(1,3),(2,1)\}$	3.8242	$\{(2,1),(2,3)\}$	3.2047
$\{(1,3),(2,1),(2,3)\}$	3.0742	$\{(2,1),(2,3),(3,1)\}$	3.0742
$\{(1,3),(2,1),(2,3),(3,1)\}$	2.3242	$\{(2,1),(2,3),(3,2)\}$	2.4547
$\{(1,3),(2,1),(2,3),(3,2)\}$	2.3242	$\{(2,1),(3,1)\}$	3.8242
$\{(1,3),(2,1),(3,1)\}$	3.0742	$\{(2,1),(3,2)\}$	3.2047
$\{(1,3),(2,1),(3,2)\}$	3.0742	$\{(2,3)\}$	2.4683
$\{(1,3),(2,3)\}$	2.2870	$\{(2,3),(3,1)\}$	2.2870
$\{(1,3),(2,3),(3,1)\}$	1.5370	$\{(2,3),(3,2)\}$	1.7183
$\{(1,3),(2,3),(3,2)\}$	1.5370	$\{(3,1)\}$	2.6915
$\{(1,3),(3,1)\}$	1.9415	$\{(3,2)\}$	2.4683

Table 1:

By the definition of characteristic function $V(S)$, find its values for all $S \subset N$. For $i \in N$ we have

$$V(\{i\}) = \max_{(g_i,d_i(g),u_i)\in G_i\times D_i(g)\times U_i} \min_{\substack{(g_j,d_j(g),u_j)\in G_j\times D_j(g)\times U_j,\\ j\neq i}} K_i(g,u) =$$

$$= \max_{(g_i,u_i)\in G_i\times U_i} K_i(g,u)|_{g_j=0,u_j=0,\ j\neq i} =$$

$$= \max_{u_i\in U_i} \left[\ln(1+u_i) - c_i u_i\right].$$

For all $i,\ j \in \{1,2,3\}$, such that either $i \in M_j$ or $j \in M_i$, $m = N \setminus \{i,j\}$,

$$V(\{i,j\}) = \max_{\substack{(g_i,d_i(g),u_i)\in G_i\times D_i(g)\times U_i\\ (g_j,d_j(g),u_j)\in G_j\times D_j(g)\times U_j}} \min_{(g_m,d_m(g),u_m)\in G_m\times D_m(g)\times U_m} [K_i(g,u)+K_j(g,u)] =$$

$$= \max_{\substack{(g_i,d_i(g),u_i)\in G_i\times D_i(g)\times U_i\\ (g_j,d_j(g),u_j)\in G_j\times D_j(g)\times U_j}} [K_i(g,u)+K_j(g,u)]_{g_m=0,u_m=0} =$$

$$= \max\{\max_{\substack{u_i\in U_i\\ u_j\in U_j}} \left[\ln(1+u_i) - c_i u_i + \ln(1+u_j) - c_j u_j\right];$$

$$\max_{\substack{u_i\in U_i\\ u_j\in U_j}} \left[2\ln(1+u_i+u_j) - c_i u_i - c_j u_j - k\right]\} =$$

$$= \max\{V(\{i\})+V(\{j\});\ \max_{\substack{u_i\in U_i\\ u_j\in U_j}} \left[2\ln(1+u_i+u_j) - c_i u_i - c_j u_j - k\right]\}.$$

Thus after solving the corresponding maximization problems, we obtain values of characteristic function $V(S)$:

S	$\{1,2,3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1\}$	$\{2\}$	$\{3\}$
$V(S)$	3.8242	2.0552	2.0552	1.6589	0.8094	0.6363	0.3163

The Shapley value $\phi(V) = (\phi_1(V), \phi_2(V), \phi_3(V))$, calculated for characteristic function $V(S)$, is

$$\phi(V) = (1.5179,\ 1.2331,\ 1.0731), \tag{2.7}$$

i.e. choosing behaviors jointly at both stages, players get the total payoff of 3.8242 and after allocation of the amount according to the Shapley value at the end of the game, each player gets $\phi_i(V)$, $i = 1, 2, 3$, as his payoff in the game.

To show that the Shapley value $\phi(V)$ is time-inconsistent, consider the subgame of the two-stage game, starting from the second stage, provided that players chose the cooperative behaviors at the first stage. Select the cooperative profile $g_1^* = (0, 0, 1)$, $g_2^* = (1, 0, 0)$, $g_3^* = (0, 0, 0)$, and $u^* = (u_1^*, u_2^*, u_3^*) = (14, 0, 0)$. The cooperative behaviors at the first stage (g_1^*, g_2^*, g_3^*) form the network $g^* = \{(1, 3), (2, 1)\}$. To prove that the Shapley value $\phi(V)$ is time-inconsistent it is sufficient to compute the Shapley value $\phi(g^*, v)$ and show that $\phi(V) \neq \phi(g^*, v)$. For this purpose calculate characteristic function $v(g^*, S)$ in the subgame for all $S \subset N$. Note that $v(g^*, S) = V(S) = 3.8242$. For $i \in \{1, 2, 3\}$, we have

$$v(g^*, \{i\}) = \max_{(d_i(g^*), u_i) \in D_i(g^*) \times U_i} \ \min_{\substack{(d_j(g^*), u_j) \in D_j(g^*) \times U_j, \\ j \neq i}} K_i(g^*, u) =$$

$$= \max_{(d_i(g^*), u_i) \in D_i(g^*) \times U_i} K_i(g^*, u)|_{d_j(g^*)=0, u_j=0,\ j \neq i} =$$

$$= \max_{u_i \in U_i} \left[\ln(1 + u_i) - c_i u_i\right] = V(\{i\}).$$

For all $i, j \in \{1, 2, 3\}$, $m = N \setminus \{i, j\}$ we get:

$$v(g^*, \{i, j\}) = \max_{\substack{(d_i(g^*), u_i) \in D_i(g^*) \times U_i \\ (d_j(g^*), u_j) \in D_j(g^*) \times U_j}} \ \min_{(d_m(g^*), u_m) \in D_m(g^*) \times U_m} \left[K_i(g^*, u) + K_j(g^*, u)\right] =$$

$$= \max_{\substack{(d_i(g^*), u_i) \in D_i(g^*) \times U_i \\ (d_j(g^*), u_j}} \left[K_i(g^*, u) + K_j(g^*, u)\right]_{d_m(g^*)=0, u_m=0} =$$

$$= \begin{cases} \max\limits_{\substack{u_i \in U_i \\ u_j \in U_j}} \left[\ln(1 + u_i) - c_i u_i + \ln(1 + u_j) - c_j u_j)\right], & \{i, j\} = \{2, 3\}, \\ \max\{ \max\limits_{\substack{u_i \in U_i \\ u_j \in U_j}} \left[\ln(1 + u_i) - c_i u_i + \ln(1 + u_j) - c_j u_j\right]; & \\ \quad \max\limits_{\substack{u_i \in U_i \\ u_j \in U_j}} \left[2 \ln(1 + u_i + u_j) - c_i u_i - c_j u_j - k\right]\}, & \text{otherwise.} \end{cases}$$

Thus after solving the corresponding maximization problems, we obtain values of characteristic function $v(g^*, S)$, $g^* = \{(1, 3), (2, 1)\}$:

S	$\{1,2,3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1\}$	$\{2\}$	$\{3\}$
$v(g^*, S)$	3.8242	2.0552	2.0552	0.9526	0.8094	0.6363	0.3163

Using the values $v(g^*, S)$, the Shapley value $\phi(g^*, v) = (\phi_1(g^*, v), \phi_2(g^*, v), \phi_3(g^*, v))$ is computed:

$$\phi(g^*, v) = (1.7533,\ 1.1154,\ 0.9554),$$

and from (2.7) we conclude that $\phi(V) \neq \phi(g^*, v)$. This shows time-inconsistency of the Shapley value $\phi(V)$. Time-consistent imputation distribution procedure β of the Shapley value $\phi(V)$ can be computed by formulas (1.6):

$$\beta = \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \\ \beta_{31} & \beta_{32} \end{pmatrix} = \begin{pmatrix} -0.2354 & 1.7533 \\ 0.1177 & 1.1154 \\ 0.1177 & 0.9554 \end{pmatrix}. \tag{2.8}$$

Similarly, it can be seen that the Shapley value $\phi(V)$ is time-inconsistent also for the second cooperative behavior profile: $g_1^* = (0,0,0)$, $g_2^* = (1,0,0)$, $g_3^* = (1,0,0)$, and $u^* = (u_1^*, u_2^*, u_3^*) = (14,0,0)$. The cooperative behaviors at the first stage (g_1^*, g_2^*, g_3^*) form the network $g^* = \{(2,1),(3,1)\}$. One can show that the characteristic function $v(g^*, S)$ in the subgame, calculated for the given network $g^* = \{(2,1),(3,1)\}$, coincides with the characteristic function $v(g, S)$ calculated for network $g = \{(1,3),(2,1)\}$. Therefore, given the network $g^* = \{(2,1),(3,1)\}$, we have time-inconsistency of the Shapley value $\phi(V)$, and time-consistent imputation procedure β for the Shapley value $\phi(V)$ will be the same as in (2.8).

2.8. Two-stage games with network formation at the first stage are considered. One of assumptions is that the payoff of any player depends only on his behavior and behavior of his neighbors in the network. The model deals with directed network that influenced the construction of characteristic function of the game. It is shown that the Shapley value—proposed cooperative solution of the two-stage game—is time-inconsistent, but with the use of a newly introduced payment mechanism—imputation distribution procedure—one can guarantee the realization of such solution in the game.

References

[1] *Basar T., Zaccour G.* Handbook of Dynamic Game Theory. New York: Springer-Verlag, 2018.

[2] *Yeung D. W. K., Petrosyan L. A.* Subgame Consistent Cooperation. A Comprehensive Treatse. Singapore: Springer-Verlag, 2016.

[3] *Petrosyan L. A., Zenkevich N. A.* Game Theory. (2nd ed.). Singapore: World Scientific Publishing, 2016.

[4] *Petrosyan L. A., Kuzutin D. V.* Stable solutions of positional games, 2008.

[5] *van Damme E.* Stability and Perfection of Nash equilibria. — Springer-Verlag, Berlin, 1991, p. 215.

[6] *Dresher M.* Games of strategy, theory and applications. — N.Y.: Prentice-Hall, Englewood Cliffs, 1961.

Zero-sum differential games. N-person differential games.

November 28, 2018

I. Zero-sum differential games

1 Differential zero-sum games with prescribed duration

Differential games are a generalization of multistage games to the case where the number of steps in a game is infinite (continuum) and the Players 1 and 2 (denoted as E and P, respectively) have the possibility of taking decisions continuously in time. In this setting the trajectories of players' motion are a solution of the system of differential equations whose right-hand sides depend on parameters that are under control of the players.

1.1. Let $x \in R^n$, $y \in R^n$, $u \in U \subset R^k$, $v \in V \subset R^l$, $f(x,u), g(y,v)$ be the vector functions of dimension n given on $R^n \times U$ and $R^n \times V$, respectively. Consider two systems of ordinary differential equations

$$\dot{x} = f(x,u), \tag{1.1}$$

$$\dot{y} = g(y,v) \tag{1.2}$$

with initial conditions x_0, y_0. Player P (E) starts his motion from the phase state x_0 (y_0) and moves in the phase space R^n in accordance with (1.1) and (1.2), choosing at each instant of time the value of parameter $u \in U (v \in V)$ to suit his objectives and in terms of information available in each current state.

The simplest to describe is the case of perfect information. In the differential game this means that at each time instant t the players choosing parameters $u \in U$, $v \in V$ know the time t and their own and the opponent's phase states x, y. Sometimes one of the players, say, Player P, is required to know at each current instant t the value of the parameter $v \in V$ chosen by Player E at the same instant of time. In this case Player E is said to be discriminated and the game is called the *game with discrimination* against Player E.

The parameters $u \in U$, $v \in V$ are called controls for the players P and E, respectively. The functions $x(t), y(t)$ which satisfy equations (1.1), (1.2) and initial conditions that are called *trajectories* for the players P, E.

1.2. Objectives in the differential game are determined by the payoff which may defined by the realized trajectories $x(t), y(t)$ in a variety of ways. For example, suppose it is assumed that the game is played during a prescribed time T. Let $x(T), y(T)$ be phase coordinates of players P and E at the time instant T the game terminates. Then the payoff to Player E is taken to be $H(x(T), y(T))$, where $H(x, y)$ is some function given on $R^n \times R^n$. In the specific case, when

$$H(x(T), y(T)) = \rho(x(T), y(T)), \tag{1.3}$$

where $\rho(x(T), y(T)) = \sqrt{\sum_{i=1}^{n}(x_i(T) - y_i(T))^2}$ is the Euclidean distance between the points $x(T), y(T)$, the game describes the process of pursuit during which the objective of Player E is to avoid Player P by moving a maximum distance from him by the time the game ends. In all cases the game is assumed to be a differential zero-sum game. Under condition (1.3), this means that the objective of Player P is to come within the shortest distance of Player E by the time T the game ends.

With such a definition, the payoff depends only on final states and the results obtained by each player during the game until the time T are not scored. It is of interest to state the problem in which the payoff to Player E is defined as a minimum distance between the players during the game:

$$\min_{0 \le t \le T} \rho(x(t), y(t)).$$

There exist games in which the constraint on the game duration is not essential and the game continues until the players obtain a particular result. Let an m-dimensional surface F be given in R^{2n}. This surface will be called terminal. Let

$$t_n = \{\min t : (x(t), y(t)) \in F\}, \tag{1.4}$$

i.e. t_n is the first time instant when the point $(x(t), y(t))$ falls on F. If for all $t \ge 0$ the point $(x(t), y(t)) \notin F$, then t_n is $+\infty$. For the realized paths $x(t), y(t)$ the payoff to Player E is t_n (the payoff to Player P is $-t_n$). In particular, if F is a sphere of radius $l \ge 0$ given by the equation

$$\sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} = l,$$

then we have the game of pursuit in which Player P seeks to come within a distance $l \ge 0$ to Player E as soon as possible. If $l = 0$, then the capture is taken to mean the coincidence of phase coordinates for the players P and E, in which case Player E seeks to postpone the capture time. Such games of pursuit are called the *time-optimal games of pursuit*.

The theory of differential games also deals with the problems of determining the set of initial states for the players from which Player P can ensure the capture of Player E within a distance l. And a definition is provided for the set of initial states of the players from which Player E can avoid in a finite time the encounter with Player P within a distance l. One set is called a *capture zone* (C, Z) and the other an *escape zone* (E, Z). It is apparent that these zones do not meet. However, a critical question arises of whether the closure of the union of the capture and the escape zones spans the entire phase space. Also the answer to this question is provided below, we now note that in order to adequately describe this process, it suffices to define the payoff as follows. If there exists $t_n < \infty$ (see (1.4)), then the payoff to Player E is -1. If, however, $t_n = \infty$, then the payoff is $+1$ (the payoff to Player P is equal to the payoff to Player E but opposite in sign, since the game is zero-sum). The games of pursuit with such a payoff are called the *pursuit games of kind*.

1.3. *Phase constraints.* If we further require that the phase point (x, y) would not leave some set $F \subset R^{2n}$ during the game, then we obtain a differential game with *phase constraints*. A special case of such a game is the "Life-line" game. The "Life-line" game is a zero-sum game of kind in which the payoff to Player E is $+1$ if he reaches the boundary of the set F ("Life-line") before Player P captures him. Thus, the objective of Player E is to reach the boundary of the set F before being captured by Player P (coming within a distance l, $l \geq 0$ with Player P). The objective of Player P, however, is to come within a distance l with Player E while the latter is still in the set F. It is assumed that Player P cannot abandon the set F.

1.4. *Example 1. (Simple motion.)* The game is played on a plane. Motions of the players P and E are described by the system of differential equations

$$\dot{x}_1 = u_1,\ \dot{x}_2 = u_2,\ u_1^2 + u_2^2 \leq \alpha^2,$$

$$\dot{y}_1 = v_1,\ \dot{y}_2 = v_2,\ v_1^2 + v_2^2 \leq \beta^2,$$

$$x_1(0) = x_1^0,\ x_2(0) = x_2^0,\ y_1(0) = y_1^0,\ y_2(0) = y_2^0, \alpha \geq \beta. \tag{1.5}$$

The physical implication of equation (1.5) is that the players P and E are moving in a plane at limited velocities, the maximum velocities α and β are constant in value and the velocity of Player E does not exceed the velocity of Player P.

Player P can change the direction of his motion (the velocity vector) by choosing at each time instant the control $u = (u_1, u_2)$ constrained by $u_1^2 + u_2^2 \leq \alpha^2$ (the set U). Similarly, Player E can also change the direction of his motion by choosing at each time instant the control $v = (v_1, v_2)$ constrained by $v_1^2 + v_2^2 \leq \beta^2$ (the set V). It is obvious that if $\alpha > \beta$, then the capture zone (C, Z) coincides with the entire space, i.e. Player P can always ensure the capture of Player E within any distance l in a finite time. To this end, it suffices to choose the motion with the maximum velocity α and to direct the velocity vector at each time instant t towards the pursued point $y(t)$, i.e. to carry out pursuit along the pursuit line. If $\alpha \leq \beta$, the escape set (E, Z) coincides with the entire space of the game except the points (x, y) for which $\rho(x, y) \leq l$. Indeed, if at the initial instant $\rho(x_0, y_0) > l$, then Player E can always avoid capture by moving away from Player P along the straight line joining the initial points x_0, y_0, the maximum velocity being β.

The special property manifested here will also be encountered in what follows. In order to form the control which ensures the avoidance of capture for Player E, we need only to know the initial states x_0, y_0 while, to form the control which ensures the capture of Player E in the case $\alpha > \beta$, Player P needs information on his own and the opponent's state at each current instant of time.

Example 2. The players P and E are the material points with unit masses moving on the plane under the control of the modulus-constrained and frictional forces. The equations of motion for the players are

$$\begin{gathered} \dot{x}_1 = x_3,\ \dot{x}_2 = x_4,\ \dot{x}_3 = \alpha u_1 - k_P x_3, \\ \dot{x}_4 = \alpha u_2 - k_P x_4,\ u_1^2 + u_2^2 \leq \alpha^2, \\ \dot{y}_1 = y_3,\ \dot{y}_2 = y_4,\ \dot{y}_3 = \beta v_1 - k_E y_3, \\ \dot{y}_4 = \beta v_2 - k_E y_4,\ v_1^2 + v_2^2 \leq \beta^2, \end{gathered} \tag{1.6}$$

where $(x_1, x_2), (y_1, y_2)$ are geometric coordinates, $(x_3, x_4), (y_3, y_4)$ are respectively momenta of the points P and E, k_P and k_E are friction coefficients, α and β are maximum forces which can be applied to the material points P and E. The motion starts from the states $x_i(0) = x_i^0$, $y_i(0) = y_i^0$, $i = 1, 2, 3, 4$. Here, by the state is meant not the locus of the players P and E, but their phase state in the space of coordinates and momenta. The sets U, V are the circles $U = \{u = (u_1, u_2) : u_1^2 + u_2^2 \leq \alpha^2\}$, $V = \{v = (v_1, v_2) : v_1^2 + v_2^2 \leq \beta^2\}$. This means that at each instant the players P and E may choose the direction of applied forces. However, the maximum values of these forces are restricted by the constants α and β. In this formulation as shown below, the condition $\alpha > \beta$ (power superiority) is not adequate for Player P to accomplish pursuit from any initial state.

1.5. We did not define yet the ways of selecting controls $u \in U$, $v \in V$ by the players P and E in terms of the incoming information. In other words, the notion of a strategy in the differential game remains to be defined.

Although there exist several approaches to this notion, we shall focus on those intuitively obvious game-theoretic properties which the notion must possess. As noted in Ch. 5, the strategy must describe the behavior of a player in all information states in which he may find himself during the game. In what follows the information state of each player will be determined by the phase vectors $x(t), y(t)$ at the current time instant t. Then it would be natural to regard the strategy for Player P (E) as a function $u(x, y, t)$ $(v(x, y, t))$ with values in the set of controls U (V). That is how the strategy is defined in [1]. Strategies of this type are called *synthesizing*. However, this method of defining a strategy suffers from some grave disadvantages. Indeed, suppose the players P and E have chosen strategies $u(x, y, t), v(x, y, t)$, respectively. Then, to determine the paths for the players, and hence the payoff (which is dependent of the paths), we substitute the functions $u(x, y, t), v(x, y, t)$ into equations (1.1),(1.2) in place of the control parameters u, v and integrate them with initial conditions x_0, y_0 on the time interval $[0, T]$. We obtain the following system of ordinary differential equations:

$$\dot{x} = f(x, u(x, y, t)), \quad \dot{y} = g(y, v(x, y, t)). \tag{1.7}$$

For the existence and uniqueness of the solution to system (1.7) it is essential that some conditions be imposed on the functions $f(x, u), g(y, v)$ and the strategies $u(x, y, t), v(x, y, t)$. The first group of conditions places no limitations on the players' capabilities, refers to the statement of the problem and is justified by the physical nature of the process involved. The case is different from the constraints on the class of functions (strategies) $u(x, y, t), v(x, y, t)$. Such constraints on the players' capabilities contradict the notions adopted in game theory that the players are at liberty to choose a behavior. In some cases this leads to substantial impoverishment of the sets of strategies. For example, if we restrict ourselves to continuous functions $u(x, y, t), v(x, y, t)$, the problems arise where there are no solutions in the class of continuous functions. The assumption of a more general class of strategies makes impossible the unique solution of system (1.7) on the interval $[0, T]$. At times, to overcome this difficulty, one considers the sets of strategies $u(x, y, t), v(x, y, t)$ under which the system (1.7) has a unique solution extendable to the interval $[0, T]$. However, such an approach (aside from the nonconstructivity of the definition of the strategy sets) is not adequately justified, since the set of all pairs of strategies $u(x, y, t), v(x, y, t)$ under which the system (1.7) has a unique solution is found to be nonrectangular.

1.6. We shall consider the strategies in the differential game to be *piecewise open-loop strategies.*

The piecewise open-loop strategy $u(\cdot)$ for Player P consists of a pair $\{\sigma, a\}$, where σ is some partition $0 = t'_0 < t'_1 < \ldots < t'_n < \ldots$ of the time interval $[0, \infty)$ by the points t'_k which have no finite accumulation points; a is the map which places each point t'_k and phase coordinates $x(t'_k), y(t'_k)$ in correspondence with some measureable open-loop control $u(t) \in U$ for $t \in [t'_k, t'_{k+1})$ (the measurable function $u(t)$ taking values from the set U). Similarly, the piecewise open-loop strategy $v(\cdot)$ for Player E consists of a pair $\{\tau, b\}$ where τ is some partition $0 = t''_0 < t''_1 < \ldots < t''_n < \ldots$ of the time interval $[0, \infty)$ by the points t''_k which have no accumulation points; b is the map which places each point t''_k and positions $x(t''_k), y(t''_k)$ in correspondence with some measurable open-loop control $v(t) \in V$ on the interval $[t''_k, t''_{k+1})$ (the measurable function $v(t)$ taking values from the set V). Using a piecewise open-loop strategy the player responds to changes in information not continuously in time, but at the time instants $t_k \in \tau$ which are determined by the player himself.

Denote the set of all piecewise open-loop strategies for Player P by P, and the set of all possible piecewise open-loop strategies for Player E by E.

Let $u(t), v(t)$ be a pair of measurable open-loop controls for the players P and E (measurable functions with values in the control sets U, V). Consider a system of ordinary differential equations

$$\dot{x} = f(x, u(t)), \ \dot{y} = g(y, v(t)), \ t \geq 0. \tag{1.8}$$

Impose the following constraints on the right-hand sides of system (1.8). The vector functions $f(x, u), g(y, v)$ are continuous in all their independent variables and are uniformly bounded, i.e. $f(x, u)$ is continuous on the set $R^n \times U$, while $g(y, v)$ is continuous on the set $R^n \times V$ and $\|f(x, u)\| \leq \alpha$, $\|g(y, v)\| \leq \beta$ (here

$\|z\|$ is the vector norm in R^n). Furthermore, the vector functions $f(x, u)$ and $g(y, v)$ satisfy the Lipschitz condition in x and y uniformly with respect to u and v, respectively, that is

$$\|f(x_1, u) - f(x_2, u)\| \leq \alpha_1 \|x_1 - x_2\|, \ u \in U,$$

$$\|g(y_1, v) - g(y_2, v)\| \leq \beta_1 \|y_1 - y_2\|, \ v \in V.$$

From the Karatheodory existence and uniqueness theorem it follows that, under the above conditions, for any initial states x_0, y_0 any measurable open-loop controls $u(t), v(t)$ given on the interval $[t_1, t_2]$, $0 \leq t_1 < t_2$, there exist unique vector functions $x(t)$,$y(t)$ which satisfy the following system of differential equations almost everywhere (i.e. everywhere except the set of measure zero). On the interval $[t_1, t_2]$

$$\dot{x}(t) = f(x(t), u(t)), \ \dot{y}(t) = g(y(t), v(t)) \tag{1.9}$$

and the initial conditions are $x(t_1) = x_0$, $y(t_1) = y_0$.

1.7. Let (x_0, y_0) be a pair of initial conditions for equations (1.8). The system $S = \{x_0, y_0; u(\cdot), v(\cdot)\}$, where $u(\cdot) \in P$, $v(\cdot) \in E$, is called a *situation* in the differential game. For each situation S there is a unique pair of paths $x(t), y(t)$ such that $x(0) = x_0$, $y(0) = y_0$ and relationships (1.9) hold for almost all $t \in [0, T]$, $T > 0$.

Indeed, let $u(\cdot) = \{\delta, a\}$, $v(\cdot) = \{\tau, b\}$. Furthermore, let $0 = t_0 < t_1 < \ldots < t_k < \ldots$ be a partition of the interval $[0, \infty)$ that is the union of partitions δ, τ. The solution to system (1.9) is constructed as follows. On each interval $[t_k, t_{k+1})$, $k = 0, 1, \ldots$, the images of the maps a, b are the measurable open-loop controls $u(t), v(t)$; hence on the interval $[t_0, t_1)$ the system of equations (1.9) with $x(0) = x_0$, $y(0) = y_0$ has a unique solution. On the interval $[t_1, t_2)$ with $x(t_1) = \lim_{t \to t_1 - 0} x(t)$, $y(t_1) = \lim_{t \to t_1 - 0} y(t)$ as initial conditions, we construct the solution to (1.9) by reusing the measurability of controls $u(t), v(t)$ as images of the maps a and b on the intervals $[t_k, t_{k+1})$, $k = 1, 2, \ldots$. Setting $x(t_2) = \lim_{t \to t_2 - 0} x(t)$, $y(t_2) = \lim_{t \to t_2 - 0} y(t)$ we continue this process to find a unique solution $x(t), y(t)$ such that $x(0) = x_0$, $y(0) = y_0$. Any path $x(t)$ ($y(t)$) corresponding to some situation $\{x_0, y_0, u(\cdot), v(\cdot)\}$ is called the path of the Player P (Player E).

1.8. *Payoff function.* As shown above, each situation $S = \{x_0, y_0; u(\cdot), v(\cdot)\}$ in piecewise open-loop strategies uniquely determines the paths $x(t), y(t)$ for Player P and E, respectively. The priority degree of these paths will be estimated by the payoff function K which places each situation in correspondence with some real number — a payoff to Player E. The payoff to Player P is $-K$ (this means that the game is zero-sum, since the sum of payoffs to players P and E in each situation is zero). We shall consider the games with payoff functions of four types.

Terminal payoff. Let there be given some number $T > 0$ and a function $H(x, y)$ that is continuous in (x, y). The payoff in each situation $S = \{x_0, y_0; u(\cdot), v(\cdot)\}$ is determined as follows:

$$K(x_0, y_0; u(\cdot), v(\cdot)) = H(x(T), y(T)),$$

where $x(T) = x(t)|_{t=T}$, $y(T) = y(t)|_{t=T}$ (here $x(t), y(t)$ are the paths of players P and E in a situation S). We have the game of pursuit when the function $H(x, y)$ is a Euclidean distance between the points x and y.

Minimum result. Let $H(x, y)$ be a real-valued continuous function. In the situation $S = \{x_0, y_0; u(\cdot), v(\cdot)\}$ the payoff to Player E is taken to be $\min_{0 \le t \le T} H(x(t), y(t))$, where $T > 0$ is a given number. If $H(x, y) = \rho(x, y)$ then the game describes the process of pursuit.

Integral payoff. Some manifold F of dimension m and a continuous function $H(x, y)$ are given in $R^n \times R^n$. Suppose in the situation $S = \{x_0, y_0; u(\cdot), v(\cdot)\}$, t_n is the first instant at which the path $(x(t), y(t))$ falls on F. Then

$$K(x_0, y_0; u(\cdot), v(\cdot)) = \int_0^{t_n} H(x(t), y(t))dt$$

(if $t_n = \infty$, then $K \equiv \infty$), where $x(t), y(t)$ are the paths of players P and E corresponding to the situation S. In the case $H \equiv 1, K = t_n$, we have the time optimal game of pursuit.

Qualitative payoff. The payoff function K can take only one of the three values $+1, 0, -1$ depending on a position of $(x(t), y(t))$ in $R^n \times R^n$. Two manifolds F and L of dimensions m_1 and m_2 respectively are given in $R^n \times R^n$. Suppose that in the situation $S = \{x_0, y_0; u(\cdot), v(\cdot)\}$, t_n is the first instant at which the path $(x(t), y(t))$ falls on F. Then

$$K(x_0, y_0; u(\cdot), v(\cdot)) = \begin{cases} -1, & if\ (x(t_n), y(t_n)) \in L, \\ 0, & if\ t_n = \infty, \\ +1, & if\ (x(t_n), y(t_n)) \notin L \end{cases}$$

1.9. Having defined the strategy sets for the players P and E and the payoff function, we may define the differential game as the game in normal form. We interpreted the normal form Γ as the triple $\Gamma =< X, Y, K >$, where $X \times Y$ is the space of pairs of all possible strategies in the game Γ, and K is the payoff function defined on $X \times Y$. In the case involved, the payoff function is defined not only on the set of pairs of all possible strategies in the game, but also on the set of all pairs of initial positions x_0, y_0. Therefore, for each pair $(x_0, y_0) \in R^n \times R^n$ there is the corresponding game in normal form, i.e. in fact some family of games in normal form that are dependent on parameters $(x_0, y_0) \in R^n \times R^n$ are defined.

Definition 1. *The normal form of the differential game $\Gamma(x_0, y_0)$ given on the space of strategy pairs $P \times E$ means the system*

$$\Gamma(x_0, y_0) = \langle x_0, y_0; P, E, K(x_0, y_0; u(\cdot), v(\cdot))\rangle,$$

where $K(x_0, y_0; u(\cdot), v(\cdot))$ is the payoff function defined by any one of the above methods.

Note that in this chapter maximizing player (1 Player) is player E, and K is the payoff of the player E. If the payoff function K (payoff of E, payoff of P being $(-K)$) in the game Γ is terminal, then the corresponding game Γ is called the game with terminal payoff. If the function K is defined by the second method, then we have the game for achievement of a minimum result. If the function K in the game Γ is integral, then the corresponding game Γ is called the game with integral payoff. When the payoff function in the game Γ is qualitative, then the corresponding game Γ is called the *game of kind.*

1.10. It appears natural that optimal strategies cannot exist in the class of piecewise open-loop strategies (in view of the open structure of the class). However, we can show that in a sufficiently large number of cases, for any $\epsilon > 0$ there is an ϵ-equilibrium point.

Recall the definition of the ϵ-equilibrium point.

Definition 2. *Let $\epsilon > 0$ be given. The situation $s_\epsilon = \{x_0, y_0; u_\epsilon(\cdot), v_\epsilon(\cdot)\}$ is called an ϵ-equilibrium in the game $\Gamma(x_0, y_0)$ if for all $u(\cdot) \in P$ and $v(\cdot) \in E$ there is*

$$K(x_0, y_0; u(\cdot), v_\epsilon(\cdot)) + \epsilon \geq K(x_0, y_0; u_\epsilon(\cdot), v_\epsilon(\cdot)) \qquad (1.10)$$
$$\geq K(x_0, y_0; u_\epsilon(\cdot), v(\cdot)) - \epsilon.$$

The strategies $u_\epsilon(\cdot), v_\epsilon(\cdot)$ determined in (1.10) are called *ϵ-optimal strategies* for players P and E respectively.

Lemma 1. *Suppose that in the game $\Gamma(x_0, y_0)$ for every $\epsilon > 0$ there is an ϵ-equilibrium. Then there exists a limit*

$$\lim_{\epsilon \to 0} K(x_0, y_0; u_\epsilon(\cdot), v_\epsilon(\cdot)).$$

Definition 3. *The function $V(x, y)$ defined at each point (x, y) of some set $D \subset R^n \times R^n$ by the rule*

$$\lim_{\epsilon \to 0} K(x, y; u_\epsilon(\cdot), v_\epsilon(\cdot)) = V(x, y), \qquad (1.11)$$

is called the value of the game $\Gamma(x, y)$ with initial conditions $(x, y) \in D$.

The existence of an ϵ-equilibrium in the game $\Gamma(x_0, y_0)$ for any $\epsilon > 0$ is equivalent to the fulfilment of the equality

$$\sup_{v(\cdot) \in E} \inf_{u(\cdot) \in P} K(x_0, y_0; u(\cdot), v(\cdot)) = \inf_{u(\cdot) \in P} \sup_{v(\cdot) \in E} K(x_0, y_0; u(\cdot), v(\cdot)).$$

If in the game $\Gamma(x_0, y_0)$ for any $\epsilon > 0$ there are ϵ-optimal strategies for players P and E, then the game $\Gamma(x_0, y_0)$ is said to *have a solution.*

Definition 4. *Let $u^*(\cdot), v^*(\cdot)$ be the pair of strategies such that*

$$K(x_0, y_0; u(\cdot), v^*(\cdot)) \geq K(x_0, y_0; u^*(\cdot), v^*(\cdot)) \geq K(x_0, y_0; u^*(\cdot), v(\cdot)) \qquad (1.12)$$

for all $u(\cdot) \in P$ and $v(\cdot) \in E$. The situation $s^* = (x_0, y_0; u^*(\cdot), v^*(\cdot))$ is then called an equilibrium in the game $\Gamma(x_0, y_0)$. The strategies $u^*(\cdot) \in P$ and $v^*(\cdot) \in E$ from (1.12) are called optimal strategies for players P and E, respectively.

The existence of an equilibrium in the game $\Gamma(x_0, y_0)$ is equivalent to the fulfilment of the equality

$$\max_{v(\cdot)\in E} \inf_{u(\cdot)\in P} K(x_0, y_0; u(\cdot), v(\cdot)) = \min_{u(\cdot)\in E} \sup_{v(\cdot)\in P} K(x_0, y_0; u(\cdot), v(\cdot)).$$

Clearly, if there exists an equilibrium, then for any $\epsilon > 0$ it is also an ϵ-equilibrium, i.e. here the function $V(x, y)$ merely coincides with $K(x, y; u^*(\cdot), v^*(\cdot))$.

1.11. We shall now consider a synthesizing strategies.

Definition 5. *The pair $(u^*(x, y, t), v^*(x, y, t))$ is called a synthesizing strategy equilibrium in the differential game, if the inequality*

$$K(x_0, y_0; u(x, y, t), v^*(x, y, t)) \geq K(x_0, y_0; u^*(x, y, t), v^*(x, y, t))$$
$$\geq K(x_0, y_0; u^*(x, y, t), v(x, y, t)) \quad (1.13)$$

holds for all situations $(u(x, y, t), v^*(x, y, t))$ and $(u^*(x, y, t), v(x, y, t))$ for which there exists a unique solution to system (1.7) that can be prolonged on $[0, \infty)$ from the initial states x_0, y_0. The strategies $u^*(x, y, t), v^*(x, y, t)$ are called optimal strategies for players P and E.

A distinction must be made between the notions of an equilibrium in piecewise open-loop and synthesizing strategies. Note that in the ordinary sense the equilibrium in the class of functions $u(x, y, t), v(x, y, t)$ cannot be defined because of the nonrectangularity of the space of situations, i.e. in the synthesizing strategies it is impossible to require that the inequality (1.13) holds for all strategies $u(x, y, t), v(x, y, t)$, since some pairs $(u^*(x, y, t), v(x, y, t))$, $(u(x, y, t), v^*(x, y, t))$ cannot be admissible (in the corresponding situation the system of equations (1.7) may have no solution or may have no unique solution).

In what follows we shall consider the classes of piecewise open-loop strategies, unless otherwise indicated. Preparatory to proving the existence of an ϵ-equilibrium in the differential game we will first consider one auxiliary class of multistage games with perfect information.

2 Multistage perfect-information games with an infinite number of alternatives

2.1. We shall consider the class of multistage games with perfect information that is a generalization of the games with perfect information from Sec. 5.1. The game proceeds in the n-dimensional Euclidean space R^n. Denote by $x \in R^n$ the position of Player 1, and by $y \in R^n$ the position of Player 2. Suppose that the sets U_x, V_y are defined for each $x \in R^n, y \in R^n$, respectively. These are taken to be the compact sets in the Euclidean space R^n. The game starts from a position x_0, y_0. At the 1st step the players 1 and 2 choose the points $x_1 \in U_{x_0}$ and $y_1 \in V_{y_0}$. In this case the choice by Player 2 is made known to Player 1 before he chooses the point $x_1 \in U_{x_0}$. At the points x_1, y_1 the players 1 and 2 choose $x_2 \in U_{x_1}$ and $y_2 \in V_{y_1}$ and Player 2's choice is made known to Player 1 before he chooses the point $x_2 \in U_{x_1}$ and so on. In positions x_{k-1}, y_{k-1} at the k-th step the players choose $x_k \in U_{x_{k-1}}$, $y_k \in V_{y_{k-1}}$ and Player 2's choice is made known to Player 1 before he chooses the point $x_k \in U_{x_{k-1}}$. This process terminates at the N-th step by choosing $x_N \in U_{x_{N-1}}$, $y_N \in V_{y_{N-1}}$ and passing to the state x_N, y_N.

The family of sets U_x, V_y, $x \in R^n$, $y \in R^n$ is taken to be continuous in x, y in Hausdorff metric. This means that for any $\epsilon > 0$ there is $\delta > 0$ such that for $|x - x_0| < \delta$ ($|y - y_0| < \delta$)

$$(U_{x_0})_\epsilon \supset U_x, \qquad (U_x)_\epsilon \supset U_{x_0};$$

$$(V_{y_0})_\epsilon \supset V_y, \qquad (V_y)_\epsilon \supset V_{y_0}.$$

Here U_ϵ (V_ϵ) is an ϵ-neighborhood of the set U (V).

The following result is well known in analysis (see [2]).

Lemma 2. *Let $f(x', y')$ be a continuous function on the Cartesian product $U_x \times V_y$. If the families $\{U_x\}, \{V_y\}$ are Hausdorff-continuous in x, y, then the functionals*

$$F_1(x, y) = \max_{y' \in V_y} \min_{x' \in U_x} f(x', y'),$$

$$F_2(x, y) = \min_{x' \in U_x} \max_{y' \in V_y} f(x', y')$$

are continuous in x, y.

Let $\overline{x} = (x_0, \ldots, x_N)$ and $\overline{y} = (y_0, \ldots, y_N)$ be the respective paths of players 1 and 2 realized during the game. The payoff to Player 2 is

$$\max_{0 \le k \le N} f(x_k, y_k) = F(\overline{x}, \overline{y}), \tag{2.1}$$

where $f(x, y)$ is a continuous function of x, y. The payoff to Player 1 is $-F$ (the game is zero-sum).

We assume that this game is a perfect-information game, i.e. at each moment (at each step) the players know the positions x_k, y_k and the time instant $k + 1$, moreover, Player 1 is informed about the choice y_{k+1} by Player 2.

Strategies for Player 1 are all possible functions $u(x, y, t)$ such that $u(x_{k-1}, y_k, k) \in U_{x_{k-1}}$. Strategies for Player 2 are all possible functions $v(x, y, t)$ such that $v(x_{k-1}, y_{k-1}, k) \in V_{y_{k-1}}$. These strategies are called pure strategies (as distinct from mixed strategies).

Suppose that players 1 and 2 use pure strategies $u(x, y, t), v(x, y, t)$. In situation $(u(\cdot), v(\cdot))$ the game proceeds as follows. At the first step Player 2 passes from the state y_0 to the state $y_1 = v(x_0, y_0, 1)$, while Player 1 passes from the state x_0 to the state $x_1 = u(x_0, y_1, 1) = u(x_0, v(x_0, y_0, 1), 1)$ (because Player 1 is informed about the choice by Player 2). At the 2nd step the players pass to the states $y_2 = v(x_1, y_1, 2)$, $x_2 = u(x_1, y_2, 2) = u(x_1, v(x_1, y_1, 2), 2)$ and so on. At the k-th step players 1 and 2 pass from the states x_{k-1}, y_{k-1} to the states

$y_k = v(x_{k-1}, y_{k-1}, k)$, $x_k = u(x_{k-1}, y_k, k) = u(x_{k-1}, v(x_{k-1}, y_{k-1}, k), k)$. Thus to each situation $(u(\cdot), v(\cdot))$ uniquely correspond the paths of the players 1 and 2, $\overline{x} = (x_0, x_1, \ldots, x_k, \ldots, x_N)$ and $\overline{y} = (y_0, y_1, \ldots, y_k, \ldots, y_N)$; hence the payoff $K(u(\cdot), v(\cdot)) = F(x, y)$ is determined by (2.1).

This game depends on two parameters; the initial positions x_0, y_0 and the duration N. For this reason, we denote the game by $\Gamma(x_0, y_0, N)$. For the purposes of further discussion it is convenient to assign each game: $\Gamma(x_0, y_0, N)$ to the family of games $\Gamma(x, y, T)$ depending on parameters x, y, T.

2.2.

Theorem 1. *The game $\Gamma(x_0, y_0, N)$ has an equilibrium in pure strategies and the value of the game $V(x_0, y_0, N)$ satisfies the following functional equation*

$$V(x_0, y_0, k) = \max\{f(x_0, y_0), \max_{y \in V_{y_0}} \min_{x \in U_{x_0}} V(x, y, k-1)\},\ k = 1, \ldots, N;$$

$$V(x, y, 0) = f(x, y). \tag{2.2}$$

Proof is carried out by induction for the number of steps. Let $N = 1$. Define the strategies $u^*(\cdot), v^*(\cdot)$, for the players in the game $\Gamma(x_0, y_0, 1)$ in the following way:

$$\min_{x \in U_{x_0}} f(x, y) = f(u^*(x_0, y, 1), y), \quad y \in V_{y_0}.$$

If $\max_{y \in V_{y_0}} \min_{x \in U_{x_0}} f(x, y) = f(u^*(x_0, y^*, 1), y^*)$ then $v^*(x_0, y_0, 1) = y^*$. Then

$$K(u^*(\cdot), v^*(\cdot)) = \max\{f(x_0, y_0), \max_{y \in V_{y_0}} \min_{x \in U_{x_0}} f(x, y)\}$$

and for any strategies $u(\cdot), v(\cdot)$ of the players in the game $\Gamma(x_0, y_0, 1)$

$$K(u^*(\cdot), v(\cdot)) \leq K(u^*(\cdot), v^*(\cdot)) \leq K(u(\cdot), v^*(\cdot)).$$

In view of this, the assertion of Theorem holds for $N \leq 1$.

We now assume that the assertion of Theorem holds for $N < n$. We shall prove it for $N = n + 1$, i.e. for the game $\Gamma(x_0, y_0, n + 1)$. Let us consider the family of games $\Gamma(x, y, n)$ $x \in U_{x_0}, y \in V_{y_0}$. Denote by $\overline{u}^n_{xy}(\cdot), \overline{v}^n_{xy}(\cdot)$ an equilibrium in the game $\Gamma(x, y, n)$. Then $K(\overline{u}^n_{xy}(\cdot), \overline{v}^n_{xy}(\cdot)) = V(x, y, n)$, where $V(x, y, n)$ is determined by relationships (2.2). Using the continuity of the function $f(x, y)$ and Lemma 2.1, we may prove the continuity of the function $V(x, y, n)$ in x, y. We define strategies $\overline{u}^{n+1}(\cdot), \overline{v}^{n+1}(\cdot)$ for players in the game $\Gamma(x_0, y_0, n + 1)$ as follows:

$$\min_{x \in U_{x_0}} V(x, y, n) = V(\overline{u}^{n+1}(x_0, y, 1), y, n),\ y \in V_{y_0}.$$

If $\max_{y \in V_{y_0}} \min_{x \in U_{x_0}} V(x, y, n) = V(\overline{u}^{n+1}(x_0, \tilde{y}, 1), \tilde{y}, n)$, then $\overline{v}^{n+1}(x_0, y_0, 1) = \tilde{y}$ (for $x \neq x_0$, $y \neq y_0$ the functions $\overline{v}^{n+1}(x, y, 1)$ and $\overline{u}^{n+1}(x, y, 1)$ can be defined in an arbitrary way)

$$\overline{u}^{n+1}(\cdot, k) = \overline{u}^n_{x_1 y_1}(\cdot, k-1),\ k = 2, \ldots, n+1,$$

$$\overline{v}^{n+1}(\cdot, k) = \overline{v}^{n}_{x_1 y_1}(\cdot, k-1),\ k = 2, \dots, n+1.$$

Here $x_1 \in U_{x_0}$, $y_1 \in V_{y_0}$ are the positions realized after the 1st step in the game $\Gamma(x_0, y_0, n+1)$. By construction,

$$K(\overline{u}^{n+1}(\cdot), \overline{v}^{n+1}(\cdot)) = \max\{f(x_0, y_0), \max_{y \in V_{y_0}} \min_{x \in U_{x_0}} V(x, y, n)\}. \tag{2.3}$$

Let us fix an arbitrary strategy $u(\cdot)$ for Player 1 in the game $\Gamma(x_0, y_0, n+1)$. Let $u(x_0, \tilde{y}, 1) = x_1$ where $\tilde{y} = \overline{v}^{n+1}(x_0, y_0, 1)$ and $u^n_{xy}(\cdot)$ is the truncation of strategy $u(\cdot)$ to the game $\Gamma(x, y, n)$, $x \in U_{x_0}$, $y \in V_{y_0}$.

The following relationships are valid:

$$K(\overline{u}^{n+1}(\cdot), \overline{v}^{n+1}(\cdot)) \le \max\{f(x_0, y_0), V(x_1, \tilde{y}, n)\}$$

$$= \max\{f(x_0, y_0), K(\overline{u}^{n}_{x_1\tilde{y}}(\cdot), \overline{v}^{n}_{x_1\tilde{y}}(\cdot)\}$$

$$\le \max\{f(x_0, y_0), K(u^{n}_{x_1\tilde{y}}(\cdot), \overline{v}^{n}_{x_1\tilde{y}}(\cdot))\} = K(u(\cdot), \overline{v}^{n+1}(\cdot)). \tag{2.4}$$

In the same manner the inequality

$$K(\overline{u}^{n+1}(\cdot), \overline{v}^{n+1}(\cdot)) \ge K(\overline{u}^{n+1}(\cdot), v(\cdot)) \tag{2.5}$$

can be proved for any strategy $v(\cdot)$ of Player 2 in the game: $\Gamma(x_0, y_0, n+1)$. From relationships (2.3)–(2.5) it follows that the assertion of Theorem holds for $N = n+1$. This completes the proof of the theorem by induction.

We shall now consider the game $\overline{\Gamma}(x_0, y_0, N)$ which differs from the game $\Gamma(x_0, y_0, N)$ in that the information is provided by Player 1 about his choice. Now, at the step k in the game: $\overline{\Gamma}(x_0, y_0, N)$ Player 2 knows not only the states x_{k-1}, y_{k-1} and the step k, but also the state $x_k \in U_{x_{k-1}}$ chosen by Player 1. We may show that the *game* $\overline{\Gamma}(x_0, y_0, N)$ *has an equilibrium in pure strategies and the game value* $\overline{V}(x_0, y_0, N)$ *satisfies the equation*

$$\overline{V}(x_0, y_0, k) = \max\{f(x_0, y_0), \min_{x \in U_{x_0}} \max_{y \in V_{y_0}} \overline{V}(x, y, k-1)\},$$

$$k = 1, \dots, N, \quad \overline{V}(x, y, 0) = f(x, y). \tag{2.6}$$

2.3. Let us consider the games $\Gamma'(x_0, y_0, N)$ and $\overline{\Gamma}'(x_0, y_0, N)$ that are distinguished from the games $\Gamma(x_0, y_0, N)$ and $\overline{\Gamma}(x_0, y_0, N)$ correspondingly in that the payoff function is equal to a distance between Player 1 and Player 2 at the final step of the game, i.e. $\rho(x_N, y_N)$. Instead of relations (2.2),(2.6), the following equations hold:

$$V'(x, y, k) = \max_{y' \in V_y} \min_{x' \in U_x} V'(x', y', k-1),\ k = 1, \dots, N,$$

$$V'(x, y, 0) = \rho(x, y); \tag{2.7}$$

$$\overline{V}'(x, y, k) = \min_{x' \in U_x} \max_{y' \in V_y} \overline{V}'(x', y', k-1),\ k = 1, \dots, N,$$

$$\overline{V}'(x, y, 0) = \rho(x, y). \tag{2.8}$$

Example 3. Let us consider a discrete game of pursuit, in which the sets U_x are the circles of radius α centered at the point x, while the sets V_y are the circles of radius β centered at the point y ($\alpha > \beta$). This corresponds to the game in which Player 2 (Evader) moves in a plane with the speed not exceeding β, while Player 1 (Pursuer) moves with a speed not exceeding α. The pursuer has a speed advantage and the second move is made by Player 1. The game of this type is called a discrete game of "simple pursuit" with discrimination against evader. The duration of the game is N steps and the payoff to Player 2 is equal to a distance between the players at the final step.

We shall find the value of the game and optimal strategies for the players by using the functional equation (2.7).

We have

$$V(x, y, 1) = \max_{y' \in V_y} \min_{x' \in U_x} \rho(x', y'). \tag{2.9}$$

Since U_x and V_y are the circles of radii α and β with centers at x and y, we have that if $U_x \supset V_y$, then $V(x, y, 1) = 0$, if, however, $U_x \not\supset V_y$, then $V(x, y, 1) = \rho(x, y) + \beta - \alpha = \rho(x, y) - (\alpha - \beta)$. Thus,

$$V(x, y, 1) = \begin{cases} 0, & if\ U_x \supset V_y, i.e.\ \rho(x, y) - (\alpha - \beta) \leq 0, \\ \rho(x, y) - (\alpha - \beta), & if\ U_x \not\supset V_y, \end{cases}$$

in other words,

$$V(x, y, 1) = \max[0, \rho(x, y) - (\alpha - \beta)]. \tag{2.10}$$

Using induction on the number of steps k we shall prove that the following formula holds:

$$V(x, y, k) = \max[0, \rho(x, y) - k(\alpha - \beta)],\ k \geq 2. \tag{2.11}$$

Suppose (2.11) holds for $k = m - 1$. Show that this formula holds for $k = m$. Using equation (2.7) and relationships (2.9), (2.10) we obtain

$$V(x, y, m) = \max_{y' \in V_y} \min_{x' \in U_x} V(x', y', m - 1)$$

$$= \max_{y' \in V_y} \min_{x' \in U_x} \{\max[0, \rho(x', y') - (m - 1)(\alpha - \beta)]\}$$

$$= \max[0, \max_{y' \in V_y} \min_{x' \in U_x} \{\rho(x', y')\} - (m - 1)(\alpha - \beta)]$$

$$= \max[0, \max\{0, \rho(x, y) - (\alpha - \beta)\} - (m - 1)(\alpha - \beta)] = \max[0, \rho(x, y) - m(\alpha - \beta)],$$

which is what we set out to prove.

If $V(x_0, y_0, m) = \rho(x_0, y_0) - m(\alpha - \beta)$, i.e. $\rho(x_0, y_0) - m(\alpha - \beta) > 0$, then the optimal strategy dictates Player 2 to choose at the k-th step of the game the point y_k of intersection the line of centers x_{k-1}, y_{k-1} with the boundary $V_{y_{k-1}}$ that is the farthest from x_{k-1}. Here x_{k-1}, y_{k-1} are the players positions after the $(k - 1)$ step, $k = 1, \ldots, N$. The optimal strategy for Player 1 dictates him to choose at the k-th step of the game the point from the set $U_{x_{k-1}}$ that is the nearest to the point y_k. If both players are acting optimally, then the sequence of the chosen points $x_0, x_1, \ldots, x_N$, $y_0, y_1, \ldots, y_N$ lies along the straight line passing through x_0, y_0. If $V(x_0, y_0, m) = 0$, then an optimal strategy for Player 2 is arbitrary, while an optimal strategy for Player 1 remains unaffected. In this case, after some step k the equality $\max_{y \in V_{y_k}} \min_{x \in U_{x_k}} \rho(x, y) = 0$ is satisfied; therefore starting with the $(k+1)$ step the choices by Player 1 will repeat the choices by Player 2.

3 Existence of ϵ-equilibria in differential games with prescribed duration

3.1. In this section we shall prove existence of piecewise open-loop strategy ϵ-equilibria in differential games with prescribed duration. Let us discuss the case where the payoff to Player E is the distance $\rho(x(T), y(T))$ at the last instant of the game T.

Let the dynamics of the game be given by the following differential equations:

$$\text{for } P: \ \dot{x} = f(x, u); \tag{3.1}$$

$$\text{for } E: \ \dot{y} = g(y, v). \tag{3.2}$$

Here $x, y \in R^n$, $u \in U$, $v \in V$, where U, V are compact sets in the Euclidean spaces R^k and R^l, respectively, $t \in [0, \infty)$.

Definition 6. *Denote by $C_P^t(x_0)$ the set of points $x \in R^n$ for which there is a measurable open-loop control $u(t) \in U$ sending the point x_0 to x in time t, i.e. $x(t_0) = x_0$, $x(t_0 + t) = x$. The set $C_P^t(x_0)$ is called a reachability set for Player P from initial state x_0 in time t.*

In this manner we may also define the reachability set $C_E^t(y_0)$ for Player E from the initial state y_0 in time t.

We assume that the functions f, g are such that the reachability sets $C_P^t(x_0)$, $C_E^t(y_0)$ for players P and E, respectively, satisfy the following conditions:

1. $C_P^t(x_0)$, $C_E^t(y_0)$ are defined for any $x_0, y_0 \in R^n$, $t_0, t \in [0, \infty)$ $(t_0 \leq t)$ and are compact sets of the space R^n;

2. the point to set map $C_P^t(x_0)$ is continuous in all its variables in Hausdorff metric, i.e. for every $\epsilon > 0$, $x_0' \in R^n$, $t \in [0, \infty)$ there is $\delta > 0$ such that if $|t - t'| < \delta$, $\rho(x_0, x_0') < \delta$, then $\rho^*(C_P^t(x_0), C_P^{t'}(x_0')) < \epsilon$. This also applies to $C_E^t(y_0)$.

Recall that the Hausdorff metric ρ^* in the space of compact sets R^n is given as follows:

$$\rho^*(A, B) \equiv \max(\rho'(A, B), \rho'(B, A)), \quad \rho'(A, B) = \max_{a \in A} \rho(a, B)$$

and $\rho(a, B) = \min_{b \in B} \rho(a, b)$, where ρ is a standard metric in R^n.

We shall prove the existence theorem for the game of pursuit $\Gamma(x_0, y_0, T)$ with prescribed duration, where $x_0, y_0 \in R^n$ are initial positions for players P and E respectively. T is the duration of the game. The game $\Gamma(x_0, y_0, T)$ proceeds as follows. The players P and E at the time $t_0 = 0$ start their motions from positions x_0, y_0 in accordance with the chosen piecewise open-loop strategies. The game ends at the time $t = T$ and Player E receives from Player P an amount $\rho(x(T), y(T))$. At each time instant $t \in [0, T]$ of the game $\Gamma(x_0, y_0, T)$ each player knows the instant of time t, his own position and the position of his opponent. Denote by $P(x_0, t_0, t)$ $(E(y_0, t_0, t))$ the set of trajectories of system (3.1), (3.2) emanating from the point $x_0(y_0)$ and defined on the interval $[t_0, t]$.

3.2. Let us fix some natural number $n \geq 1$. We set $\delta = T/2^n$ and introduce the games $\Gamma_i^\delta(x_0, y_0, T)$, $i = 1, 2, 3$, that are auxiliary with respect to the game $\Gamma(x_0, y_0, T)$.

The game $\Gamma_1^\delta(x_0, y_0, T)$ proceeds as follows. At the 1st step Player E in position y_0 chooses y_1 from the set $C_E^\delta(y_0)$. At this step Player P knows the choice of y_1 by Player E and, in position x_0, chooses the point $x_1 \in C_P^\delta(x_0)$. At the k-th step $k = 2, 3, \ldots, 2^n$, Player E knows Player P's position $x_{k-1} \in C_P^\delta(x_{k-2})$ and his own position $y_{k-1} \in C_E^\delta(y_{k-2})$ and chooses the point $y_k \in C_E^\delta(y_{k-1})$. Player P knows x_{k-1}, y_{k-1}, y_k and chooses $x_k \in C_P^\delta(x_{k-1})$. At the 2^n-th step the game ends and Player E receives an amount $\rho(x(T), y(T))$, where $x(T) = x^{2^n}$, $y(T) = y^{2^n}$.

Note that the players' choices at the k-th step of the points x_k, y_k from the rechability sets $C_P^\delta(x_{k-1})$, $C_E^\delta(y_{k-1})$ can be interpreted as their choices of the corresponding trajectories from the sets $P((x_{k-1}, k-1)\delta, k\delta)$, $E((y_{k-1}, k-1)\delta, k\delta)$, terminating in the points x_k, y_k at the time instant $t = k\delta$ (or the choice of controls $u(\cdot), v(\cdot)$ on $[(k-1)\delta, k\delta]$ to which these trajectories correspond according to (3.1), (3.2)).

The game $\Gamma_2^\delta(x_0, y_0, T)$ differs from the game $\Gamma_1^\delta(x_0, y_0, T)$ in that at the k-th step Player P chooses $x_k \in C_P^\delta(x_{k-1})$ with a knowledge of x_{k-1}, y_{k-1}, while Player E, with an additional knowledge of x_k, chooses $y_k \in C_E^\delta(y_{k-1})$.

The game $\Gamma_3^\delta(x_0, y_0, T)$ differs from the game $\Gamma_2^\delta(x_0, y_0, T)$ in that at the 2^n step Player P chooses $x_{2^n} \in C_P^\delta(x_{2^n-1})$, then the game ends and Player E receives an amount $\rho(x(T), y(T-\delta))$, where $x(T) = x_{2^n}$, $y(T-\delta) = y_{2^n-1}$.

3.3.

Lemma 3. *The games $\Gamma_i^\delta(x_0, y_0, T)$, $i = 1, 2, 3$, have an equilibrium for all $x_0, y_0, T < \infty$ and the value of the game $Val\Gamma_i^\delta(x_0, y_0, T)$ is a continuous function $x_0, y_0 \in R^n$. For any $n \geq 0$ there is*

$$Val\Gamma_1^\delta(x_0, y_0, T) \leq Val\Gamma_2^\delta(x_0, y_0, T), \; T = 2^n\delta. \tag{3.3}$$

Proof. The games $\Gamma_i^\delta(x_0, y_0, T)$, $i = 1, 2, 3$, belong to the class of multistage games. The existence of an equilibrium in the games $\Gamma_i^\delta(x_0, y_0, T)$ and the continuity of functions $Val\Gamma_i^\delta(x_0, y_0, T)$ in x_0, y_0 immediately follows from Theorem 1 in Module 4 and its corollary. The following recursion equations hold for the values of the games $\Gamma_i^\delta(x_0, y_0, T)$, $i = 1, 2$,

$$Val\Gamma_1^\delta(x_0, y_0, T) = \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} Val\Gamma_1^\delta(x, y, T - \delta),$$

$$Val\Gamma_2^\delta(x_0, y_0, T) = \min_{x \in C_P^\delta(x_0)} \max_{y \in C_E^\delta(y_0)} Val\Gamma_2^\delta(x, y, T - \delta),$$

with the initial condition $Val\Gamma_1^\delta(x, y, 0) = Val\Gamma_2^\delta(x, y, 0) = \rho(x, y)$. Sequential application of Lemma 1 in Module 1 shows the validity of inequality (3.3).

3.4.

Lemma 4. *For any integer $n \geq 0$ the following inequalities hold:*

$$Val\Gamma_1^{\delta_n}(x_0, y_0, T) \leq Val\Gamma_1^{\delta_{n+1}}(x_0, y_0, T),$$

$$Val\Gamma_2^{\delta_n}(x_0, y_0, T) \geq Val\Gamma_2^{\delta_{n+1}}(x_0, y_0, T),$$

where $\delta_k = T/2^k$

Proof. Show the validity of the first of these inequalities. The second inequality can be proved in a similar manner. To avoid cumbersome notation, let $C^k(y_i) = C_E^{\delta_k}(y_i)$, $C^k(x_i) = C_P^{\delta_k}(x_i)$, $i = 0, 1, \ldots, 2^n - 1$. We have

$$Val\Gamma_1^{\delta_{n+1}}(x_0, y_0, T)$$

$$= \max_{y_1 \in C^{n+1}(y_0)} \min_{x_1 \in C^{n+1}(x_0)} \max_{y_2 \in C^{n+1}(y_1)} \min_{x_2 \in C^{n+1}(x_1)} Val\Gamma_1^{\delta_{n+1}}(x_2, y_2, T - 2\delta_{n+1})$$

$$\geq \max_{y_1 \in C^{n+1}(y_0)} \max_{y_2 \in C^{n+1}(y_1)} \min_{x_1 \in C^{n+1}(x_0)} \min_{x_2 \in C^{n+1}(x_1)} Val\Gamma_1^{\delta_{n+1}}(x_2, y_2, T - 2\delta_{n+1})$$

$$= \max_{y_1 \in C^n(y_0)} \min_{x_1 \in C^n(x_0)} Val\Gamma_1^{\delta_{n+1}}(x_1, y_1, T - \delta_n).$$

Continuation of this process yields

$$Val\Gamma_1^{\delta_{n+1}}(x_0, y_0, T) \geq \max_{y_1 \in C^n(y_0)} \min_{x_1 \in C^n(x_0)} \ldots \max_{y_{2^n} \in C^n(y_{2^n} - 1)} \min_{x_{2^n} \in C^n(x_{2^n} - 1)} \rho(x_{2^n}, y_{2^n})$$

$$= Val\Gamma_1^{\delta_n}(x_0, y_0, T).$$

3.5.

Theorem 2. *For all $x_0, y_0 \in R^n$, $T < \infty$ there is the limit equality*

$$\lim_{n \to \infty} Val\Gamma_1^{\delta_n}(x_0, y_0, T) = \lim_{n \to \infty} Val\Gamma_2^{\delta_n}(x_0, y_0, T),$$

where $\delta_n = T/2^n$.

Proof. Let us fix some $n \geq 0$. Let $u(\cdot), v(\cdot)$ be a pair of strategies in the game $\Gamma_2^{\delta_n}(x_0, y_0, T)$. This pair remains the same in the game $\Gamma_3^{\delta_n}(x_0, y_0, T)$. Suppose that the sequence $x_0, x_1, \ldots, x_{2^n}$, $y_0, y_1, \ldots, y_{2^n}$ is realized in situation $u(\cdot), v(\cdot)$. Denote the payoff functions in the games $\Gamma_2^{\delta_n}(x_0, y_0, T)$, $\Gamma_3^{\delta_n}(x_0, y_0, T)$ by $K_2(u(\cdot), v(\cdot)) = \rho(x_{2^n}, y_{2^n})$, $K_3(u(\cdot), v(\cdot)) = \rho(x_{2^n}, y_{2^n-1})$, respectively. Then

$$K_2(u(\cdot), v(\cdot)) \leq K_3(u(\cdot), v(\cdot)) + \rho(y_{2^n-1}, y_{2^n}).$$

Hence, by the arbitrariness of $u(\cdot), v(\cdot)$, we have:

$$Val\Gamma_2^{\delta_n}(x_0, y_0, T) \leq Val\Gamma_3^{\delta_n}(x_0, y_0, T) + \max_{y \in C_E^{T-\delta_n}(y_0)} \max_{y' \in C_E^{\delta_n}(y)} \rho(y, y'). \quad (3.4)$$

Let $y_1^{\delta_n} \in C_E^{\delta_n}(y_0)$, then $C_E^{T-\delta_n}(y_1^{\delta_n}) \subset C_E^T(y_0)$. We now write inequality (3.4) for the games with the initial state $x_0, y_1^{\delta_n}$. In view of the previous inclusion, we have

$$Val\Gamma_2^{\delta_n}(x_0, y_1^{\delta_n}, T) \leq Val\Gamma_3^{\delta_n}(x_0, y_1^{\delta_n}, T) + \max_{y \in C_E^T(y_0)} \max_{y' \in C_E^{\delta}(y)} \rho(y, y'). \quad (3.5)$$

From the definition of the games $\Gamma_1^{\delta_n}(x_0, y_0, T)$ and $\Gamma_3^{\delta_n}(x_0, y_0, T)$ follows the equality

$$Val\Gamma_1^{\delta_n}(x_0, y_0, T) = \max_{y_1^{\delta_n} \in C_E^{\delta_n}(y_0)} Val\Gamma_3^{\delta_n}(x_0, y_1^{\delta_n}, T). \quad (3.6)$$

Since the function $C_E^t(y)$ is continuous in t and the condition $C_E^0(y) = y$ is satisfied, the second term in (3.5) tends to zero as $n \to \infty$. Denote it by $\epsilon_1(n)$. From (3.5), (3.6) we obtain

$$Val\Gamma_1^{\delta_n}(x_0, y_0, T) \geq Val\Gamma_2^{\delta_n}(x_0, y_1^{\delta_n}, T) - \epsilon_1(n). \quad (3.7)$$

By the continuity of the function $Val\Gamma_2^{\delta_n}(x_0, y_0, T)$, from (3.7) we obtain

$$Val\Gamma_1^{\delta_n}(x_0, y_0, T) \geq Val\Gamma_2^{\delta_n}(x_0, y_0, T) - \epsilon_1(n) - \epsilon_2(n), \quad (3.8)$$

where $\epsilon_2(n) \to 0$ as $n \to \infty$. Passing in (3.8) to the limit as $n \to \infty$ (which is possible in terms of the Lemmas in 3.3, 3.4 and the limit existence theorem for a monotone bounded sequence) we obtain

$$\lim_{n \to \infty} Val\Gamma_1^{\delta_n}(x_0, y_0, T) \geq \lim_{n \to \infty} Val\Gamma_2^{\delta_n}(x_0, y_0, T). \quad (3.9)$$

From Lemma 3.3 the inverse inequality follows. Hence both limits in (3.9) coincide.

3.6. The statement of Theorem 3.5 is proved on the assumption that the partition sequence of the interval $[0, T]$

$$\sigma_n = \{t_0 = 0 < t_1 < \ldots < t_N = T\}, \ n = 1, \ldots,$$

satisfies the condition $t_{j+1} - t_j = T/2^n$, $j = 0, 1, \ldots, 2^n - 1$. The statements of Theorem 3.5 and Lemmas 3.3, 3.4 hold for any sequence σ_n of refined partitions of the interval $[0, T]$, i.e. such that $\sigma_{n+1} \supset \sigma_n$ (this means that the partition σ_{n+1} is obtained from σ_n by adding new points)

$$\gamma(\sigma_n) = \max_i (t_{i+1} - t_i) \to_{n \to +\infty} 0.$$

We shall now consider any such partition sequences of the interval $[0, T]$ $\{\sigma_n\}$ and $\{\sigma'_n\}$.

Lemma 5. *The following equality holds:*

$$\lim_{n\to\infty} Val\Gamma_1^{\sigma_n}(x_0, y_0, T) = \lim_{n\to\infty} Val\Gamma_1^{\sigma'_n}(x_0, y_0, T),$$

where $x_0, y_0 \in R^n$. $T < \infty$.

Proof is carried out by reductio ad absurdum. Suppose the statement of this lemma is not true. For definiteness assume that the following inequality is satisfied:

$$\lim_{n\to\infty} Val\Gamma_1^{\sigma_n}(x_0, y_0, T) > \lim_{n\to\infty} Val\Gamma_1^{\sigma'_n}(x_0, y_0, T).$$

Then, by Theorem 3.5, we have

$$\lim_{n\to\infty} Val\Gamma_1^{\sigma_n}(x_0, y_0, T) > \lim_{n\to\infty} Val\Gamma_2^{\sigma'_n}(x_0, y_0, T).$$

Hence we may find natural numbers m_1, n_1 such that the following inequality is satisfied:

$$Val\Gamma_1^{\sigma_{m_1}}(x_0, y_0, T) > Val\Gamma_2^{\sigma'_{n_1}}(x_0, y_0, T).$$

Denote by $\overline{\sigma}$ the partition of the interval $[0, T]$ by the points belonging to both the partitions σ_{m_1} and σ'_{n_1}. For this partition

$$Val\Gamma_2^{\overline{\sigma}}(x_0, y_0, T) \le Val\Gamma_2^{\sigma'_{n_1}}(x_0, y_0, T) < Val\Gamma_1^{\sigma_{m_1}}(x_0, y_0, T) \le Val\Gamma_1^{\overline{\sigma}}(x_0, y_0, T),$$

whence

$$Val\Gamma_2^{\overline{\sigma}}(x_0, y_0, T) < Val\Gamma_1^{\overline{\sigma}}(x_0, y_0, T).$$

This contradicts (3.3), hence the above assumption is not true and the statement of the lemma is true.

3.7

Theorem 3. *For all* $x_0, y_0, T < \infty$ *in the game* $\Gamma(x_0, y_0, T)$ *there exists an* ϵ*-equilibium for any* $\epsilon > 0$. *In this case*

$$Val\Gamma(x_0, y_0, T) = \lim_{n\to\infty} Val\Gamma_1^{\sigma_n}(x_0, y_0, T), \tag{3.10}$$

where $\{\sigma_n\}$ *is any sequence of refined partitions of the interval* $[0, T]$.

Proof. Let us specify an arbitrarily chosen number $\epsilon > 0$ and show that for the players P and E there are respective strategies $u_\epsilon(\cdot)$ and $v_\epsilon(\cdot)$ such that for all strategies $u(\cdot) \in P$ and $v(\cdot) \in E$ the following inequalities hold:

$$K(x_0, y_0, u_\epsilon(\cdot), v(\cdot)) - \epsilon \leq K(x_0, y_0, u_\epsilon(\cdot), v_\epsilon(\cdot)) \leq K(x_0, y_0, u(\cdot), v_\epsilon(\cdot)) + \epsilon. \tag{3.11}$$

By Theorem 3.5, there is a partition σ of the interval $[0, T]$ such that

$$Val\Gamma_2^\sigma(x_0, y_0, T) - \lim_{n\to\infty} Val\Gamma_2^{\sigma_n}(x_0, y_0, T) < \epsilon/2,$$

$$\lim_{n\to+\infty} Val\Gamma_1^{\sigma_n}(x_0, y_0, T) - Val\Gamma_1^\sigma(x_0, y_0, T) < \epsilon/2.$$

Let $u^\epsilon(\cdot) = (\sigma, a_{u^\epsilon})$, $v^\epsilon(\cdot) = (\sigma, b_{v^\epsilon})$, where a_{u^ϵ}, b_{v^ϵ} are the optimal strategies for the players P and E, respectively, in the games $\Gamma_2^\sigma(x_0, y_0, T)$ and $\Gamma_1^\sigma(x_0, y_0, T)$.

Then the following relationships are valid:

$$K(x_0, y_0, u^\epsilon(\cdot), v(\cdot)) \leq Val\Gamma_2^\sigma(x_0, y_0, T)$$

$$< \lim_{n\to\infty} Val\Gamma_2^{\sigma_n}(x_0, y_0, T) + \frac{\epsilon}{2}, \quad v(\cdot) \in E, \tag{3.12}$$

$$K(x_0, y_0, u(\cdot), v^\epsilon(\cdot)) \geq Val\Gamma_1^\sigma(x_0, y_0, T)$$

$$> \lim_{n\to\infty} Val\Gamma_1^{\sigma_n}(x_0, y_0, T) - \frac{\epsilon}{2}, \quad u(\cdot) \in P. \tag{3.13}$$

From (3.12), (3.13) and Theorem 3.5 we have

$$-\frac{\epsilon}{2} < K(x_0, y_0, u^\epsilon(\cdot), v^\epsilon(\cdot)) - \lim_{n\to\infty} Val\Gamma_1^{\sigma_n}(x_0, y_0, T) < \frac{\epsilon}{2}. \tag{3.14}$$

From relationships (3.12)–(3.14) follows (3.11). By the arbitrariness of ϵ from (3.14) follows (3.10). This completes the proof of theorem.

3.8. *Remark.* The specific type of the payoff was not used in the proof of the existence theorem. Only continuous dependence of the payoff on the realized trajectories is essential. Therefore, Theorem 3.7 holds if any continuous functional of the trajectories $x(t), y(t)$ is considered in place of $\rho(x(T), y(T))$. In particular, such a functional can be $\min_{0\leq t<T} \rho(x(t), y(t))$, i.e. a minimum distance between players during the game. Thus the result of this section also holds for the differential minimum result game of pursuit with prescribed duration.

4 Differential time-optimal games of pursuit

4.1. Differential *time-optimal games of pursuit* are special case of differential games with integral payoff. The classes of strategies P and E are the same as in the game with prescribed duration. We assume that the set $F = \{(x, y) : \rho(x, y) \leq l,\ l > 0\}$ is given in $R^n \times R^n$ and $x(t), y(t)$ are trajectories for the players P and E in situation $(u(\cdot), v(\cdot))$ from initial conditions x_0, y_0. Denote

$$t_n(x_0, y_0; u(\cdot), v(\cdot)) = \min\{t : (x(t), y(t)) \in F\}. \tag{4.1}$$

If there is no t such that $(x(t), y(t)) \in F$, then $t_n(x_0, y_0; u(\cdot), v(\cdot))$ is $+\infty$. In the differential time-optimal game of pursuit the payoff to Player E is

$$K(x_0, y_0; u(\cdot), v(\cdot)) = t_n(x_0, y_0; u(\cdot), v(\cdot)). \tag{4.2}$$

The game depends on the initial conditions x_0, y_0, therefore it is denoted by $\Gamma(x_0, y_0)$.

From the definition of the payoff function (4.2) it follows that the objective of Player E in the game $\Gamma(x_0, y_0)$ is to maximize the time of approaching Player P within a given distance $l \geq 0$. Conversely, Player P wishes to minimize this time.

4.2. There is a close relation between the time-optimal game of pursuit $\Gamma(x_0, y_0, T)$ and the minimum result game of pursuit with prescibed duration. Let $\Gamma(x_0, y_0, T)$ be the game of pursuit with prescribed duration T for achievement of a minimum result (the payoff to Player E is $\min_{0 \leq t < T} \rho(x(t), y(t)))$. It was shown that for the games of this type there is an ϵ-equilibrium in the class of piecewise open-loop strategies for any $\epsilon > 0$ (see 3.8). Let $V(x_0, y_0, T)$ be the value of the game $\Gamma(x_0, y_0, T)$ and $V(x_0, y_0)$ be the value of the game $\Gamma(x_0, y_0)$ if it exists.

Lemma 6. *With x_0, y_0 fixed, the function $V(x_0, y_0, T)$ is continuous and does not increase in T on the interval* $[0, \infty]$.

Proof. Let $T_1 > T_2 > 0$. Denote by $v_\epsilon^{T_1}$ a strategy for Player E in the game $\Gamma(x_0, y_0, T)$ which guarantees that a distance between Player E and Player P on the interval $[0, T_1]$ will be at least $\max[0, V(x_0, y_0, T_1) - \epsilon]$. Hence it does ensure a distance $\max[0, V(x_0, y_0, T_1) - \epsilon]$ between the players on the interval $[0, T_2]$, where $T_2 < T_1$. Therefore

$$V(x_0, y_0, T_2) \geq \max[0, V(x_0, y_0, T_1) - \epsilon] \tag{4.3}$$

(the strategy ϵ-optimal in the game $\Gamma(x_0, y_0, T_1)$ is not necessarily ϵ-optimal in the game $\Gamma(x_0, y_0, T_2)$). Since ϵ can be chosen to be arbitrary, the statement of this Lemma follows from (4.3). The continuity of $V(x_0, y_0, T)$ in T will be left without proof. To be noted only is that this property can be obtained by using the continuity of $V(x_0, y_0, T)$ in x_0, y_0.

4.3. Let us consider the equation

$$V(x_0, y_0, T) = l \tag{4.4}$$

with respect to T. Three cases are possible here:

1) equation (4.4) has no roots;

2) it has a single root;

3) it has more than one root.

In case 3), the monotonicity and the continuity of the function $V(x_0, y_0, T)$ in T imply that the equation (4.4) has the whole segment of roots, the function $V(x_0, y_0, T)$, as a function of T, has a constancy interval. Let us consider each case individually.

Case 1. In this case the following is possible: a) $V(x_0, y_0, T) < l$ for all $T \geq 0$; b) $\inf_{T\geq 0} V(x_0, y_0, T) > l$; c) $\inf_{T\geq 0} V(x_0, y, T) = l$.

In case a) we have

$$V(x_0, y_0, 0) = \rho(x_0, y_0) < l,$$

i.e. $t_n(x_0, y_0; u(\cdot), v(\cdot)) = 0$ for all $u(\cdot), v(\cdot)$. The value of the game $\Gamma(x_0, y_0)$ is then $V(x_0, y_0) = 0$.

In case b) the following equality holds:

$$\inf_{T\geq 0} V(x_0, y_0, T) = \lim_{T\to\infty} V(x_0, y_0, T) > l.$$

Hence for any $T > 0$ (arbitrary large) Player E has a suitable strategy $v^T(\cdot) \in E$ which guarantees him l capture avoidance on the interval $[0, T]$. But Player P then has no strategy which could guarantee him l-capture of Player E in finite time. However, we cannot claim that Player E has a strategy which ensures l-capture avoidance in finite time. The problem of finding initial states in which such a strategy exists reduces to solving the game of kind for player E. Thus, for $l < \lim_{T\to\infty} V(x_0, y_0, T)$ it can be merely stated that the value of the game $\Gamma(x_0, y_0)$, if any, is larger than any previously given T, i.e. it is $+\infty$.

c) is considered together with case 3).

Case 2. Let T_0 be a single root of equation (4.4). Then it follows from the monotonicity and the continuity of the function $V(x_0, y_0, T)$ in T that

$$V(x_0, y_0, T) > V(x_0, y_0, T_0) \text{ for all } T < T_0,$$

$$V(x_0, y_0, T) < V(x_0, y_0, T_0) \text{ for all } T > T_0, \tag{4.5}$$

$$\lim_{T\to T_0} V(x_0, y_0, T) = V(x_0, y_0, T_0). \tag{4.6}$$

Let us fix an arbitrary $T > T_0$ and consider the game of pursuit $\Gamma(x_0, y_0, T)$. The game has an ϵ-equilibrium in the class of piecewise open-loop strategies for any $\epsilon > 0$. This, in particular, means that for any $\epsilon > 0$ there is Player P's strategy $u_\epsilon(\cdot) \in P$ which ensures the capture of Player E within a distance $V(x_0, y_0, T) + \epsilon$, i.e.

$$K(u_\epsilon(\cdot), v(\cdot)) \leq V(x_0, y_0, T) + \epsilon, \quad v(\cdot) \in E, \tag{4.7}$$

where $K(u(\cdot), v(\cdot))$ is the payoff function in the game $\Gamma(x_0, y_0, T)$. Then (4.5), (4.6) imply the existence of $\bar{\epsilon} > 0$ such that for any $\epsilon < \bar{\epsilon}$ there is a number $\tilde{T}(\epsilon), T_0 < \tilde{T}(\epsilon) \leq T$ for which

$$\epsilon = V(x_0, y_0, T_0) - V(x_0, y_0, \tilde{T}(\epsilon)). \tag{4.8}$$

From (4.7), (4.8) it follows that for any $\epsilon < \bar{\epsilon}$

$$K(u_\epsilon(\cdot), v(\cdot)) \leq V(x_0, y_0, T) + \epsilon \leq V(x_0, y_0, \tilde{T}(\epsilon)) + \epsilon = V(x_0, y_0, T_0) = l, \; v(\cdot) \in E,$$

i.e. the strategy $u_\epsilon(\cdot)$ ensures l-capture in time T. Hence, by the arbitrariness of $T > T_0$, it follows that for any $T > T_0$ there is a corresponding strategy $u^T(\cdot) \in P$ which ensures l-capture in time T. In other words, for $\delta > 0$ there is $u_\delta(\cdot) \in P$ such that

$$t_n(x_0, y_0; u_\delta(\cdot), v(\cdot)) \leq T_0 + \delta \text{ for all } v(\cdot) \in E. \tag{4.9}$$

In a similar manner we may prove the existence of $v_\delta(\cdot) \in E$ such that

$$t_n(x_0, y_0; u(\cdot), v_\delta(\cdot)) \geq T_0 - \delta \text{ for all } u(\cdot) \in P. \tag{4.10}$$

It follows from (4.9), (4.10) that in the time-optimal game of pursuit $\Gamma(x_0, y_0)$ for any $\epsilon > 0$ there is an ϵ-equilibrium in piecewise open-loop strategies and the value of the game is equal to T_0, with T_0 as a single root of equation (4.4).

Case 3. Denote by T_0 the minimal root of equation (4.4). Generally speaking, we cannot now state that the value of the game $Val\Gamma(x_0, y_0) = T_0$. Indeed, $V(x_0, y_0, T_0) = l$ merely implies that in the game $\Gamma(x_0, y_0, T_0)$ for any $\epsilon > 0$ Player P has a strategy $u_\epsilon(\cdot)$ which ensures for him, in time T_0, the capture of Player E within a distance of at most $l + \epsilon$. From the existence of more than one root of equation (4.4), and from the monotonicity of $V(x_0, y_0, T)$ in T we obtain the existence of the interval of constancy of the function $V(x_0, y_0, T)$ in $T \in [T_0, T_1]$. Therefore, an increase in the duration of the game $\Gamma(x_0, y_0, T_0)$ by δ, where $\delta < T_1 - T_0$, does not involve a decrease in the guaranteed approach to Player E, i.e. for all $T \in [T_0, T_1]$ Player P can merely ensure approaching Player E within a distance $l + \epsilon$ (for any $\epsilon > 0$), and it is beyond reason to hope for this quantity to become zero for some $T \in [T_0, T_1]$. If the game $\Gamma(x_0, y_0, T_0)$ had an equilibrium (but not an ϵ-equilibrium), then the value of the game $\Gamma(x_0, y_0)$ would also be equal to T_0 in Case 3.

4.4. Let us modify the notion of an equilibrium in the game $\Gamma(x_0, y_0)$. Further, in this section it may be convenient to use the notation $\Gamma(x_0, y_0, l)$ instead of $\Gamma(x_0, y_0)$ emphasizing the fact that the game $\Gamma(x_0, y_0, l)$ terminates when the players come within a distance l of each other.

Let $t_n^l(x_0, y_0; u(\cdot), v(\cdot))$ be the time until coming within a distance l in situation $(u(\cdot), v(\cdot))$ and let there be $\epsilon \geq 0$, $\delta \geq 0$.

Definition 7. *We say that the pair of strategies $\overline{u}_\epsilon^\delta(\cdot), \overline{v}_\epsilon^\delta(\cdot)$ constitutes an ϵ, δ-equilibrium in the game $\Gamma(x_0, y_0, l)$ if*

$$t_n^{l+\delta}(x_0, y_0; u(\cdot), \overline{v}_\epsilon^\delta(\cdot)) + \epsilon \geq t_n^{l+\delta}(x_0, y_0; \overline{u}_\epsilon^\delta(\cdot), \overline{v}_\epsilon^\delta(\cdot)) \geq t_n^{l+\delta}(x_0, y_0; \overline{u}_\epsilon^\delta(\cdot), v(\cdot)) - \epsilon$$

for all strategies $u(\cdot) \in P$, $v(\cdot) \in E$.

Definition 8. *Let there be a sequence $\{\delta_k\}$, $\delta_k \geq 0$, $\delta_k \to 0$ such that in all of the games $\Gamma(x_0, y_0, l + \delta_k)$ for every $\epsilon > 0$ there is an ϵ-equilibrium. Then the limit*

$$\lim_{k\to\infty} V(x_0, y_0, l + \delta_k) = V'(x_0, y_0, l)$$

is called the value of the game $\Gamma(x_0, y_0, l)$ in the generalized sense.

Note that the quantity $V'(x_0, y_0, l)$ does not depend on the choice of a sequence $\{\delta_k\}$ because of the monotone decrease of the function $V(x_0, y_0, l)$ in l.

Definition 9. *We say that the game* $\Gamma(x_0, y_0, l)$ *has the value in the generalized sense if there exists a sequence* $\{\delta_k\}$, $\delta_k \to 0$ *such that for every* $\epsilon > 0$ *and* $\delta_k \in \{\delta_k\}$ *in the game* $\Gamma(x_0, y_0, l)$ *there exists an* ϵ, δ_k*-equilibrium.*

It can be shown that if the game $\Gamma(x_0, y_0, l)$ has the value in the ordinary sense, then its value $V'(x_0, y_0, l)$ (in the generalized sense) exists and is

$$\lim_{\epsilon \to 0, \delta_k \to 0} t_n^{l+\delta_k}(x_0, y_0; \overline{u}_\epsilon^\delta(\cdot), \overline{v}_\epsilon^\delta(\cdot)) = V'(x_0, y_0, l).$$

From the definition of the value and solution of the game $\Gamma(x_0, y_0, l)$ (in the generalized sense) it follows that if in the game $\Gamma(x_0, y_0, l)$ for every $\epsilon > 0$ there is an ϵ-equilibrium in the ordinary sense (i.e. the solution in the ordinary sense), then $V(x_0, y_0, l) = V'(x_0, y_0, l)$ (it suffices to take a sequence $\delta_k \equiv 0$ for all k).

Theorem 4. *Let equation (4.4) have more than one root and let* T_0 *be the least root,* $T_0 < \infty$. *Then there exists the value* $V'(x_0, y_0, l)$ *(in the generalized sense) of the time-optimal game of pursuit* $\Gamma(x_0, y_0, l)$ *and* $V'(x_0, y_0, l) = T_0$.

Proof. The monotonicity and continuity of the function $V(x_0, y_0, T)$ in T imply the existence of a sequence $T_k \to T_0$ on the left such that $V(x_0, y_0, T_k) \to V(x_0, y_0, T_0) = l$ and the function $V(x_0, y_0, T_k)$ is strictly monotone in the points T_k. Let

$$\delta_k = V(x_0, y_0, T_k) - l \geq 0.$$

The strict monotonicity of the function $V(x_0, y_0, T)$ in the points T_k implies that the equation $V(x_0, y_0, T) = l + \delta_k$ has a single root T_k. This means that for every $\delta_k \in \{\delta_k\}$ in the games $\Gamma(x_0, y_0, l + \delta_k)$ there is an ϵ-equilibrium for every $\epsilon > 0$ (see Case 2 in 4.3). The game $\Gamma(x_0, y_0, l)$ then has a solution in the generalized sense:

$$\lim_{k \to \infty} V(x_0, y_0, l + \delta_k) = \lim_{k \to \infty} T_k = T_0 = V'(x_0, y_0, l).$$

This completes the proof of the theorem.

We shall now consider Case 1c in 4.3. We have $\inf_T V(x_0, y_0, T) = l$. Let $T_k \to \infty$. Then $\lim_{k \to \infty} V(x_0, y_0, T_K) = l$. From the monotonicity and continuity of $V(x_0, y_0, T)$ in T it follows that the sequence $\{T_k\}$ can be chosen to be such that the function $V(x_0, y_0, T)$ is strictly monotone in the points T_k. Then, as in the proof of Theorem in 4.4, it can be shown that there exists a sequence $\{\delta_k\}$ such that

$$\lim_{k \to \infty} V(x_0, y_0, l + \delta_k) = \lim_{k \to \infty} T_k = T_0 = \infty.$$

Thus, in this case there also exists a generalized solution, while the generalized value of the game $\Gamma(x_0, y_0, l)$ is infinity.

4.5. It is often important to find out whether Player P can guarantee l-capture from the given initial positions x, y in finite time T. If it is impossible, then we have to find out whether Player E can guarantee l-capture avoidance within a specified period of time.

Let $V(x, y, T)$ be the value of the game with prescribed duration T from initial states $x, y \in R^n$ with the payoff $\min_{0 \le t \le T} \rho(x(t), y(t))$. Then the following alternatives are possible: 1) $V(x, y, T) > l$; 2) $V(x, y, T) \le l$.

Case 1. From the definition of the function $V(x, y, T)$ it follows that for every $\epsilon > 0$ there is a strategy for Player E such that for all strategies $u(\cdot)$

$$K(x, y; u(\cdot), v_\epsilon^*(\cdot)) \ge V(x, y, T) - \epsilon.$$

Having chosen the ϵ to be sufficiently small we may ensure that

$$K(x, y; u(\cdot), v_\epsilon^*(\cdot)) \ge V(x, y, T) - \epsilon > l$$

holds for all strategies $u(\cdot) \in P$ of Player P. From the form of the payoff function K it follows that, by employing a strategy $v_\epsilon^*(\cdot)$, Player E can ensure that the inequality $\min_{0 \le t \le T_0} \rho(x(t), y(t)) > l$ would be satisfied no matter what Player P does. That is, in this case Player E ensures l-capture avoidance on the interval $[0, T]$ no matter what Player P does.

Case 2. Let T_0 be a minimal root of the equation $V(x, y, T) = l$ with x, y fixed (if $\rho(x, y) < l$, then T_0 is taken to be 0). From the definition of $V(x, y, T_0)$ it then follows that in the game $\Gamma(x, y, T_0)$ for every $\epsilon > 0$ Player P has a strategy u_ϵ^* which ensures that

$$K(x, y; u_\epsilon^*(\cdot), v(\cdot)) \le V(x, y; T_0) + \epsilon = l + \epsilon$$

for all strategies $v(\cdot) \in E$ of Player E. From the form of the payoff function K it follows that, by employing a strategy $u_\epsilon^*(\cdot)$, Player P can ensure that the inequality $\min_{0 \le t \le T} \rho(x(t), y(t)) \le l + \epsilon$ would be satisfied no matter what Player E does. Extending arbitrarily the strategy $u_\epsilon^*(\cdot)$ to the interval $[T_0, T]$ we have that, in Case 2, for every $\epsilon > 0$ Player P can ensure $(l + \epsilon)$-capture of Player E in time T no matter what the latter does.

This in fact proves the following theorem (of alternative).

Theorem 5. *For every $x, y \in R^n$, $T > 0$ one of the following assertions holds:*

1. *from initial conditions x, y Player E can ensure l-capture avoidance during the time T no matter what Player P does;*
2. *for any $\epsilon > 0$ Player P can ensure $(l + \epsilon)$-capture of Player E from initial states x, y during the time T no matter what Player E does.*

4.6. For each fixed $T > 0$ the entire space $R^n \times R^n$ is divided into three nonoverlapping regions: region $A = \{x, y : V(x, y, T) < l\}$ which is called the capture zone; region $B = \{x, y : V(x, y, T) > l\}$ which is naturally called the escape zone; region $C = \{x, y : V(x, y, T) = l\}$ is called the indifference zone.

Let $x, y \in A$. By the definition of A, for any $\epsilon > 0$ Player P has a strategy $u^*_\epsilon(\cdot)$ such that

$$K(x, y; u^*_\epsilon(\cdot), v(\cdot)) \leq V(x, y, T) + \epsilon$$

for all strategies $v(\cdot)$ of Player E. By a proper choice of $\epsilon > 0$ it is possible to ensure that the following inequality be satisfied:

$$K(x, y; u^*_\epsilon(\cdot), v(\cdot)) \leq V(x, y, T) + \epsilon < l.$$

This means that the strategy u^*_ϵ of Player P guarantees him l-capture of Player E from initial states during the time T. We thus obtain the following refinement of Theorem 4.5.

Theorem 6. *For every fixed $T > 0$ the entire space is divided into three nonoverlapping regions A, B, C possessing the following properties:*

1. *for any $x, y \in A$ Player P has a strategy $u^*_\epsilon(\cdot)$ which ensures l-capture of Player E on the interval $[0, T]$ no matter what the latter does;*
2. *for $x, y \in B$ Player E has a strategy $v^*_\epsilon(\cdot)$ which ensures l-capture avoidance of Player P on the interval $[0, T]$ no matter what the latter does;*
3. *if $x, y \in C$ and $\epsilon > 0$, then Player P has a strategy $u^*_\epsilon(\cdot)$ which ensures $(l + \epsilon)$ capture of Player E during the time T no matter what the latter does.*

5 Necessary and sufficient condition for existence of optimal open-loop strategy for Evader

5.1. An important subclass of games of pursuit is represented by the games in which an optimal strategy for evader is a function of time only (this is what is called a regular case).

We shall restrict consideration to the games of pursuit with prescribed duration, although all of the results below can be extended to the time-optimal games of pursuit. Let $C^T_P(x)(C^T_E(y))$ be a reachability set for Player $P(E)$ from initial state $x(y)$ by the time T, i.e. the set of those positions at which Player $P(E)$ can arrive from the initial state $x(y)$ at the time T by employing all possible measurable open-loop controls $u(t), (v(t))$, $t \in [0, T]$ provided the motion occurs in terms of the system $\dot{x} = f(x, u)$ $(\dot{y} = g(y, v))$. Let us introduce the quantity

$$\hat{\rho}_T(x_0, y_0) = \max_{y \in C^T_E(y_0)} \min_{x \in C^T_P(x_0)} \rho(x, y), \tag{5.1}$$

which may at times also be called a *hypothetical mismatch* of the sets $C^T_E(y_0)$ and $C^T_P(x_0)$.

The function $\hat{\rho}_T(x_0, y_0)$ has the following properties:

1^0. $\hat{\rho}_T(x_0, y_0) \geq 0, \quad \hat{\rho}_T(x_0, y_0)|_{T=0} = \rho(x_0, y_0)$;
2^0. $\hat{\rho}_T(x_0, y_0) = 0$ if $C_P^T(x_0) \supset C_E^T(y_0)$;
3^0. If $V(x_0, y_0, T)$ is the value of the game $\Gamma(x_0, y_0, T)$ with prescribed duration and terminal payoff $\rho(x(T), y(T))$, then

$$V(x_0, y_0, T) \geq \hat{\rho}_T(x_0, y_0).$$

Indeed, property 1^0 follows from non-negativity of the function $\rho(x, y)$. Let $C_P^T(x_0) \supset C_E^T(y_0)$. Then for every $y' \in C_E^T(y_0)$ there is $x' \in C_P^T(x_0)$ such that $\rho(x', y') = 0$, $(x' = y')$, whence follows 2^0. Property 3^0 follows from the fact that Player E can always guarantee himself an amount $\hat{\rho}_T(x_0, y_0)$ by choosing the motion directed towards the point $M \in C_E^T(y_0)$ for which

$$\hat{\rho}_T(x_0, y_0) = \min_{x \in C_P^T(x_0)} \rho(x, M).$$

The point M is called the *center of pursuit.*

5.2. Let $\Gamma_\delta(x_0, y_0, T)$ be a discrete game of pursuit with step δ $(\delta = t_{k+1} - t_k)$, prescribed duration T, discrimination against Player E, and initial states x_0, y_0. Then the following theorem holds.

Theorem 7. *In order for the following equality to hold for any $x_0, y_0 \in R^n$ and $T = \delta \cdot k$, $k = 1, 2, \ldots$:*

$$\hat{\rho}_T(x_0, y_0) = Val\Gamma_\delta(x_0, y_0, T), \tag{5.2}$$

it is necessary and sufficient that for all $x_0, y_0 \in R^n$, $\delta > 0$ and $T = \delta \cdot k, k = 1, 2, \ldots$, there be

$$\hat{\rho}_T(x_0, y_0) = \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T-\delta}(x, y) \tag{5.3}$$

($Val\Gamma_\delta(x_0, y_0, T)$ is the value of the game $\Gamma_\delta(x_0, y_0, T)$).

The proof of this theorem is based on the following result.

Lemma 7. *The following inequality holds for any $x_0, y_0 \in R^n$, $T \geq \delta$:*

$$\hat{\rho}_T(x_0, y_0) \leq \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T-\delta}(x, y).$$

Proof. By the definition of the function $\hat{\rho}_T$, we have

$$\max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T-\delta}(x, y) = \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \max_{\overline{y} \in C_E^{T-\delta}(y)} \min_{\overline{x} \in C_P^{T-\delta}(x)} \rho(\overline{x}, \overline{y}).$$

For all $x \in C_P^\delta(x_0)$ there is an inclusion $C_P^{T-\delta}(x) \subset C_P^T(x_0)$. Hence for any $x \in C_T^\delta(x_0)$, $\overline{y} \in C_E^{T-\delta}(y)$.

$$\min_{\overline{x} \in C_P^{T-\delta}(x)} \rho(\overline{x}, \overline{y}) \geq \min_{\overline{x} \in C_P^T(x_0)} \rho(\overline{x}, \overline{y}).$$

Then for all $x \in C_P^\delta(x_0)$, $y \in C_E^\delta(y)$

$$\max_{\overline{y} \in C_E^{T-\delta}(y)} \min_{\overline{x} \in C_P^{T-\delta}(x)} \rho(\overline{x}, \overline{y}) \geq \max_{\overline{y} \in C_E^{T-\delta}(y)} \min_{\overline{x} \in C_P^{T}(x_0)} \rho(\overline{x}, \overline{y})$$

and

$$\min_{x \in C_P^\delta(x_0)} \max_{\overline{y} \in C_E^{T-\delta}(y)} \min_{\overline{x} \in C_P^{T-\delta}(x)} \rho(\overline{x}, \overline{y}) \geq \max_{\overline{y} \in C_E^{T-\delta}(y)} \min_{\overline{x} \in C_P^{T}(x_0)} \rho(\overline{x}, \overline{y}).$$

Thus

$$\max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T-\delta}(x, y) \geq \max_{y \in C_E^\delta(y_0)} \max_{\overline{y} \in C_E^{T-\delta}(y)} \min_{\overline{x} \in C_P^{T}(x_0)} \rho(\overline{x}, \overline{y})$$

$$= \max_{y \in C_E^T(y_0)} \min_{x \in C_P^T(x_0)} \rho(x, y) = \hat{\rho}_T(x_0, y_0).$$

This completes the proof of lemma.

We shall now prove the Theorem.

Necessity. Suppose that condition (5.2) is satisfied and condition (5.3) is not. Then, by Lemma, there exist $\delta > 0$, $x_0, y_0 \in R^n$, $T_0 = \delta k_0$, $k_0 \geq 1$ such that

$$\hat{\rho}_{T_0}(x_0, y_0) < \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T_0-\delta}(x, y). \tag{5.4}$$

Let $u^0(\cdot)$ be an optimal strategy for Player P in the game $\Gamma_\delta(x_0, y_0, T_0)$ and suppose that at the 1st step Player E chooses the point $y^* \in C_E^\delta(y_0)$ for which

$$\min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T_0-\delta}(x, y^*) = \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T_0-\delta}(x, y). \tag{5.5}$$

Let $x^0(\delta)$ be the state to which Player P passes at the 1st step when he uses strategy $u^0(\cdot)$, and let $\overline{v}^0(\cdot)$ be an optimal strategy for Player E in the game $\Gamma_\delta(x_0(\delta), y^*, T_0 - \delta)$. Let us consider the strategy $\tilde{v}(\cdot)$ for Player E in the game $\Gamma_\delta(x_0, y_0, T_0)$: at the time instant $t = 0$ he chooses the point y^* and from the instant $t = \delta$ uses strategy $\overline{v}^0(\cdot)$.

Denote by $\hat{u}^0(\cdot)$ the truncation of strategy $u^0(\cdot)$ on the interval $[\delta, T_0]$. From (5.2), (5.4), (5.5) (by (5.2) $\hat{\rho}_T(x_0, y_0)$ is the value of the game $\Gamma_\delta(x_0, y_0, T)$) we find

$$\hat{\rho}_{T_0}(x_0, y_0) \geq K(u^0(\cdot), \tilde{v}(\cdot); x_0, y_0, T_0) = K(\hat{u}^0(\cdot), \overline{v}^0(\cdot); x^0(\delta), y^*, T_0 - \delta)$$

$$= \hat{\rho}_{T_0-\delta}(x^0(\delta), y^*) \geq \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T_0-\delta}(x, y^*) =$$

$$= \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T_0-\delta}(x, y) > \hat{\rho}_{T_0}(x_0, y_0).$$

This contradiction proves the necessity of condition (5.3).

Sufficiency. Note that the condition (5.3), in conjunction with the condition $\hat{\rho}_T(x_0, y_0)|_{T=0} = \rho(x_0, y_0)$, shows that the function $\hat{\rho}_T(x_0, y_0)$ satisfies the functional equation for the function of the value of the game $\Gamma_\delta(x_0, y_0, T)$. As follows from the proof of Theorem in 2.2, this condition is sufficient for $\hat{\rho}_T(x_0, y_0)$ to be the value of the game $\Gamma_\delta(x_0, y_0, T)$.

5.3.

Lemma 8. *In order for Player E's optimal open-loop strategy (i.e. the strategy which is the function of time only) to exist in the game $\Gamma(x_0, y_0, T)$ it is necessary and sufficient that*

$$Val\Gamma(x_0, y_0, T) = \hat{\rho}_T(x_0, y_0). \tag{5.6}$$

Proof. Sufficiency. Let $v^*(t)$, $t \in [0, T]$ be an admissible control for Player E which sends the point y_0 to a point M such that

$$\hat{\rho}_T(x_0, y_0) = \min_{x \in C_P^T(x_0)} \rho(x, M).$$

Denote $v^*(\cdot) = \{\sigma, v^*(t)\}$, where the partition σ of the interval $[0, T]$ consists of two points $t_0 = 0, t_1 = T$. Evidently, $v^*(\cdot) \in E$. $v^*(\cdot) \in E$ is an optimal strategy for Player E in the game $\Gamma(x_0, y_0, T)$ if

$$Val\Gamma(x_0, y_0, T) = \inf_{u(\cdot) \in P} K(u(\cdot), v^*(\cdot); x_0, y_0, T).$$

But this equality follows from (5.6), since

$$\inf_{u(\cdot) \in P} K(u(\cdot), v^*(\cdot); x_0, y_0, T) = \hat{\rho}_T(x_0, y_0).$$

Necessity. Suppose that in the game $\Gamma(x_0, y_0, T)$ there exists an optimal open-loop strategy for Player E. Then

$$Val\Gamma(x_0, y_0, T) = \sup_{v(\cdot) \in E} \inf_{u(\cdot) \in P} K(u(\cdot), v(\cdot); x_0, y_0, T)$$

$$= \max_{y \in C_E^T(y_0)} \inf_{u(\cdot) \in P} \rho(x(T), y) = \hat{\rho}_T(x_0, y_0).$$

This completes the proof of lemma.

Theorem 8. *In order for Player E to have an optimal open-loop strategy for any $x_0, y_0 \in R^n$, $T > 0$ in the game $\Gamma(x_0, y_0, T)$ it is necessary and sufficient that for any $\delta > 0$, $x_0, y_0 \in R^n$, $T \geq \delta$*

$$\hat{\rho}_T(x_0, y_0) = \max_{y \in C_E^\delta(y_0)} \min_{x \in C_P^\delta(x_0)} \hat{\rho}_{T-\delta}(x, y). \tag{5.7}$$

Proof. Sufficiency. By Theorem in 5.2, condition (5.7) implies relationship (5.2) from which, by passing to the limit (see Theorem in 3.7) we obtain

$$\hat{\rho}_T(x_0, y_0) = Val\Gamma(x_0, y_0, T).$$

By Lemma in 5.3, this implies existence of an optimal open-loop strategy for Player E.

Necessity of condition (5.7) follows from Theorem in 5.2, since the existence of an optimal open-loop strategy for Player E in the game $\Gamma(x_0, y_0, T)$ involves the existence of such a strategy in all games $\Gamma_\delta(x_0, y_0, T)$, $T = \delta k$, $k \geq 1$ and the validity of relationship (5.3).

6 Fundamental equation

In this section we will show that, under some particular conditions, the value function of the differential game satisfies a partial differential equation which is called fundamental. Although in monographic literatures [1] was the first to consider this equation, it is often referred to as the *Isaacs-Bellman equation*.

6.1. By employing Theorem in 5.3, we shall derive a partial differential equation for the value function of the differential game. We assume that the conditions of Theorem in 5.3 hold for the game $\Gamma(x, y, T)$. Then the function $\hat{\rho}_T(x, y)$ is the value of the game $\Gamma(x, y, T)$ of duration T from initial states x, y.

Suppose that in some domain Ω of the space $R^n \times R^n \times [0, \infty)$ the function $\hat{\rho}_T(x, y)$ has continuous partial derivatives in all its variables. We shall show that in this case the function $\hat{\rho}_T(x, y)$ in domain Ω satisfies the extremal differential equation

$$\frac{\partial \hat{\rho}}{\partial T} - \max_{v \in V} \sum_{i=1}^{n} \frac{\partial \hat{\rho}}{\partial y_i} g_i(y, v) - \min_{u \in U} \sum_{i=1}^{n} \frac{\partial \hat{\rho}}{\partial x_i} f_i(x, u) = 0, \tag{6.1}$$

where the functions $f_i(x, u), g_i(y, v)$, $i = 1, \dots, n$ determine the behavior of players in the game Γ (see (3.1), (3.2)).

Suppose that (6.1) fails to hold in some point $(x, y, T) \in \Omega$. For definiteness, let

$$\frac{\partial \hat{\rho}}{\partial T} - \max_{v \in V} \sum_{i=1}^{n} \frac{\partial \hat{\rho}}{\partial y_i} g_i(y, v) - \min_{u \in U} \sum_{i=1}^{n} \frac{\partial \hat{\rho}}{\partial x_i} f_i(x, u) < 0.$$

Let $\overline{v} \in V$ be such that in the point involved, $(x, y, T) \in \Omega$, the following relationship is satisfied:

$$\sum_{i=1}^{n} \frac{\partial \hat{\rho}}{\partial y_i} g_i(y, \overline{v}) = \max_{v \in V} \sum_{i=1}^{n} \frac{\partial \hat{\rho}}{\partial y_i} g_i(y, v).$$

Then the following inequality holds for any $u \in U$ in the point $(x, y, T) \in \Omega$:

$$\frac{\partial \hat{\rho}}{\partial T} - \sum_{i=1}^{n} \frac{\partial \hat{\rho}}{\partial y_i} g_i(y, \overline{v}) - \sum_{i=1}^{n} \frac{\partial \hat{\rho}}{\partial x_i} f_i(x, u) < 0. \tag{6.2}$$

From the continuous differentiability of the function $\hat{\rho}$ in all its variables it follows that the inequality (6.2) also holds in some neighbourhood S of the point (x, y, T). Let us choose a number $\delta > 0$ so small that the point $(x(\tau), y(\tau), T - \tau) \in S$ for all $\tau \in [0, \delta]$. Here

$$x(\tau) = x + \int_0^{\tau} f(x(t), u(t))dt,$$

$$y(\tau) = y + \int_0^{\tau} g(y(t), \overline{v}(t))dt$$

are the trajectories of systems (3.1), (3.2) corresponding to some admissible control $u(t)$ and $\overline{v}(t) \equiv \overline{v}$ and initial conditions $x(0) = x$, $y(0) = y$, respectively. Let us define the function

$$G(\tau) = \frac{\partial \hat{\rho}}{\partial T}|_{(x(\tau),y(\tau),T-\tau)} - \sum_{i=1}^{n} \frac{\partial \hat{\rho}}{\partial y_i}|_{(x(\tau),y(\tau),T-\tau)} g_i(y(\tau), \overline{v})$$

$$- \sum_{i=1}^{n} \frac{\partial \hat{\rho}}{\partial x_i}|_{(x(\tau),y(\tau),T-\tau)} f_i(x(\tau), u(\tau)), \quad \tau \in [0, \delta].$$

The function $G(\tau)$ is continuous in τ, therefore there is a number $c < 0$ such that $G(\tau) \leq c$ for $\tau \in [0, \delta]$. Hence we have

$$\int_0^\delta G(\tau) d\tau \leq c\delta. \tag{6.3}$$

It can be readily seen that

$$G(\tau) = -\frac{d\hat{\rho}}{d\tau}|_{(x(\tau),y(\tau),T-\tau)}.$$

From (6.3) we obtain

$$\hat{\rho}_T(x, y) - \hat{\rho}_{T-\delta}(x(\delta), y(\delta)) \leq c\delta.$$

Hence, by the arbitrariness of $u(t)$, it follows that

$$\hat{\rho}_T(x, y) < \max_{y' \in C_E^\delta(y)} \min_{x' \in C_P^\delta(x)} \hat{\rho}_{T-\delta}(x', y').$$

But this contradicts (5.7).

We have thus shown that in the case when Player E in the game $\Gamma(x, y, T)$ has an optimal open-loop strategy for any $x, y \in R^n$, $T > 0$, the value of the game $V(x, y, T)$ (it coincides with $\hat{\rho}_T(x, y)$ by Lemma in 5.3) in the domain of the space $R^n \times R^n \times [0, \infty)$, where this function has continuous partial derivatives, satisfies the equation

$$\frac{\partial V}{\partial T} = \max_{v \in V} \sum_{i=1}^{n} \frac{\partial V}{\partial y_i} g_i(y, v) + \min_{u \in U} \sum_{i=1}^{n} \frac{\partial V}{\partial x_i} f_i(x, u) \tag{6.4}$$

with the initial condition $V(x, y, T)|_{T=0} = \rho(x, y)$. Suppose we have defined $\overline{u}, \overline{v}$ computing max and min to (6.4) as functions of x, y and $\frac{\partial V}{\partial x}$, $\frac{\partial V}{\partial y}$ that is

$$\overline{u} = \overline{u}\big(x, \frac{\partial V}{\partial x}\big), \quad \overline{v} = \overline{v}\big(y, \frac{\partial V}{\partial y}\big). \tag{6.5}$$

Substituting expressions (6.5) into (6.4) we obtain

$$\sum_{i=1}^{n} \frac{\partial V}{\partial y_i} g_i\Big(y, \overline{v}\big(y, \frac{\partial V}{\partial y}\big)\Big) + \sum_{i=1}^{n} \frac{\partial V}{\partial x_i} f_i\Big(x, \overline{u}\big(x, \frac{\partial V}{\partial x}\big)\Big) = \frac{\partial V}{\partial T} \tag{6.6}$$

subject to

$$V(x, y, T)|_{T=0} = \rho(x, y). \tag{6.7}$$

Thus, to define $V(x, y, T)$ we have the initial value problem for the first order partial differential equation (6.6) with the initial condition (6.7).

Remark. In the derivation of the functional equations (6.4), (6.6), and in the proof of Theorem in 5.3, no use was made of a specific payoff function, therefore this theorem holds for any continuous terminal payoff $H(x(T), y(T))$. In this case, however, instead of the quantity $\hat{\rho}_T(x, y)$ we have to consider the quantity

$$\hat{H}_T(x, y) = \max_{y' \in C_E^T(y)} \min_{x' \in C_P^T(x)} H(x', y').$$

Equation (6.4) also holds for the value of the differential game with prescribed duration and any terminal payoff, i.e. if in the differential game $\Gamma(x, y, T)$ with prescribed duration and terminal payoff $H(x(T), y(T))$ there is an optimal open-loop strategy for Player E, then the value of the game $V(x, y, T)$ in the domain of the space $R^n \times R^n \times [0, \infty)$, where there exist continuous partial derivatives, satisfies equation (6.4) with the initial condition $V(x, y, T)|_{T=0} = H(x, y)$ or equation (6.6) with the same initial condition.

6.2. We shall now consider the games of pursuit in which the payoff function is equal to the time-to-capture. For definiteness, we assume that terminal manifold F is a sphere $\rho(x, y) = l$, $l > 0$. We also assume that the sets $C_P^t(x)$ and $C_E^t(y)$ are t-continuous in zero uniformly with respect to x and y.

Suppose the following quanitity makes sense:

$$\theta(x, y, l) = \max_{v(t)} \min_{u(t)} t_n^l(x, y; u(t), v(t)),$$

where $t_n^l(x, y; u(t), v(t))$ is the time of approach within l-distance for the players P and E moving from initial points x, y and using measurable open-loop controls $u(t)$ and $v(t)$, respectively. Also, suppose the function $\theta(x, y, l)$ is continuous in all its independent variables.

Let us denote the time-optimal game by $\Gamma(x_0, y_0)$. As in Secs. 5.4, 5.5, we may derive necessary and sufficient conditions for existence of an optimal open-loop strategy for Player E in the time-optimal game. The following theorem holds.

Theorem 9. *In order for Player E to have an optimal open-loop strategy for any $x_0, y_0 \in R^n$ in the game $\Gamma(x_0, y_0)$ it is necessary and sufficient that for any $\delta > 0$ and any $x_0, y_0 \in R^n$*

$$\theta(x_0, y_0, l) = \delta + \max_{y' \in C_E^\delta(y_0)} \min_{x' \in C_P^\delta(x_0)} \theta(x', y', l).$$

For the time-optimal game of pursuit the equation (6.4) becomes

$$\max_{v \in V} \sum_{i=1}^{n} \frac{\partial \theta}{\partial y_i} g_i(y, v) + \min_{u \in U} \sum_{i=1}^{n} \frac{\partial \theta}{\partial x_i} f_i(x, u) = -1 \tag{6.8}$$

with the initial condition

$$\theta(x,y,l)|_{\rho(x,y)=l} = 0. \tag{6.9}$$

Here it is assumed that there exist the first order continuous partial derivatives of the function $\theta(x,y,l)$ with respect to x, y. Assuming that the $\overline{u}, \overline{v}$ sending max and min to (6.8) can be defined as functions of $x, y, \partial\theta/\partial x, \partial\theta/\partial y$, i.e. $\overline{u} = \overline{u}(x, \frac{\partial\theta}{\partial x}), \overline{v} = \overline{v}(y, \frac{\partial\theta}{\partial y})$, we can rewrite equation (6.8) as

$$\sum_{i=1}^{n} \frac{\partial\theta}{\partial y_i} g_i\Big(y, \overline{v}(y, \frac{\partial\theta}{\partial y})\Big) + \sum_{i=1}^{n} \frac{\partial\theta}{\partial x_i} f_i\Big(x, \overline{u}(x, \frac{\partial\theta}{\partial x})\Big) = -1 \tag{6.10}$$

subject to

$$\theta(x,y,l)|_{\rho(x,y)=l} = 0. \tag{6.11}$$

The derivation of equation (6.8) is analogous to the derivation of equation (6.4) for the game of pursuit with prescribed duration.

Both initial value problems (6.4), (6.7) and (6.8), (6.9) are nonlinear in partial derivatives, therefore their solution presents serious difficulties.

6.3. We shall now derive equations of characteristics for (6.4). We assume that the function $V(x,y;T)$ has continuous mixed second derivatives over the entire space, the functions $g_i(y,v)$, $f_i(x,u)$ and the functions $\overline{u} = \overline{u}(x, \frac{\partial V}{\partial x})$, $\overline{v} = \overline{v}(y, \frac{\partial V}{\partial y})$ have continuous first derivatives with respect to all their variables, and the sets U, V have the aspect of parallelepipeds $a_m \leq u_m \leq b_m$, $m = 1, \ldots, k$ and $c_q \leq v_q \leq d_q$, $q = 1, \ldots, l$. where $u = (u_1, \ldots, u_k) \in U$, $v = (v_1, \ldots, v_l) \in V$. Denote

$$B(x,y,T) = \frac{\partial V}{\partial T} - \sum_{i=1}^{n} \frac{\partial V}{\partial x_i} f_i(x, \overline{x}) - \sum_{i=1}^{n} \frac{\partial V}{\partial y_i} g_i(y, \overline{v}).$$

The function $B(x,y,T) \equiv 0$, thus taking partial derivatives with respect to $x_1, \ldots, x_n$ we obtain

$$\frac{\partial B}{\partial x_k} = \frac{\partial^2 V}{\partial T \partial x_k} - \sum_{i=1}^{n} \frac{\partial^2 V}{\partial x_i \partial x_k} f_i - \sum_{i=1}^{n} \frac{\partial V}{\partial x_i} \frac{\partial f_i}{\partial x_k} - \sum_{i=1}^{n} \frac{\partial^2 V}{\partial y_i \partial x_k} g_i -$$

$$- \sum_{m=1}^{k} \frac{\partial}{\partial u_m} \Big(\sum_{i=1}^{n} \frac{\partial V}{\partial x_i} f_i \Big) \frac{\partial \overline{u}_m}{\partial x_k} - \sum_{i=1}^{l} \frac{\partial}{\partial v_q} \Big(\sum_{i=1}^{n} \frac{\partial V}{\partial y_i} g_i \Big) \frac{\partial \overline{v}_q}{\partial x_k} = 0, \quad k = 1, \ldots, n. \tag{6.12}$$

For every fixed point $(x,y,T) \in R^n \times R^n \times [0, \infty)$ the maximizing value $\overline{v}$ and the minimizing value $\overline{u}$ in (6.4) lie either inside or on the boundary of the interval of constraints. If this is an interior point, then

$$\frac{\partial}{\partial u_m} \Big(\sum_{i=1}^{n} \frac{\partial V}{\partial x_i} f_i \Big)|_{u=\overline{u}} = 0, \quad \frac{\partial}{\partial v_q} \Big(\sum_{i=1}^{n} \frac{\partial V}{\partial y_i} g_i \Big)|_{v=\overline{v}} = 0.$$

If, however, $\overline{u}(\overline{v})$ is at the boundary, then two cases are possible. Let us discuss these cases for one of the components $\overline{u}_m(x, \frac{\partial V}{\partial x})$ of the vector $\overline{u}$. The other components of vector $\overline{u}$ and vector $\overline{v}$ can be investigated in a similar manner. For simplicity assume that at some point (x', y', T')

$$\overline{u}_m = \overline{u}_m\left(x', \frac{\partial V(x', y', T')}{\partial x}\right) = a_m.$$

Case 1. In the space R^n there exists a ball with its center at the point x' and the following equality holds for all points x:

$$\overline{u}_m = \overline{u}_m\left(x, \frac{\partial V(x, y', T')}{\partial x}\right) = a_m.$$

The function $\overline{u}_m$ assumes on the ball a constant value; therefore in the point x' we have

$$\frac{\partial \overline{u}_m}{\partial x_i} = 0, \quad i = 1, \ldots, n.$$

Case 2. Such a ball does not exist. Then there is a sequence of interior points x_r, $\lim_{r\to\infty} x_r = x'$ such that

$$\overline{u}_m\left(x_r, \frac{\partial V(x_r, y', T')}{\partial x}\right) \neq a_m.$$

Hence

$$\frac{\partial}{\partial u_m}\left(\sum_{i=1}^{n} \frac{\partial V}{\partial x_i}|_{(x_r, y', T')} f_i(x_r, \overline{u})\right) = 0.$$

From the continuity of derivatives $\partial V/\partial x_i$, $\partial f_i/\partial u_m$ and function $\overline{u} = \overline{u}(x, \frac{\partial V(x,y,T)}{\partial x})$ it follows that the preceding equality also holds in the point (x', y', T').

Thus, the last two terms in (6.12) are zero and the following equality holds for all $(x, y, T) \in R^n \times [0, \infty)$:

$$\frac{\partial B}{\partial x_k} = \frac{\partial^2 V}{\partial T \partial x_k} - \sum_{i=1}^{n} \frac{\partial^2 V}{\partial x_i \partial x_k} f_i(x, \overline{u})$$

$$- \sum_{i=1}^{n} \frac{\partial V}{\partial x_i}\frac{\partial f_i}{\partial x_k} - \sum_{i=1}^{n} \frac{\partial^2 V}{\partial y_i \partial x_k} g_i(y, \overline{v}) = 0, \quad k = 1, 2, \ldots, n.$$

Let $\overline{x}(t), \overline{y}(t)$, $t \in [0, T]$ be a solution of the system

$$\dot{x} = f\left(x, \overline{u}\big(x, \frac{\partial V(x, y, T-t)}{\partial x}\big)\right),$$

$$\dot{y} = g\left(y, \overline{v}\big(y, \frac{\partial V(x, y, T-t)}{\partial y}\big)\right)$$

with the initial condition $x(0) = x_0$, $y(0) = y_0$. Along the solution $\overline{x}(t), \overline{y}(t)$ we have

$$\frac{\partial^2 V(\overline{x}(t), \overline{y}(t), T-t)}{\partial T \partial x_k} - \sum_{i=1}^{n} \frac{\partial^2 V(\overline{x}(t), \overline{y}(t), T-t)}{\partial x_i \partial x_k} f_i(\overline{x}(t), \tilde{u}(t))$$

$$-\sum_{i=1}^{n}\frac{\partial V(\overline{x}(t),\overline{y}(t),T-t)}{\partial x_i}\frac{\partial f_i(\overline{x}(t),\tilde{u}(t))}{\partial x_k}$$

$$-\sum_{i=1}^{n}\frac{\partial^2 V(\overline{x}(t),\overline{y}(t),T-t)}{\partial y_i \partial x_k}g_i(\overline{y}(t),\tilde{v}(t))=0, \qquad k=1,\dots,n, \tag{6.13}$$

where

$$\tilde{u}(t)=\overline{u}\left(\overline{x}(t),\frac{\partial V(\overline{x}(t),\overline{y}(t),T-t)}{\partial x}\right),$$

$$\tilde{v}(t)=\overline{v}\left(\overline{y}(t),\frac{\partial V(\overline{x}(t),\overline{y}(t),T-t)}{\partial y}\right).$$

However,

$$\frac{d}{dt}\left(\frac{\partial V(\overline{x}(t),\overline{y}(t),T-t)}{\partial x_k}\right)=\sum_{i=1}^{n}\frac{\partial^2 V(\overline{x}(t),\overline{y}(t),T-t)}{\partial x_k \partial x_i}f_i(\overline{x}(t),\tilde{u}(t))$$

$$+\sum_{i=1}^{n}\frac{\partial^2 V(\overline{x}(t),\overline{y}(t),T-t)}{\partial x_k \partial y_i}g_i(\overline{y}(t),\tilde{v}(t))$$

$$-\frac{\partial^2 V(\overline{x}(t),\overline{y}(t),T-t)}{\partial x_k \partial T}, \qquad k=1,\dots,n. \tag{6.14}$$

Note that for the twice continuously differentiable function we may reverse the order of differentiation. Now (6.13) can be rewritten in terms of (6.14) as

$$\frac{d}{dt}\left(\frac{\partial V(\overline{x}(t),\overline{y}(t),T-t)}{\partial x_k}\right)=-\sum_{i=1}^{n}\frac{\partial V(\overline{x}(t),\overline{y}(t),T-t)}{\partial x_i}\frac{\partial f_i(\overline{x}(t),\tilde{u}(t))}{\partial x_k},$$

$$k=1,...,n.$$

In a similar manner we obtain the equations

$$\frac{d}{dt}\left(\frac{\partial V(\overline{x}(t),\overline{y}(t),T-t)}{\partial y_i}\right)=-\sum_{j=1}^{n}\frac{\partial V(\overline{x}(t),\overline{y}(t),T-t)}{\partial y_j}\frac{\partial g_j(\overline{y}(t),\tilde{v}(t))}{\partial y_i},$$

$$i=1,\dots,n$$

Since for $t\in[0,T]$

$$V(\overline{x}(t),\overline{y}(t),T-t)=H(\overline{x}(T),\overline{y}(T)),$$

we have

$$\frac{d}{dt}\left(\frac{\partial V(\overline{x}(t),\overline{y}(t),T-t)}{\partial T}\right)=0.$$

Let us introduce the following notation:

$$V_{x_i}(t)\equiv\frac{\partial V(\overline{x}(t),\overline{y}(t),T-t)}{\partial x_i},$$

$$V_{y_i}(t) \equiv \frac{\partial V(\overline{x}(t), \overline{y}(t), T-t)}{\partial y_i}, \quad i = 1, \ldots, n;$$

$$V_x(t) \equiv \{V_{x_i}(t)\}, \quad V_y(t) \equiv \{V_{y_i}(t)\},$$

$$V_T(t) \equiv \frac{\partial V(\overline{x}(t), \overline{y}(t), T-t)}{\partial T}.$$

As a result we obtain the following system of ordinary differential equations for the functions $x(t), y(t), V_x(t), V_y(t)$:

$$\begin{aligned} \dot{x}_i &= f_i(x, \overline{u}(x, V_x)), \\ \dot{y}_i &= g_i(y, \overline{v}(y, V_y)), \\ \dot{V}_{x_k} &= -\sum_{i=1}^{n} V_{x_i} \frac{\partial f_i(x, \overline{u}(x, V_x))}{\partial x_k}, \\ \dot{V}_{y_k} &= -\sum_{i=1}^{n} V_{y_i} \frac{\partial g_i(y, \overline{v}(y, V_y))}{\partial y_k}, \\ V_T &= 0, \quad i, k = 1, \ldots, n, \end{aligned} \tag{6.15}$$

and, by (6.6), we have

$$V_T = \sum_{i=1}^{n} V_{y_i} g_i(y, \overline{v}(y, V_y)) + \sum_{i=1}^{n} V_{x_i} f_i(x, \overline{u}(x, V_x)).$$

In order to solve the system of nonlinear equations (6.15) with respect to the functions $x(t), y(t), V_{x_k}(t), V_{y_k}(t), V_T(t)$, we need to define initial conditions. For the function $V(\overline{x}(t), \overline{y}(t), T-t)$ such conditions are given at the time instant $t = T$, therefore we introduce the variable $\tau = T - t$ and write the equation of characteristics as a regression. Let us introduce the notation $\overset{\circ}{x} = -\dot{x}, \overset{\circ}{y} = -\dot{y}$. The equation of characteristics become

$$\begin{aligned} \overset{\circ}{x}_i &= -f_i(x, \overline{u}), \\ \overset{\circ}{y}_i &= -g_i(y, \overline{v}), \\ \overset{\circ}{V}_{x_k} &= \sum_{i=1}^{n} V_{x_i} \frac{\partial f_i(x, \overline{u})}{\partial x_k}, \\ \overset{\circ}{V}_{y_k} &= \sum_{i=1}^{n} V_{y_i} \frac{\partial g_i(y, \overline{v})}{\partial y_k}, \\ \overset{\circ}{V}_T &= 0. \end{aligned} \tag{6.16}$$

In the specification of initial conditions for system (6.16), use is made of the relationship $V(x, y, T)|_{T=0} = H(x, y)$. Let $x|_{\tau=0} = s$, $y_{\tau=0} = s'$. Then

$$V_{x_i}|_{\tau=0} = \frac{\partial H}{\partial x_i}|_{x=s, y=s'},$$

$$V_{y_i}|_{\tau=0} = \frac{\partial H}{\partial y_i}|_{x=s,y=s'}, \tag{6.17}$$

$$V_T|_{\tau=0} = \sum_{i=1}^{n} V_{y_i}|_{\tau=0} g_i(s', \overline{v}(s', V_y|_{\tau=0})) + \sum_{i=1}^{n} V_{x_i}|_{\tau=0} f_i(s, \overline{u}(s, V_x|_{\tau=0})).$$

Possible ways of solving system (6.16)-(6.17) are discussed in detail in [1].

In a similar manner, using equation (6.8) we may write the equation of characteristics for the problem of time-optimal pursuit.

7 Methods of successive approximations for solving differential games of pursuit

7.1. Let $\Gamma_\delta(x, y, T)$ be a discrete form of the differential game $\Gamma(x, y, T)$ of duration $T > 0$ with a fixed step of partition δ and discrimination against Player E for the time $\delta > 0$ in advance. Denote by $V_\delta(x, y, T)$ the value of the game $\Gamma_\delta(x, y, T)$.[1] Then

$$\lim_{\delta \to 0} V_\delta(x, y, T) = V(x, y, T)$$

and optimal strategies in the game $\Gamma_\delta(x, y, T)$ for sufficiently small δ can be efficiently used to construct ϵ-equilibria in the game $\Gamma(x, y, T)$.

7.2. The essence of the numerical method is to construct an algorithm of finding a solution of the game $\Gamma_\delta(x, y, T)$. We shall now expound this method.

Zero-order approximation. A zero-order approximation for the function of the value of the game $V_\delta(x, y, T)$ is taken to be the function

$$V_\delta^0(x, y, T) = \max_{\eta \in C_E^T(y)} \min_{\xi \in C_P^T(x)} \rho(\xi, \eta), \tag{7.1}$$

where $C_P^T(x), C_E^T(y)$ are reachability sets for the players P and E from initial states $x, y \in R^n$ by the time T.

The choice of the function $V_\delta^0(x, y, T)$ as an initial approximation is justified by the fact that in a sufficiently large class of games (what is called a regular case) it turns out to be the value of the game $\Gamma(x, y, T)$. The following approximations are constructed by the rule:

$$V_\delta^1(x, y, T) = \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} V_\delta^0(\xi, \eta, T - \delta),$$

$$V_\delta^2(x, y, T) = \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} V_\delta^1(\xi, \eta, T - \delta),$$

$$\dots$$

$$V_\delta^k(x, y, T) = \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} V_\delta^{k-1}(\xi, \eta, T - \delta) \tag{7.2}$$

for $T > \delta$ and $V_\delta^k(x, y, T) = V_\delta^0(x, y, T)$ for $T \leq \delta$, $k \geq 1$.

[1] The terminal payoff is equal to $\rho(x(T), y(T))$, where $\rho(x, y)$ is a distance in R^n.

As may be seen from formulas (7.2), the max min operation is taken over the reachability sets $C_E^\delta(y), C_P^\delta(x)$ for the time δ, i.e. for one step of the discrete game $\Gamma_\delta(x, y, T)$.

7.3.

Theorem 10. *For the fixed x, y, T, δ the numerical sequence $\{V_\delta^k(x, y, T)\}$ does not decrease with the growth of k.*

Proof. First we prove the inequality

$$V_\delta^1(x, y, T) \geq V_\delta^0(x, y, T).$$

For all $\xi \in C_P^\delta(x)$ there is $C_P^{T-\delta}(\xi) \subset C_P^T(x)$. For any $\overline{\eta} \in C_E^{T-\delta}(\eta)$, $\xi \in C_P^\delta(x)$ we have

$$\min_{\overline{\xi} \in C_P^{T-\delta}(\xi)} \rho(\overline{\xi}, \overline{\eta}) \geq \min_{\overline{\xi} \in C_P^T(x)} \rho(\overline{\xi}, \overline{\eta}).$$

Hence

$$V_\delta^1(x, y, T) = \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} \max_{\overline{\eta} \in C_E^{T-\delta}(\eta)} \min_{\overline{\xi} \in C_P^{T-\delta}(\xi)} \rho(\overline{\xi}, \overline{\eta})$$

$$\geq \max_{\eta \in C_E^\delta(y)} \max_{\overline{\eta} \in C_E^{T-\delta}(\eta)} \min_{\overline{\xi} \in C_P^T(x)} \rho(\overline{\xi}, \overline{\eta}) = \max_{\eta \in C_E^T(y)} \min_{\xi \in C_P^T(x)} \rho(\xi, \eta) = V_\delta^0(x, y, T).$$

We now assume that for $l \leq k$ there is

$$V_\delta^l(x, y, T) \geq V_\delta^{l-1}(x, y, T). \tag{7.3}$$

We prove this inequality for $l = k + 1$. From relationships (7.2) and (7.3) it follows that

$$V_\delta^{k+1}(x, y, T) = \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} V_\delta^k(\xi, \eta, T - \delta)$$

$$\geq \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} V_\delta^{k-1}(\xi, \eta, T - \delta) = V_\delta^k(x, y, T).$$

Thus, in case $T > \delta$, by induction, the statement of the theorem is proved (in case $T \leq \delta$ the statement of the theorem is obvious).

7.4.

Theorem 11. *The sequence $\{V_\delta^k(x, y, T)\}$ converges in a finite number of steps N, with the estimate $N \leq [\frac{T}{\delta}] + 1$, where the brackets stand for the integer part.*

Proof. Let $N = [T/\delta] + 1$. We show that

$$V_\delta^N(x, y, T) = V_\delta^{N+1}(x, y, T). \tag{7.4}$$

Equation (7.4) can be readily obtained from construction of the sequence $\{V_\delta^k(x, y, T)\}$. Indeed,

$$V_\delta^N(x, y, T) = \max_{\eta^1 \in C_E^\delta(y)} \min_{\xi^1 \in C_P^\delta(x)} V_\delta^{N+1}(\xi^1, \eta^1, T - \delta)$$

$$= \max_{\eta^1 \in C_E^\delta(y)} \min_{\xi^1 \in C_P^\delta(x)} \max_{\eta^2 \in C_E^\delta(\eta^1)} \dots$$

$$\dots \max_{\eta^{N-1}\in C_E^{\delta}(\eta^{N-2})} \min_{\xi^{N-1}\in C_P^{\delta}(\xi^{N-2})} V_\delta^1(\xi^{N-1},\eta^{N-1},T-(N-1)\delta).$$

Similarly we get

$$V_\delta^{N+1}(x,y,T)$$
$$= \max_{\eta_1\in C_E^{\delta}(y)} \min_{\xi^1\in C_P^{\delta}(x)} \max_{\eta^2\in C_E^{\delta}(\eta^1)} \dots$$
$$\dots \max_{\eta^{N-1}\in C_E^{\delta}(\eta^{N-2})} \min_{\xi^{N-1}\in C_P^{\delta}(\xi^{N-2})} V_\delta^2(\xi^{N-1},\eta^{N-1},T-(N-1)\delta).$$

But $T-(N-1)\delta=\alpha<\delta$, therefore

$$V_\delta^1(\xi^{N-1},\eta^{N-1},\alpha)=V_\delta^2(\xi^{N-1},\eta^{N-1},\alpha)=V_\delta^0(\xi^{N-1},\eta^{N-1},\alpha),$$

whence equality (7.4) follows.

The coincidence of members of the sequence V_δ^k for $k\geq N$ is derived from (7.4) by induction. This completes the proof of the theorem.

7.5.

Theorem 12. *The limit of the sequence* $\{V_\delta^k(x,y,T)\}$ *coincides with the value of the game* $\Gamma_\delta(x,y,T)$.

Proof. This theorem is essentially a corollary to Theoren in 7.4. Indeed, let

$$V_\delta(x,y,T)=\lim_{k\to\infty} V_\delta^k(x,y,T).$$

Convergence takes place in a finite number of steps not exceeding $N=[T/\delta]+1$; therefore in the recursion equation (7.2) we may pass to the limit as $k\to\infty$. The limiting function $V_\delta(x,y,T)$ satisfies the equation

$$V_\delta(x,y,T)=\max_{\eta\in C_E^{\delta}(y)} \min_{\xi\in C_P^{\delta}(x)} V_\delta(\xi,\eta,T-\delta) \tag{7.5}$$

with initial condition

$$V_\delta(x,y,T)|_{0\leq T\leq\delta}=\max_{\eta\in C_E^{T}(y)} \min_{\xi\in C_P^{T}(x)} \rho(\xi,\eta), \tag{7.6}$$

which is a sufficient condition for the function $V_\delta(x,y,T)$ to be the value of the game $\Gamma_\delta(x,y,T)$, (this is also a "regularity" criterion).

7.7. We shall now provide a modification of the method of successive approximation discussed above.

The initial approximation is taken to be the function $\tilde V_\delta^0(x,y,T) = V_\delta^0(x,y,T)$, where $V_\delta^0(x,y,T)$ is defined by (7.1). The following approximations are constructed by the rule:

$$\tilde V_\delta^{k+1}(x,y,T)=\max_{i\in[1:N]} \max_{\eta\in C_E^{i\delta}(y)} \min_{\xi\in C_P^{i\delta}(x)} \tilde V_\delta^k(\xi,\eta,T-i\delta)$$

for $T>\delta$, where $N=[T/\delta]$, and $\tilde V_\delta^{k+1}(x,y,T)=\tilde V_\delta^0(x,y,T)$ for $T\leq\delta$.

The statements of the theorems in 7.3–7.5 hold for the sequence of functions $\{\tilde{V}_\delta^k(x,y,T)\}$ and the sequence of functions $\{V_\delta^k(x,y,T)\}$.

The proof of these statements for the sequence of functions $\{\tilde{V}_\delta^k(x,y,T)\}$ is almost an exact replica of a similar argument for the sequence of functions $\{V_\delta^k(x,y,T)\}$. In the region $\{(x,y,T)|T>\delta\}$ the functional equation for the function of the value of the game $\Gamma_\delta(x,y,T)$ becomes

$$V_\delta(x,y,T) = \max_{i\in[1:N]} \max_{\eta\in C_E^{i\delta}(y)} \min_{\xi\in C_P^{i\delta}(x)} V_\delta(\xi,\eta,T-i\delta), \tag{7.7}$$

where $N=[T/\delta]$, while the initial condition remains unaffected, i.e. it is of the form (7.6).

7.7. We shall now prove the equivalence of equations (7.5) and (7.7).

Theorem 13. *Equations (7.5) and (7.7) with initial condition (7.6) are equivalent.*

Proof. Suppose the function $V_\delta(x,y,T)$ satisfies equation (7.5) and initial condition (7.6). Show that this function satisfies equation (7.7) in the region $\{(x,y,T)|T>\delta\}$.

Indeed, the following relationships hold:

$$V_\delta(x,y,T) = \max_{\eta\in C_E^{\delta}(y)} \min_{\xi\in C_P^{\delta}(x)} V_\delta(\xi,\eta,T-\delta) =$$

$$= \max_{\eta\in C_E^{\delta}(y)} \min_{\xi\in C_P^{\delta}(x)} \max_{\overline{\eta} C_E^{\delta}(\eta)} \min_{\overline{\xi}\in C_P^{\delta}(\xi)} V_\delta(\overline{\xi},\overline{\eta},T-2\delta) \geq$$

$$\geq \max_{\eta\in C_E^{\delta}(y)} \max_{\overline{\eta}\in C_E^{\delta}(\eta)} \min_{\xi\in C_P^{\delta}(x)} \min_{\overline{\xi}\in C_P^{\delta}(\xi)} V_\delta(\overline{\xi},\overline{\eta},T-2\delta) =$$

$$= \max_{\eta\in C_E^{2\delta}(y)} \min_{\xi\in C_P^{2\delta}(x)} V_\delta(\xi,\eta,T-2\delta) \geq \ldots \geq \max_{\eta\in C_E^{i\delta}(y)} \min_{\xi\in C_P^{i\delta}(x)} V_\delta(\xi,\eta,T-i\delta) \geq \ldots .$$

When $i=1$ we have

$$V_\delta(x,y,T) = \max_{\eta\in C_E^{\delta}(y)} \min_{\xi\in C_P^{\delta}(x)} V_\delta(\xi,\eta,T-\delta),$$

hence

$$V_\delta(x,y,T) = \max_{i\in[1:N]} \max_{\eta\in C_E^{i\delta}(y)} \min_{\xi\in C_P^{i\delta}(x)} V_\delta(\xi,\eta,T-i\delta),$$

where $N=[T/\delta]$, which proves the statement.

Now suppose the function $V_\delta(x,y,T)$ in the region $\{(x,y,T)|T>\delta\}$ satisfies equation (7.7) and initial condition (7.6). Show that this function also satisfies equation (7.5). Suppose the opposite is true. Then the following inequality must hold in the region $\{(x,y,T)|T>\delta\}$:

$$V_\delta(x,y,T) > \max_{\eta\in C_E^{\delta}(y)} \min_{\xi\in C_P^{\delta}(x)} V_\delta(\xi,\eta,T-\delta).$$

However,

$$\max_{\eta\in C_E^{\delta}(y)} \min_{\xi\in C_P^{\delta}(x)} V_\delta(\xi,\eta,T-\delta) =$$

$$= \max_{\eta \in C_E^\delta(y)} \min_{\xi \in C_P^\delta(x)} \max_{i \in [1:N-1]} \max_{\overline{\eta} \in C_E^{i\delta}(\eta)} \min_{\overline{\xi} \in C_P^{i\delta}(\xi)} V_\delta(\overline{\xi}, \overline{\eta}, T-(i+1)\delta) \geq$$

$$\geq \max_{\eta \in C_E^\delta(y)} \max_{i \in [1:N-1]} \max_{\overline{\eta} \in C_E^{i\delta}(\eta)} \min_{\xi \in C_P^\delta(x)} \min_{\overline{\xi} \in C_P^{i\delta}(\xi)} V_\delta(\overline{\xi}, \overline{\eta}, T-(i+1)\delta) =$$

$$= \max_{i \in [1:N-1]} \max_{\eta \in C_E^\delta(y)} \max_{\overline{\eta} \in C_E^{i\delta}(\eta)} \min_{\xi \in C_P^\delta(x)} \min_{\overline{\xi} \in C_P^{i\delta}(\xi)} V_\delta(\overline{\xi}, \overline{\eta}, T-(i+1)\delta) =$$

$$= \max_{i \in [2:N]} \max_{\eta \in C_E^{i\delta}(y)} \min_{\xi \in C_P^{i\delta}(x)} V_\delta(\xi, \eta, T-i\delta) = V_\delta(x, y, T).$$

Since for $i = 1$ the strong inequality holds, this contradiction proves the theorem.

8 Examples of solutions to differential games of pursuit

8.1. *Example 4. Simple motion.* Let us consider the differential game $\Gamma(x_0, y_0, T)$ in which the motion by the Players P and E in the Euclidean space R^n is governed by the following equations:

$$\text{for } P: \ \dot{x} = \alpha u, \ \|u\| < 1, \ x(0) = x_0,$$

$$\text{for } E: \ \dot{y} = \beta v, \ \|v\| < 1, \ y(0) = y_0, \tag{8.1}$$

where α, β are constants $\alpha > \beta > 0$, $x, y, u, v \in R^n$.

The payoff to Player E is

$$H(x(T), y(T)) = \|x(T) - y(T)\|.$$

Let $\Gamma_\delta(x, y, T)$ be a discrete form of the differential game $\Gamma(x, y, T)$ with the partition step $\delta > 0$ and discrimination against Player E. The game $\Gamma_\delta(x, y, T)$ has N steps, where $N = T/\delta$. The game $\Gamma_\delta(x, y, T)$ has the value

$$V_\delta(x, y, T) = \max\{0, \|x - y\| - N \cdot \delta \cdot (\alpha - \beta)\} = \max\{0, \|x - y\| - T(\alpha - \beta)\},$$

and the optimal motion by players is along the straight line connecting the initial states x, y.

By the results of 5.3, the value of the original differential game

$$V(x, y, T) = \lim_{\delta \to 0} V_\delta(x, y, T) = \max\{0, \|x - y\| - T(\alpha - \beta)\}. \tag{8.2}$$

It can be seen that

$$V(x, y, T) = \max_{y' \in C_E^T(y)} \min_{x' \in C_P^T(x)} \|x' - y'\| = \hat{\rho}_T(x, y),$$

where $C_E^T(y) = S(y, \beta T)$ is the ball in R^n of radius βT with its center at the point y, similarly $C_P^T(x) = S(x, \alpha T)$. Thus, by Lemma in 5.3, Player E in the game $\Gamma(x_0, y_0, T)$ has the optimal open-loop strategy $v^*(t)$, $t \in [0, T]$, which leads Player E's trajectory to the point $y^* \in C_E^T(y_0)$ for which

$$\hat{\rho}_T(x_0, y_0) = \min_{x' \in C_P^T(x_0)} \|x' - y^*\|.$$

Evidently,

$$v^*(t) \equiv v^* = \begin{cases} \frac{y_0 - x_0}{\|y_0 - x_0\|}, & with\ y_0 \neq x_0, \\ v, & with\ y_0 = x_0, \end{cases}$$

where $v \in R^n$ is an arbitrary vector such that $\|v\| = 1$. From the results of 5.6 it follows that in the region Δ

$$\Delta = \{(x, y, T) : \|x - y\| - T(\alpha - \beta) > 0\},$$

where there exist continuous partial derivatives

$$\frac{\partial V}{\partial T} = -(\alpha - \beta), \quad \frac{\partial V}{\partial x} = -\frac{\partial V}{\partial y} = \frac{x - y}{\|x - y\|},$$

the function $V(x, y, T)$ satisfies equation (7.4):

$$\frac{\partial V}{\partial T} - \alpha \min_{\|u\|<1} (\frac{\partial V}{\partial x}, u) - \beta \max_{\|v\| \leq 1} (\frac{\partial V}{\partial y}, v) = 0. \tag{8.3}$$

In equation (8.3), minimum and maximum are achieved under controls

$$\overline{u}(x, \frac{\partial V}{\partial x}) = -\frac{\frac{\partial V}{\partial x}}{\|\frac{\partial V}{\partial x}\|} = \frac{y - x}{\|y - x\|}; \tag{8.4}$$

$$\overline{v}(y, \frac{\partial V}{\partial y}) = \frac{\frac{\partial V}{\partial y}}{\|\frac{\partial V}{\partial y}\|} = \frac{y - x}{\|y - x\|}. \tag{8.5}$$

Strategies (8.4), (8.5) are optimal in the differential game (8.1). The strategy $\overline{u}(x, y)$ determined by relationship (8.4) is called a "pursuit strategy", since at each instant of time for Player P using this strategy the vector of his velocity is pointing towards Evader E.

8.2. *Example 5. Game of pursuit with frictional forces.* The pursuit takes place over the plane. Equations of motion are of the form:

for P:

$$\dot{q}_i = p_i,$$
$$\dot{p}_i = \alpha u_i - k_P p_i, \quad i = 1, 2, \quad \|u\| \leq 1; \tag{8.6}$$

for E:

$$\dot{r}_i = s_i,$$
$$\dot{s}_i = \beta v_i - k_E s_i, \quad i = 1, 2, \quad \|v\| \leq 1; \tag{8.7}$$
$$q_i(0) = q_i^0, \ p_i(0) = p_i^0, \ r_i(0) = r_i^0,$$

$$s_i(0) = s_i^0, \quad i = 1, 2, \quad \alpha, \beta, k_E, k_P > 0. \tag{8.8}$$

Here $q = (q_1, q_2)$ and $r = (r_1, r_2)$ are positions on the plane for players P and E, respectively; $p = (p_1, p_2)$ and $s = (s_1, s_2)$ are the players' momenta; k_P, k_E are some constants interpreted to mean friction coefficients.

The payoff to Player E is taken to be

$$H(q(T), r(T)) = \|q(T) - r(T)\| = \sqrt{[q_1(T) - r_1(T)]^2 + [q_2(T) - r_2(T)]^2}.$$

In the plane $q = (q_1, q_2)$, the reachability set $C_P^T(q^0, p^0)$ for Player P from the initial states $p(0) = p^0$, $q(0) = q^0$ in the time T is the circle (Exercise 18) of radius

$$R_P(T) = \frac{\alpha}{k_P^2}(e^{-k_P T} + k_P T - 1)$$

with its center at the point

$$a(q^0, p^0, T) = q^0 + p^0 \frac{1 - e^{-k_P T}}{k_P}.$$

Similarly, the set $C_E^T(r^0, s^0)$ is the circle of radius

$$R_E(T) = \frac{\beta}{k_E^2}(e^{-k_E T} + k_E T - 1)$$

with its center at the point

$$b(r^0, s^0, T) = r^0 + s^O \frac{1 - e^{-k_E T}}{k_E}.$$

For the quantity $\hat{\rho}_T(q^0, p^0, r^0, s^0)$ determined by relationship (5.1), in this differential game there is

$$\hat{\rho}_T(q^0, p^0, r^0, s^0) = \max_{r \in C_E^T(r^0, s^0)} \min_{q \in C_P^T(q^0, p^0)} \|q - r\|.$$

Hence (see formula (2.10)) we have

$$\begin{aligned} \hat{\rho}_T(q, p, r, s) &= \max\{0, \|a(q, p, T) - b(r, s, T)\| - (R_P(T) - R_E(T))\} \\ &= \max\left\{0, \sqrt{\sum_{i=1}^{2}\left(q_i - r_i + p_i \frac{1 - e^{-k_P T}}{k_P} - s_i \frac{1 - e^{-k_E T}}{k_E}\right)^2} \right. \\ &\quad \left. - \left(\alpha \frac{e^{-k_P T} + k_P T - 1}{k_P^2} - \beta \frac{e^{-k_E T} + k_E T - 1}{k_E^2}\right)\right\}. \end{aligned} \tag{8.9}$$

In particular, the conditions $\alpha > \beta$, $\frac{\alpha}{k_P} > \frac{\beta}{k_E}$ suffice to ensure that for any initial states q, p, r, s there is a suitable T for which $\hat{\rho}_T(q, p, r, s) = 0$.

The function $\hat{\rho}_T(q,p,r,s)$ satisfies the extremal differential equation (7.1) in the domain $\Omega = \{(q,p,r,s,T) : \hat{\rho}_T(q,p,r,s) > 0\}$. In fact, in the domain Ω there are continuous partial derivatives

$$\frac{\partial \hat{\rho}}{\partial T}, \ \frac{\partial \hat{\rho}}{\partial q_i}, \ \frac{\partial \hat{\rho}}{\partial p_i}, \ \frac{\partial \hat{\rho}}{\partial r_i}, \ \frac{\partial \hat{\rho}}{\partial s_i}, \quad i = 1, 2. \tag{8.10}$$

Equation (7.1) becomes

$$\frac{\partial \hat{\rho}}{\partial T} - \sum_{i=1}^{2} \Big(\frac{\partial \hat{\rho}}{\partial q_i} p_i + \frac{\partial \hat{\rho}}{\partial r_i} s_i - \frac{\partial \hat{\rho}}{\partial p_i} k_P p_i - \frac{\partial \hat{\rho}}{\partial s_i} k_E s_i \Big)$$

$$-\beta \max_{\|v\| \le 1} \sum_{i=1}^{2} \frac{\partial \hat{\rho}}{\partial s_i} v_i - \alpha \min_{\|u\| \le 1} \sum_{i=1}^{2} \frac{\partial \hat{\rho}}{\partial p_i} u_i = 0. \tag{8.11}$$

Here extrema are achieved on the controls $\overline{u}, \overline{v}$ determined by the following formulas:

$$\overline{u}_i = -\frac{\frac{\partial \hat{\rho}}{\partial p_i}}{\sqrt{(\frac{\partial \hat{\rho}}{\partial p_1})^2 + (\frac{\partial \hat{\rho}}{\partial p_2})^2}}, \tag{8.12}$$

$$\overline{v}_i = \frac{\frac{\partial \hat{\rho}}{\partial s_i}}{\sqrt{(\frac{\partial \hat{\rho}}{\partial s_1})^2 + (\frac{\partial \hat{\rho}}{\partial s_2})^2}}, \quad i = 1, 2. \tag{8.13}$$

Substituting these controls into (8.11) we obtain the nonlinear first-order partial differential equation

$$\frac{\partial \hat{\rho}}{\partial T} - \sum_{i=1}^{2} \Big(\frac{\partial \hat{\rho}}{\partial q_i} p_i + \frac{\partial \hat{\rho}}{\partial r_i} s_i - \frac{\partial \hat{\rho}}{\partial p_i} k_P p_i - \frac{\partial \hat{\rho}}{\partial s_i} k_E s_i \Big) - \beta \sqrt{\big(\frac{\partial \hat{\rho}}{\partial s_1}\big)^2 + \big(\frac{\partial \hat{\rho}}{\partial s_2}\big)^2}$$

$$+\alpha \sqrt{(\frac{\partial \hat{\rho}}{\partial p_1})^2 + (\frac{\partial \hat{\rho}}{\partial p_2})^2} = 0. \tag{8.14}$$

Computing the partial derivatives (8.10) we see that the function $\hat{\rho}_T(q,p,r,s)$ in the domain Ω satisfies equation (8.14). Note that the quantity $\hat{\rho}_T(q^0,p^0,r^0,s^0)$ is the value of the differential game (8.6)–(8.8) and the controls determined by relationships (8.12), (8.13) are optimal in the domain Ω.

From formulas (8.12), (8.13), (8.9) we find

$$\overline{u}_i = \frac{r_i - q_i + s_i \frac{1-e^{-k_E T}}{k_E} - p_i \frac{1-e^{-k_P T}}{k_P}}{\sqrt{\sum_{i=1}^{2} (r_i - q_i + s_i \frac{1-e^{-k_E T}}{k_E} - p_i \frac{1-e^{-k_P T}}{k_P})^2}}, \quad \overline{v}_i = \overline{u}_i, \quad i = 1, 2. \tag{8.15}$$

In the situation $\overline{u}, \overline{v}$ the force direction for each of the players is parallel to the line connecting the centers of reachability circles (as follows from formula (8.15)) and remains unaffected, since in this situation the centers of reachability circles move along the straight line.

9 Games of pursuit with delayed information for Pursuer

9.1. In this chapter we investigated games examined conflict-controlled processes where each participant (player) has perfect information, i.e. at each current instant of time Player $P(E)$ is aware of his state $x(t)[y(t)]$ and the opponent's state $y(t)[x(t)]$. Existence theorems were obtained for pure strategy ϵ-equilibria in such games and various methods for constructing solutions were illustrated. This was made possible by the fact that the differential games with perfect information are the limiting case of multistage games with perfect information where the time interval between two sequential moves tends to zero. In differential games with incomplete information, where mixed strategies play an important role, we have a completely different situation. Without analyzing the entire problem we will deal with game of pursuit with prescribed duration, terminal payoff and delayed information for Player P on the phase state (state variable) of Player E, the time of delay being $l > 0$.

9.2. Let there be given some number $l > 0$ referred to as the information delay. For $0 \leq t \leq l$, Pursuer P at each instant of time t knows his own state $x(t)$, the time t and the initial position y_0 of Evader E. For $l \leq t \leq T$, Player P at each instant of time t knows his own state $x(t)$, the time t and the state $y(t-l)$ of Player E at the time instant $t-l$. Player E at each instant of time t knows his own state $y(t)$, the opponent's state $x(t)$ and the time t. His payoff is equal to a distance between the players at the time instant T, the payoff to Player P is equal to the payoff to Player E but opposite in sign (the game is zero-sum). Denote this game by $\Gamma(x_0, y_0, T)$.

Definition 10. *The pure piecewise open-loop strategy $v(\cdot)$ for Player E means the pair $\{\tau, b\}$, where τ is a partitioning of the time interval $[0, T]$ by a finite number of points $0 = t_1 < \ldots < t_k = T$, and b is the map which places each state $x(t_i), y(t_i), t_i$ in correspondence with the measurable open-loop control $v(t)$ of Player E for $t \in [t_i, t_{i+1})$.*

Definition 11. *The pure piecewise open-loop strategy $u(\cdot)$ for Player P means the pair $\{\sigma, a\}$, where σ is an arbitrary partitioning of the time interval $[0, T)$ by a finite number of points $0 = t'_1 < t'_2 < \ldots < t'_s = T$, and a is the map which places each state $x(t'_i), y(t'_i - l)$, t'_i for $l \leq t'_i$ in correspondence with the segment of Player P's measurable open-loop control $u(t)$ for $t \in [t'_i, t'_{i+1})$. For $t'_i \leq l$, the map a places each state $x(t'_i), y_0, t'_i$ in correspondence with the segment of Player P's measurable control $u(t)$ for $t \in [t'_i, t'_{i+1})$.*

The sets of all pure piecewise open-loop strategies for the players P and E are denoted by P and E, respectively.

Equations of motion are of the form

$$\dot{x} = f(x, u),\ u \in U \subset R^p,\ x \in R^n,$$

$$\dot{y} = g(y, v),\ v \in V \subset R^q,\ y \in R^n. \tag{9.1}$$

We assume that the conditions which ensure the existence and uniqueness of a solution to system (9.1) for any pair of measurable open-loop controls $u(t), v(t)$ with the given initial conditions x_0, y_0 are satisfied. This ensures the existence of a unique solution to system (9.1) where the players P and E use piecewise open-loop strategies $u(\cdot) \in P$, $v(\cdot) \in E$ with the given initial conditions x_0, y_0. Thus, in any situation $(u(\cdot), v(\cdot))$ with the given initial conditions x_0, y_0 the payoff function for Player E is determined in a unique way

$$K(x_0, y_0; u(\cdot), v(\cdot)) = \rho(x(T), y(T)), \tag{9.2}$$

where $x(t), y(t)$ is a solution to system (9.1) with initial conditions x_0, y_0 in situation $(u(\cdot), v(\cdot))$, and ρ is the Euclidean distance.

9.3. We can demonstrate with simple examples that in the game under study $\Gamma(x_0, y_0, T)$ the ϵ-equilibria do not exist for all $\epsilon > 0$. For this reason, to construct equilibria, we shall follow the way proposed by [3] for finite positional games with incomplete information. The strategy spaces of the players P and E will be extended to what are called *mixed piecewise open-loop behavior strategies* (MPOLBS) which allows for a random choice of control at each step.

Example 6. Equations of motion are of the form

$$\begin{gathered} \text{for } P: \ \dot{x} = u, \ \|u\| \leq \alpha, \\ \text{for } E: \ \dot{y} = v, \ \|v\| \leq \beta, \\ \alpha > \beta > 0, \ x, y, \in R^2, \ u, v \in R^2. \end{gathered} \tag{9.3}$$

The payoff to Player E is $\rho(x(T), y(T))$, where $x(t), y(t)$ is a solution to system (9.3) with the initial conditions $x(t_0) = x_0$, $y(t_0) = y_0$. Player P is informed only about the initial state y_0 of his opponent, while Player E is completely informed about Player P's state ($l = T$).

Let $\overline{v}(x, y, t)$ be some piecewise open-loop strategy for Player E. For each strategy $\overline{v}$ there is a strategy $\overline{u}(x, t)$ of Player P using only information about the initial position of Player E, his current position and the time from the start of the game, for which a payoff of $\rho(x(T), y(T)) \leq \epsilon$ for $T \geq \rho(x_0, y_0)/(\alpha - \beta)$. Indeed, let $u^*(x, y, t)$ be a strategy for Player P in the game with perfect information. The strategy is as follows: Player E is pursued until the capture time t_n (while the capture of E takes place) while for $t_n \leq t \leq T$ the point $x(t)$ is kept in some ϵ-neighbourhood of the evading point. It is an easy matter to describe analytically such a strategy in the game with perfect information (see Example 4, 8.1). Let us construct the players' trajectories $\overline{x}(t), \overline{y}(t)$ in situation $(u^*(x, y, t), \overline{v}(x, y, t))$ from the initial states x_0, y_0. To do this, it suffices to integrate the system

$$\begin{gathered} \dot{x} = u^*(x, y, t), \ \overline{x}(t_0) = x_0, \\ \dot{y} = \overline{v}(x, y, t), \ \overline{y}(t_0) = y_0. \end{gathered} \tag{9.4}$$

By construction $\rho(\overline{x}(T), \overline{y}(T)) \leq \epsilon$. Now let $\tilde{u}(t) = u^*(\overline{x}(t), \overline{y}(t), t)$. Although the strategy $u^*(x, y, t)$ using the information about E's position is inadmissible, the strategy $\tilde{u}(t)$ is admissible since it uses only information about the time from the start of the game and information about the initial state of Player E. It is apparent that in situations $(\tilde{u}(t), \overline{v}(x, y, t))$ and $(u^*(x, y, t), \overline{v}(x, y, t))$ the players' paths coincide since the strategy $\overline{v}(x, y, t)$ responds to the strategy $u^*(x, y, t)$ and the strategy $\tilde{u}(t)$ by choosing the same control $\overline{v}(\overline{x}(t), \overline{y}(t), t)$.

We have thus shown that for each strategy $\overline{v}(x, y, t)$ there is an open-loop control $\tilde{u}(t)$ which is an admissible strategy in the game with incomplete information and is such that $\rho(\overline{x}(T), \overline{y}(T)) \leq \epsilon$, where $\overline{x}(t), \overline{y}(t)$ are the corresponding trajectories. The choice of $\overline{v}(x, y, t)$ is made in an arbitrary way, hence it follows that

$$\sup\inf \rho(x(T), y(T)) = 0, \tag{9.5}$$

where sup inf is taken over the players' strategy sets in the game with incomplete information.

For any strategy $u(x, t)$ of Player P, however, we may construct a strategy $v(x, y, t)$ for Player E such that in situation $(u(x, t), v(x, y, t))$ the payoff ρ to Player E will exceed βT. Indeed, let $\overline{u}(x, t)$ be a strategy for Player P. Since his motion is independent of $y(t)$, the path of Player P can be obtained by integrating the system

$$\dot{x} = \overline{u}(x, t), \quad x(t_0) = x_0 \tag{9.6}$$

irrespective of what motion is made by Player E. Let $\overline{x}(t)$ be a trajectory resulting from integration of system (9.6). The points $\overline{x}(T)$ and y_0 are connected and the motion by Player E is oriented along the straight line $[\overline{x}(T), y_0]$ away from the point $\overline{x}(T)$. His speed is taken to be maximum. Evidently, the motion by Player E ensures a distance between him and the point $\overline{x}(T)$ which is greater than or equal to βT. Denote the thus constructed strategy for Player E by $\overline{v}(t)$. In the situation $(\overline{u}(x, t), \overline{v}(t))$, the payoff to Player E is then greater than or equal to βT. From this it follows that

$$\inf\sup \rho(x(T), y(T)) \geq \beta T, \tag{9.7}$$

where inf sup is taken over the players' strategy sets in the game with incomplete information.

It follows from (9.5) and (9.7) that the value of the game in the class of pure strategies does not exist in the game under study.

9.4.

Definition 12. *The mixed piecewise open-loop behavior strategy (MPOLBS) for Player P means the pair $\mu(\cdot) = \{\tau, d\}$, where τ is an arbitrary partitioning of the time interval $[0, T]$ by a finite number of points $0 = t_1 < t_2 < \ldots < t_k = T$, and d is the map which places each state $x(t_i), y(t_i - l), t_i$ for $t_i > l$ and the state $x(t_i), y_0, t_i$ for $t_i \leq l$ in correspondence with the probability distribution $\mu'_i(\cdot)$ concentrated on a finite number of measurble open-loop controls $u(t)$ for $t \in [t_i, t_{i+1})$.*

Similarly, MPOLBS for Player E means the pair $\nu(\cdot) = \{\sigma, c\}$, where σ is an arbitrary partitioning of the time interval $[0, T]$ by a finite number of points $0 = t'_1 < t'_2 < \ldots < t'_s = T$, and c is the map which places the state $x(t'_i), y(t'_i), t'_i$ in correspondence with the probability distribution $\nu'_i(\cdot)$ concentrated on a finite number of measurable open-loop controls $v(t)$ for $t \in [t_i, t_{i+1})$.

MPOLBS for the players P and E are denoted respectively by $\overline{P}$ and $\overline{E}$ (compare these strategies with "behavior strategies").

Each pair of MPOLBS $\mu(\cdot), \nu(\cdot)$ induces the probability distribution over the space of trajectories $x(t), x(0) = x_0$; $y(t), y(0) = y_0$. For this reason, the payoff $\overline{K}(x_0, y_0; \mu(\cdot), \nu(\cdot))$ in MPOLBS is interpreted to mean the mathematical expectation of the payoff averaged over the distributions over the trajectory spaces that are induced by MPOLBS $\mu(\cdot), \nu(\cdot)$. Having determined the strategy spaces $\overline{P}, \overline{E}$ and the payoff $\overline{K}$ we have determined the mixed extension $\overline{\Gamma}(x_0, y_0, T)$ of the game $\Gamma(x_0, y_0, T)$.

9.5. Denote by $C_P^T(x)$ and $C_E^T(y)$ the respective reachability sets of the players P and E from initial states x and y at the instant of time T, and by $\overline{C}_E^T(y)$ the convex hull of the set $C_E^T(y)$. We assume that the reachability sets are compact, and introduce the quantity

$$\gamma(y, T) = \min_{\xi \in \overline{C}_E^T(y)} \max_{\eta \in C_E^T(y)} \rho(\xi, \eta).$$

Let $\gamma(y, T) = \rho(\tilde{y}, \overline{y})$, where $y \in \overline{C}_E^T(y)$, $\overline{y} \in C_E^T(y)$. From the definition of the point $\tilde{y}$ it follows that it is a center of the minimal sphere containing the set $C_E^T(y)$. Hence it follows that this point is unique. At the same time, there exist at least two points of tangency of the set $C_E^T(y)$ to the minimal sphere containing it, these points coinciding with the points $\overline{y}$.

Let $y(t)$ be a trajectory $(y(0) = y_0)$ of Player E for $0 \le t \le T$. When Player E moves along this trajectory the value of the quanitity $\gamma(y(t), T - t)$ changes, the point $\tilde{y}$ also changes. Let $\tilde{y}(t)$ be a trajectory of the point $\tilde{y}$ corresponding to the trajectory $y(t)$. The point $M \in C_E^{T-l}(y_0)$ will be referred to as the *center of pursuit* if

$$\gamma(M, l) = \max_{y' \in C_E^{T-l}(y_0)} \gamma(y', l).$$

9.6. We shall now consider an auxiliary simultaneous zero-sum game of pursuit over a convex hull of the set $C_E^T(y)$. Pursuer chooses a point $\xi \in \overline{C}_E^T(y)$ and Evader chooses a point $\eta \in C_E^T(y)$. The choices are made simultaneously. When choosing the point ξ, Player P has no information on the choice of η by Player E, and conversely. Player E receives a payoff $\rho(\xi, \eta)$. We denote the value of this game by $V(y, T)$ in order to emphasize the dependence of the game value on the parameters y and T which determine the strategy sets $\overline{C}_E^T(y)$ and $C_E^T(y)$ for players P and E, respectively. The game in normal form can be written as follows:

$$\Gamma(y, T) = \langle \overline{C}_E^T(y), C_E^T(y), \rho(y', y'') \rangle.$$

The strategy set of the minimizing player P is convex, the function $\rho(y', y'')$ is also convex in its independent variables and is continuous. Therefore the game $\Gamma(y, T)$ has an equilibrium in mixed strategies. An optimal strategy for Player P is pure, and an optimal strategy for Player E assigns the positive probability to at most $(n + 1)$ points from the set $C_E^T(y)$, with $V(y, T) = \gamma(y, T)$. An optimal strategy for Player P in the game $\Gamma(y,T)$ is the choice of a center of the minimal sphere $\tilde{y}$ containing the set $C_E^T(y)$. An optimal strategy for Player E assigns the positive probabilities to at most $(n+1)$ points among the points of tangency of the sphere to the set $C_E^T(y)$ (here n is the dimension of the space of y). The value of the game is equal to the radius of this sphere.

9.7. We shall now consider a simultaneous game $\Gamma(M,l)$, where M is the center of pursuit. Denote by $\overline{y}_1(M),\dots,\overline{y}_{n+1}(M)$ the points from the set $C_E^l(M)$ appearing in the spectrum of an optimal mixed strategy for Player E in the game $\Gamma(M,l)$ and by $y(M)$ an optimal strategy for Player P in this game.

Definition 13. *The trajectory $y^*(t)$ is called conditionally optimal if $y^*(0)=y_0$, $y^*(T-l)=M$, $y^*(T)=\overline{y}_i(M)$ for some i from the numbers $1,\dots,n+1$.*

For each i there can be several conditionally optimal strategies of Player E.

Theorem 14. *Let $T\geq l$ and suppose that for any number $\epsilon>0$ Player P can ensure by the time T the ϵ-capture of the center $\tilde{y}(T)$ of the minimal sphere containing the set $C_E^l(y(T-l))$. Then the game $\overline{\Gamma}(x_0,y_0,T)$ has the value $\gamma(M,l)$, and the ϵ-optimal strategy of Player P is pure and coincides with any one of his strategies which may ensure the $\epsilon/2$-capture of the point $\tilde{y}(T)$. An optimal strategy for Player E is mixed: during the time $0\leq t\leq T-l$ he must move to the point M along any conditionally optimal trajectory $y^*(t)$ and then, with probabilities $p_1,\dots,p_{n+1}$ (the optimal strategy for Player E in the game $\Gamma(M,l)$), he must choose one of the conditionally optimal trajectories sending the point $y^*(T-l)$ to the points $\overline{y}_i(M)$, $i=1,\dots,n+1$ which appear in the spectrum of an optimal mixed strategy for Player E in the game $\Gamma(M,l)$.*

Proof. Denote by $u_\epsilon(\cdot),\nu_*(\cdot)$ the strategies mentioned in Theorem whose optimality is to be proved. In order to prove Theorem, it suffices to verify the validity of the following relationships:

$$\overline{K}(x_0,y_0;\mu(\cdot),\nu_*(\cdot))+\epsilon\geq\overline{K}(x_0,y_0;u_\epsilon(\cdot),\nu_*(\cdot))$$

$$\geq\overline{K}(x_0,y_0;u_\epsilon(\cdot),\nu(\cdot))-\epsilon,\quad \mu(\cdot)\in\overline{P},\ \nu(\cdot)\in\overline{E}, \tag{9.8}$$

$$\lim_{\epsilon\to 0}\overline{K}(x_0,y_0;u_\epsilon(\cdot),\nu_*(\cdot))=\gamma(M,l). \tag{9.9}$$

The left-hand side of inequality (9.8) follows from the definition of strategy $u_\epsilon(\cdot)$ by which for any piecewise open-loop strategy $u(\cdot)\in P$

$$\overline{K}(x_0,y_0;u(\cdot),\nu_*(\cdot))+\epsilon\geq\overline{K}(x_0,y_0;u_\epsilon(\cdot),\nu_*(\cdot)).$$

Denote by $x^*(t)$ Pursuer's trajectory in situation $(u_\epsilon(\cdot),\nu_*(\cdot))$. Then

$$\overline{K}(x_0,y_0;u_\epsilon(\cdot),\nu_*(\cdot))=\sum_{i=1}^{n+1}p_i\rho(x^*(T),\overline{y}_i(M)). \tag{9.10}$$

Let R be a radius of the minimal sphere containing the set $C_E^l(M)$, i.e. $R=\gamma(M,l)$. Then $R-\epsilon/2\leq\rho(x^*(T),\overline{y}_i(M))\leq R+\epsilon/2$ for all $i=1,\dots,n+1$, since the point $x^*(T)$ belongs to the $\epsilon/2$-neighborhood of the point $y(M)$. Since $\sum_{i=1}^{n+1} p_i = 1$, $p_i \geq 0$, from (9.10) we get

$$R - \epsilon/2 \leq \overline{K}(x_0, y_0; u_\epsilon(\cdot), \nu_*(\cdot)) \leq R + \epsilon/2, \tag{9.11}$$

and this proves (9.9).

Suppose the state $x(T), y(T-l)$ have been realized in situation $(u_\epsilon(\cdot), \nu(\cdot))$ and $Q(\cdot)$ is the probability measure induced on the set $C_E^l(y(T-l))$. From the optimality of the mixed strategy $p = (p_1, \ldots, p_{n+1})$ in the game $\Gamma(M, l)$ we have

$$R = \sum_{i=1}^{n+1} p_i \rho(y(M), \overline{y}_i(M)) \geq \gamma(y(T-l), l) = Val\Gamma(y(T-l), l)$$

$$\geq \int_{C_E^l(y(T-l))} \rho(\tilde{y}[y(T-l)], y) dQ, \tag{9.12}$$

where $\tilde{y}[y(T-l)]$ is the center of the minimal sphere containing the set $C_E^l(y(T-l))$.

However, $\rho(x(T), y[y(T-l)]) \leq \epsilon/2$, therefore for $y \in C_E^l(y(T-l))$ we have

$$\rho(x(T), y) \leq \frac{\epsilon}{2} + \rho(\tilde{y}[y(T-l)], y) \leq R + \epsilon/2. \tag{9.13}$$

From inequalities (9.11)–(9.13) it follows that

$$\overline{K}(x_0, y_0; u_\epsilon(\cdot), \nu_*(\cdot)) \geq \int_{C_E^l(y(T-l))} \rho(x(T), y) dQ - \epsilon, \tag{9.14}$$

but

$$\int_{C_E^l(y(T-l))} \rho(x(T), y) dQ = \overline{K}(x_0, y_0; u_\epsilon(\cdot), \nu(\cdot)). \tag{9.15}$$

From formulas (9.14) and (9.15) we obtain the right-hand side of inequality (9.8). This completes the proof of the theorem.

For $T < l$ the solution of the game does not differ essentially from the case $T \geq l$ and Theorem holds if we consider $C_E^T(y_0)$, $\overline{C}_E^T(y_0)$, $\gamma(M, T)$, y_0 instead of $C_E^l(y_0)$, $\overline{C}_E^l(y_0)$, $\gamma(M, l)$, $y(T-l)$, respectively.

The diameter of the set $C_E^l(M)$ tends to zero as $l \to 0$, which is why the value of the auxiliary game $\Gamma(M, l)$ also tends to zero. But the value of this auxiliary game is equal to the value $V_l(x_0, y_0, T)$ of the game of pursuit with delayed information $\overline{\Gamma}(x_0, y_0, T)$ (here index l indicates the information delay). The optimal mixed strategy for Player E in $\Gamma(M, l)$ concentrating its mass on at most $n+1$ points from $C_E^l(M)$ concentrates in the limit its entire mass in one point M, i.e. it becomes a pure strategy. This agrees with the fact that the game $\Gamma(x_0, y_0, T)$ becomes the game with perfect information as $l \to 0$.

Example 7. Equations of motion are of the form

$$\dot{x} = u,\ \|u\| \leq \alpha; \quad \dot{y} = v,\ \|v\| \leq \beta, \quad \alpha > \beta,\ x, y \in R^2.$$

Suppose the time T satisfies the condition $T > \rho(x_0, y_0)/(\alpha - \beta) + l$. The reachability set $C^l_E(y_0) = \overline{C}^l_E(y_0)$ and coincides with the circle of radius βl with its center at y_0. The value of the game $\Gamma(y, l)$ is equal to the radius of the circle $C^l_E(y)$, i.e. $V(y, l) = \beta l$.

Since $V(y, l)$ is now independent of y, any point of the set $C^{T-l}_E(y_0)$ can be the center of pursuit M. An optimal strategy for Player P in the game $\Gamma(y, l)$ is the choice of point y, and an optimal strategy for Player E is mixed and is the choice of any two diametrically opposite points of the circle $C^l_E(y)$ with probabilities $(1/2, 1/2)$. Accordingly, an optimal strategy for Pursuer in the game $\overline{\Gamma}(x_0, y_0, T)$ is the linear pursuit of the point $y(t - l)$ for $l \leq t \leq T$ (the point y_0 for $0 \leq t \leq l$) until the capture of this point; moreover, it must remain in $\epsilon/2$-neighborhood of this point. An optimal strategy for Player E (the mixed piecewise open-loop behavior strategy) is the transition from the point y_0 to an arbitrary point $M \in C^{T-l}_E(y_0)$ during the time $T - l$ and then the equiprobable choice of a direction towards one of the two diametrically opposite points of the circle $C^l_E(M)$. In this case $Val\overline{\Gamma}(x_0, y_0, T) = \beta l$.

II. N-person differential games

10 Optimal control problem

10.1 Consider the optimal control problem in which the single decision-maker:

$$\max_u \left\{ \int_{t_0}^{T} g\left[s, x(s), u(s)\right] ds + q\left(x(T)\right) \right\}, \tag{10.1}$$

subject to the vector-valued differential equation:

$$\dot{x}(s) = f\left[s, x(s), u(s)\right] ds, \quad x(t_0) = x_0, \tag{10.2}$$

where $x(s) \in X \subset R^m$ denotes the state variables of game, and $u \in U$ is the control.

The functions $f[s,x,u]$, $g[s,x,u]$ and $q(x)$ are differentiable functions.

Dynamic programming and optimal control are used to identify optimal solutions for the problem (10.1)–(10.2) .

10.2. Dynamic Programming. A frequently adopted approach to dynamic optimization problems is the technique of dynamic programming. The technique was developed by [4]. The technique is given in Theorem below.

Theorem 15. *A set of controls $u^*(t) = \phi^*(t,x)$ constitutes an optimal solution to the control problem (10.1)–(10.2) if there exist continuously differentiable functions $V(t,x)$ defined on $[t_0,T] \times R^m \to R$ and satisfying the following Bellman equation:*

$$\begin{aligned} -V_t(t,x) &= \max_u \left\{ g[t,x,u] + V_x(t,x) f[t,x,u] \right\} \\ &= \left\{ g[t,x,\phi^*(t,x)] + V_x(t,x) f[t,x,\phi^*(t,x)] \right\}, \\ V(T,x) &= q(x). \end{aligned}$$

Proof. Define the maximized payoff at time t with current state x as a *value Bellman function* in the form:

$$\begin{aligned} V(t,x) &= \max_u \left[\int_t^T g(s, x(s), u(s))\, ds + q(x(T)) \right] \\ &= \int_t^T g\left[s, x^*(s), \phi^*(s, x^*(s))\right] ds + q(x^*(T)) \end{aligned}$$

satisfying the boundary condition

$$V(T, x^*(T)) = q(x^*(T)),$$

and

$$\dot{x}^*(s) = f\left[s, x^*(s), \phi^*(s, x^*(s))\right], \quad x^*(t_0) = x_0.$$

If in addition to $u^*(s) \equiv \phi^*(s,x)$, we are given another set of strategies, $u(s) \in U$, with the corresponding terminating trajectory $x(s)$, then from conditions of the theorem we have

$$g(t,x,u) + V_x(t,x) f(t,x,u) + V_t(t,x) \le 0, \text{ and}$$
$$g(t,x^*,u^*) + V_{x^*}(t,x^*) f(t,x^*,u^*) + V_t(t,x^*) = 0.$$

Integrating the above expressions from t_0 to T, we obtain

$$\int_{t_0}^{T} g(s,x(s),u(s))\,ds + V(T,x(T)) - V(t_0,x_0) \le 0, \text{ and}$$
$$\int_{t_0}^{T} g(s,x^*(s),u^*(s))\,ds + V(T,x^*(T)) - V(t_0,x_0) = 0.$$

Elimination of $V(t_0,x_0)$ yields

$$\int_{t_0}^{T} g(s,x(s),u(s))\,ds + q(x(T)) \le \int_{t_0}^{T} g(s,x^*(s),u^*(s))\,ds + q(x^*(T)),$$

from which it readily follows that u^* is the optimal strategy.

Upon substituting the optimal strategy $\phi^*(t,x)$ into (10.2) yields the dynamics of optimal state trajectory as

$$\dot{x}(s) = f[s,x(s),\phi^*(s,x(s))]\,ds, \qquad x(t_0) = x_0. \tag{10.3}$$

Let $x^*(t)$ denote the solution to (10.3) . The optimal trajectory $\{x^*(t)\}_{t=t_0}^{T}$ can be expressed as:

$$x^*(t) = x_0 + \int_{t_0}^{t} f[s,x^*(s),\psi^*(s,x^*(s))]\,ds. \tag{10.4}$$

For notational convenience, we use the terms $x^*(t)$ and x_t^* interchangeably. The value (Bellman) function $V(t,x)$ where $x = x_t^*$ can be expressed as

$$V(t,x_t^*) = \int_{t}^{T} g[s,x^*(s),\phi^*(s)]\,ds + q(x^*(T)).$$

Example 8. Consider the optimization problem:

$$\max_u \left\{ \int_0^T \exp[-rs] \left[-x(s) - cu(s)^2 \right] ds + \exp[-rT]\, qx(T) \right\} \tag{10.5}$$

subject to

$$\dot{x}(s) = a - u(s)(x(s))^{1/2}, \qquad x(0) = x_0, \qquad u(s) \ge 0, \tag{10.6}$$

where a, c, x_0 are positive parameters.

Invoking the above theorem we have

$$-V_t(t,x) = \max_u \left\{ \left[-x - cu^2\right] \exp\left[-rt\right] + V_x(t,x) \left[a - ux^{1/2}\right] \right\},$$

$$\text{and} \tag{10.7}$$

$$V(T,x) = \exp\left[-rT\right] qx.$$

Performing the indicated maximization in (10.7) yields

$$\phi(t,x) = \frac{-V_x(t,x)\, x^{1/2}}{2c} \exp\left[rt\right].$$

Substituting $\phi(t,x)$ into (10.7) and upon solving (10.7), one obtains

$$V(t,x) = \exp\left[-rt\right]\left[A(t)\,x + B(t)\right],$$

where $A(t)$ and $B(t)$ satisfy:

$$\begin{aligned} \dot{A}(t) &= rA(t) - \frac{A(t)^2}{4c} + 1, \\ \dot{B}(t) &= rB(t) - aA(t), \\ A(T) &= q \ \text{ and } \ B(T) = 0. \end{aligned}$$

The optimal control can be solved explicitly as

$$\phi(t,x) = \frac{-A(t)\, x^{1/2}}{2c} \exp\left[rt\right].$$

Now, consider the infinite-horizon dynamic optimization problem with a constant discount rate

$$\max_u \left\{ \int_{t_0}^{\infty} g\left[x(s), u(s)\right] \exp\left[-r(s - t_0)\right] ds \right\}, \tag{10.8}$$

subject to the vector-valued differential equation

$$\dot{x}(s) = f\left[x(s), u(s)\right] ds, \qquad x(t_0) = x_0. \tag{10.9}$$

Since s does not appear in $g\left[x(s), u(s)\right]$ and the state dynamics explicitly, the problem (10.8)–(10.9) is an autonomous problem.

Consider the alternative problem:

$$\max_u \int_t^{\infty} g\left[x(s), u(s)\right] \exp\left[-r(s - t)\right] ds, \tag{10.10}$$

subject to

$$\dot{x}(s) = f\left[x(s), u(s)\right], \qquad x(t) = x. \tag{10.11}$$

The infinite-horizon autonomous problem (10.10)–(10.11) is independent of the choice of t and dependent only upon the state at the starting time, that is x.

Define the value function to the problem (10.8)–(10.9) by

$$V(t,x) = \max_u \left\{ \int_t^\infty g\left[x(s), u(s)\right] \exp\left[-r(s-t_0)\right] ds \middle| x(t) = x = x_t^* \right\},$$

where x_t^* is the state at time t along the optimal trajectory. Moreover, we can write

$$V(t,x) =$$
$$\exp\left[-r(t-t_0)\right] \max_u \left\{ \int_t^\infty g\left[x(s), u(s)\right] \exp\left[-r(s-t)\right] ds \middle| x(t) = x = x_t^* \right\}.$$

Since the problem

$$\max_u \left\{ \int_t^\infty g\left[x(s), u(s)\right] \exp\left[-r(s-t)\right] ds \middle| x(t) = x = x_t^* \right\}$$

depends on the current state x only, we can write

$$W(x) = \max_u \left\{ \int_t^\infty g\left[x(s), u(s)\right] \exp\left[-r(s-t)\right] ds \middle| x(t) = x = x_t^* \right\}.$$

It follows that

$$\begin{gathered} V(t,x) = \exp\left[-r(t-t_0)\right] W(x), \\ V_t(t,x) = -r \exp\left[-r(t-t_0)\right] W(x), \text{ and} \\ V_x(t,x) = -r \exp\left[-r(t-t_0)\right] W_x(x). \end{gathered} \tag{10.12}$$

Substituting the results from (10.12) into Bellman equation from the theorem yields

$$rW(x) = \max_u \left\{ g\left[x,u\right] + W_x(x) f\left[x,u\right] \right\}. \tag{10.13}$$

Since time is not explicitly involved (10.13), the derived control u will be a function of x only. Hence one can obtain the theorem.

Theorem 16. *A set of controls $u = \phi^*(x)$ constitutes an optimal solution to the infinite-horizon control problem (10.10)–(10.11) if there exists continuously differentiable function $W(x)$ defined on $R^m \to R$ which satisfies the following equation:*

$$\begin{aligned} rW(x) &= \max_u \left\{ g\left[x,u\right] + W_x(x) f\left[x,u\right] \right\} \\ &= \left\{ g\left[x, \phi^*(x)\right] + W_x(x) f\left[x, \phi^*(x)\right] \right\}. \end{aligned}$$

Substituting the optimal control into (10.9) yields the dynamics of the optimal state path as

$$\dot{x}(s) = f\left[x(s), \phi^*(x(s))\right] ds, \qquad x(t_0) = x_0.$$

Solving the above dynamics yields the optimal state trajectory $\{x^*(t)\}_{t\geq t_o}$ as

$$x^*(t) = x_0 + \int_{t_0}^{t} f\left[x^*(s), \psi^*(x^*(s))\right] ds, \quad \text{for } t \geq t_0.$$

We denote term $x^*(t)$ by x_t^*. The optimal control to the infinite-horizon problem (10.8)–(10.9) can be expressed as $\psi^*(x_t^*)$ in the time interval $[t_0, \infty)$.

Example 9. Consider the infinite-horizon dynamic optimization problem:

$$\max_u \int_0^\infty \exp\left[-rs\right]\left[-x(s) - cu(s)^2\right] ds \tag{10.14}$$

subject to dynamics (10.6).

Invoking previous theorem we have

$$rW(x) = \max_u \left\{\left[-x - cu^2\right] + W_x(x)\left[a - ux^{1/2}\right]\right\}. \tag{10.15}$$

Performing the indicated maximization in (10.15) yields

$$\phi^*(x) = \frac{-V_x(x)\, x^{1/2}}{2c}.$$

Substituting $\phi(x)$ into (10.15) and upon solving (10.15), one obtains

$$V(t, x) = \exp\left[-rt\right]\left[Ax + B\right],$$

where A and B satisfy

$$0 = rA - \frac{A^2}{4c} + 1 \;\text{ and }\; B = \frac{-a}{r}A.$$

Solving A to be $2c\left[r \pm \left(r^2 + c^{-1}\right)^{1/2}\right]$. For a maximum, the negative root of A holds. The optimal control can be obtained as

$$\phi^*(x) = \frac{-Ax^{1/2}}{2c}.$$

Substituting $\phi^*(x) = -Ax^{1/2}/(2c)$ into (10.6) yields the dynamics of the optimal trajectory as

$$\dot{x}(s) = a + \frac{A}{2c}(x(s)), \qquad x(0) = x_0.$$

Upon the above dynamical equation yields the optimal trajectory $\{x^*(t)\}_{t\geq t_0}$ as

$$x^*(t) = \left[x_0 + \frac{2ac}{A}\right]\exp\left(\frac{A}{2c}t\right) - \frac{2ac}{A} = x_t^*, \qquad \text{for } t \geq t_0.$$

The optimal control of problem (10.14)–(10.15) is then

$$\phi^*(x_t^*) = \frac{-A(x_t^*)^{1/2}}{2c}.$$

10.3. Optimal Control. The maximum principle of optimal control was developed by Pontryagin (details in [5]). Consider again the dynamic optimization problem (10.1)–(10.2).

Theorem 17. *Pontryagin's Maximum Principle. A set of controls* $u^*(s) = \zeta^*(s, x_0)$ *provides an optimal solution to control problem (10.1)–(10.2), and* $\{x^*(s), t_0 \le s \le T\}$ *is the corresponding state trajectory, if there exist costate functions* $\Lambda(s) : [t_0, T] \to R^m$ *such that the following relations are satisfied:*

$$\zeta^*(s, x_0) \equiv u^*(s) = \arg\max_u \{g[s, x^*(s), u(s)] + \Lambda(s) f[s, x^*(s), u(s)]\},$$
$$\dot{x}^*(s) = f[s, x^*(s), u^*(s)], \quad x^*(t_0) = x_0,$$
$$\dot{\Lambda}(s) = -\frac{\partial}{\partial x}\{g[s, x^*(s), u^*(s)] + \Lambda(s) f[s, x^*(s), u^*(s)]\},$$
$$\Lambda(T) = \frac{\partial}{\partial x^*} q(x^*(T)).$$

Proof. First define the function (Hamiltonian)

$$H(t, x, u) = g(t, x, u) + V_x(t, x) f(t, x, u).$$

The theorem from 10.2 gives us

$$-V_t(t, x) = \max_u H(t, x, u).$$

This yields the first condition of the above theorem. Using u^* to denote the payoff maximizing control, we obtain

$$H(t, x, u^*) + V_t(t, x) \equiv 0,$$

which is an identity in x. Differentiating this identity partially with respect to x yields

$$V_{tx}(t, x) + g_x(t, x, u^*) + V_x(t, x) f_x(t, x, u^*) + V_{xx}(t, x) f(t, x, u^*)$$
$$+ [g_u(t, x, u^*) + V_x(t, x) f_u(t, x, u^*)] \frac{\partial u^*}{\partial x} = 0.$$

If u^* is an interior point, then $[g_u(t, x, u^*) + V_x(t, x) f_u(t, x, u^*)] = 0$ according to the condition $-V_t(t, x) = \max_u H(t, x, u)$. If u^* is not an interior point, then it can be shown that

$$[g_u(t, x, u^*) + V_x(t, x) f_u(t, x, u^*)] \frac{\partial u^*}{\partial x} = 0$$

(because of optimality, $[g_u(t, x, u^*) + V_x(t, x) f_u(t, x, u^*)]$ and $\partial u^*/\partial x$ are orthogonal; and for specific problems we may have $\partial u^*/\partial x = 0$). Moreover, the expression $V_{tx}(t,x) + V_{xx}(t,x) f(t,x,u^*) \equiv V_{tx}(t,x) + V_{xx}(t,x)\dot{x}$ can be written as $[dV_x(t,x)](dt)^{-1}$. Hence, we obtain:

$$\frac{dV_x(t,x)}{dt} + g_x(t,x,u^*) + V_x(t,x) f_x(t,x,u^*) = 0.$$

By introducing the *costate vector*, $\Lambda(t) = V_{x^*}(t,x^*)$, where x^* denotes the state trajectory corresponding to u^*, we arrive at

$$\frac{dV_x(t,x^*)}{dt} = \dot{\Lambda}(s) = -\frac{\partial}{\partial x}\left\{g\left[s,x^*(s),u^*(s)\right] + \Lambda(s) f\left[s,x^*(s),u^*(s)\right]\right\}.$$

Finally, the boundary condition for $\Lambda(t)$ is determined from the terminal condition of optimal control in Theorem from 10.2 as

$$\Lambda(T) = \frac{\partial V(T,x^*)}{\partial x} = \frac{\partial q(x^*)}{\partial x}.$$

Hence theorem follows.

Example 10. Consider the problem in Example 8. Invoking theorem, we first solve the control $u(s)$ that satisfies

$$\arg\max_u \left\{\left[-x^*(s) - cu(s)^2\right]\exp[-rs] + \Lambda(s)\left[a - u(s)x^*(s)^{1/2}\right]\right\}.$$

Performing the indicated maximization:

$$u^*(s) = \frac{-\Lambda(s)x^*(s)^{1/2}}{2c}\exp[rs]. \tag{10.16}$$

We also obtain

$$\dot{\Lambda}(s) = \exp[-rs] + \frac{1}{2}\Lambda(s)u^*(s)x^*(s)^{-1/2}. \tag{10.17}$$

Substituting $u^*(s)$ from (10.16) into (10.6) and (10.17) yields a pair of differential equations:

$$\dot{x}^*(s) = a + \frac{1}{2c}\Lambda(s)(x^*(s))\exp[rs], \tag{10.18}$$

$$\dot{\Lambda}(s) = \exp[-rs] + \frac{1}{4c}\Lambda(s)^2\exp[rs],$$

with boundary conditions:

$$x^*(0) = x_0 \text{ and } \Lambda(T) = \exp[-rT]q.$$

Solving (10.18) yields

$$\Lambda(s) = 2c\left(\theta_1 - \theta_2\frac{q-2c\theta_1}{q-2c\theta_2}\exp\left[\frac{\theta_1-\theta_2}{2}(T-s)\right]\right)\exp(-rs)$$

$$\div \left(1 - \frac{q-2c\theta_1}{q-2c\theta_2} \exp\left[\frac{\theta_1-\theta_2}{2}(T-s)\right]\right),$$

and

$$x^*(s) = \varpi(0,s)\left[x_0 + \int_0^s \varpi^{-1}(0,t)\, a\; dt\right], \quad \text{for} \;\; s \in [0,T],$$

where

$$\theta_1 = r - \sqrt{r^2 + \frac{1}{c}} \;\; \text{and} \;\; \theta_2 = r + \sqrt{r^2 + \frac{1}{c}},$$

$$\varpi(0,s) = \exp\left[\int_0^s H(\tau)\, d\tau\right],$$

and

$$H(\tau) = \left(\theta_1 - \theta_2 \frac{q-2c\theta_1}{q-2c\theta_2} \exp\left[\frac{\theta_1-\theta_2}{2}(T-\tau)\right]\right)$$

$$\div \left(1 - \frac{q-2c\theta_1}{q-2c\theta_2} \exp\left[\frac{\theta_1-\theta_2}{2}(T-\tau)\right]\right).$$

Upon substituting $\Lambda(s)$ and $x^*(s)$ into (10.16) yields $u^*(s) = \zeta^*(s, x_0)$ which is a function of s and x_0.

Consider the infinite-horizon dynamic optimization problem (10.8)–(10.9). The Hamiltonian function can be expressed as

$$H(t,x,u) = g(x,u)\exp[-r(t-t_0)] + \Lambda(t) f(x,u).$$

Define $\lambda(t) = \Lambda(t)\exp[r(t-t_0)]$ and the current value Hamiltonian

$$\hat{H}(t,x,u) = H(t,x,u)\exp[r(t-t_0)] \tag{10.19}$$

$$= g(x,u) + \lambda(t) f(x,u).$$

Using the previous theorem and (10.19) we get the maximum principle for the game (10.10)–(10.11).

Theorem 18. *A set of controls $u^*(s) = \zeta^*(s, x_t)$ provides an optimal solution to the infinite-horizon control problem (10.10)–(10.11), and $\{x^*(s),\; s \geq t\}$ is the corresponding state trajectory, if there exist costate functions $\lambda(s) : [t, \infty) \to R^m$ such that the following relations are satisfied*

$$\zeta^*(s, x_t) \equiv u^*(s) = \arg\max_u \{g[x^*(s), u(s)] + \lambda(s) f[x^*(s), u(s)]\},$$

$$\dot{x}^*(s) = f[x^*(s), u^*(s)], \qquad x^*(t) = x_t,$$

$$\dot{\lambda}(s) = r\lambda(s) - \frac{\partial}{\partial x}\{g[x^*(s), u^*(s)] + \lambda(s) f[x^*(s), u^*(s)]\}.$$

Example 11 Consider the infinite-horizon problem in Example 9.
Invoking the previous theorem we have

$$\zeta^*(s, x_t) \equiv u^*(s) =$$

$$\arg\max_u \left\{ \left[-x^*(s) - cu(s)^2 \right] + \lambda(s) \left[a - u(s) x^*(s)^{1/2} \right] \right\},$$
$$\dot{x}^*(s) = a - u^*(s) \left(x^*(s) \right)^{1/2}, \qquad x^*(t) = x_t, \tag{10.20}$$
$$\dot{\lambda}(s) = r\lambda(s) + \left[1 + \frac{1}{2} \lambda(s) u^*(s) x^*(s)^{-1/2} \right].$$

Performing the indicated maximization yields

$$u^*(s) = \frac{-\lambda(s) x^*(s)^{1/2}}{2c}.$$

Substituting $u^*(s)$ into (10.20), one obtains

$$\dot{x}^*(s) = a + \frac{\lambda(s)}{2c} u^*(s) x^*(s), \quad x^*(t) = x_t, \tag{10.21}$$
$$\dot{\lambda}(s) = r\lambda(s) + \left[1 - \frac{1}{4c} \lambda(s)^2 \right].$$

Solving (10.21) in a manner similar to that in Example 10 yields the solutions of $x^*(s)$ and $\lambda(s)$. Upon substituting them into $u^*(s)$ gives the optimal control of the problem.

11 Differential games and their solution concepts

11.1. One particularly complex but fruitful branch of game theory is dynamic or differential games, which investigates interactive decision making over time under different assumptions regarding pre-commitment (of actions), information, and uncertainty. The origin of differential games traces back to the late 1940s. Rufus Isaacs (whose work was published in 1965) formulated missile versus enemy aircraft pursuit schemes in terms of descriptive and navigation variables (state and control), and established a fundamental principle: the tenet of transition. The seminal contributions of Isaacs together with the classic research of Bellman on dynamic programming and Pontryagin et al. on optimal control laid the foundations of deterministic differential zero-sum games.

Differential games or continuous-time infinite dynamic games study a class of decision problems, under which the evolution of the state is described by a differential equation and the players act throughout a time interval. In this section we follow [6], [7].

In particular, in the general n-person differential game, Player i seeks to:

$$\max_{u_i} \int_{t_0}^{T} g^i \left[s, x(s), u_1(s), u_2(s), \ldots, u_n(s) \right] ds + q^i \left(x(T) \right),$$
$$\text{for } \; i \in N = \{1, 2, \ldots, n\}, \tag{11.1}$$

subject to the deterministic dynamics

$$\dot{x}(s) = f[s, x(s), u_1(s), u_2(s), \ldots, u_n(s)], \qquad x(t_0) = x_0, \tag{11.2}$$

where $x(s) \in X \subset R^m$ denotes the state variables of game, and $u_i \in U^i$ is the control of Player i, for $i \in N$.

The functions $f[s, x, u_1, u_2, \ldots, u_n]$, $g^i[s, \cdot, u_1, u_2, \ldots, u_n]$ and $q^i(\cdot)$, for $i \in N$, and $s \in [t_0, T]$ are differentiable functions.

A set-valued function $\eta^i(\cdot)$ defined for each $i \in N$ as

$$\eta^i(s) = \left\{x(t), \quad t_0 \le t \le \epsilon_s^i\right\}, \qquad t_0 \le \epsilon_s^i \le s,$$

where ϵ_s^i is nondecreasing in s, and $\eta^i(s)$ determines the state information gained and recalled by Player i at time $s \in [t_0, T]$. Specification of $\eta^i(\cdot)$(in fact, ϵ_s^i in this formulation) characterizes the *information structure* of Player i and the collection (over $i \in N$) of these information structures is the *information structure* of the game.

A sigma-field N_s^i in S_0 generated for each $i \in N$ by the cylinder sets $\{x \in S_0, x(t) \in B\}$ where B is a Borel set in S^0 and $0 \le t \le \epsilon_s$. N_s^i, $s \ge t_0$, is called the *information field* of Player i.

A pre-specified class Γ^i of mappings $\upsilon_i : [t_0, T] \times S_0 \to S^i$, with the property that $u_i(s) = \upsilon_i(s, x)$ is n_s^i-measurable (i.e. it is adapted to the information field N_s^i). U^i is the strategy space of Player i and each of its elements υ_i is a permissible strategy for Player i.

Definition 14. *A set of strategies $\{\upsilon_1^*(s), \upsilon_2^*(s), \ldots, \upsilon_n^*(s)\}$ is said to constitute a non-cooperative Nash equilibria solution for the n-person differential game (11.1) –(11.2) , if the following inequalities are satisfied for all $\upsilon_i(s) \in U^i$, $i \in N$:*

$$\int_{t_0}^{T} g^1[s, x^*(s), \upsilon_1^*(s), \upsilon_2^*(s), \ldots, \upsilon_n^*(s)]\, ds + q^1(x^*(T)) \ge$$
$$\int_{t_0}^{T} g^1\left[s, x^{[1]}(s), \upsilon_1(s), \upsilon_2^*(s), \ldots, \upsilon_n^*(s)\right] ds + q^1\left(x^{[1]}(T)\right),$$
$$\int_{t_0}^{T} g^2[s, x^*(s), \upsilon_1^*(s), \upsilon_2^*(s), \ldots, \upsilon_n^*(s)]\, ds + q^2(x^*(T)) \ge$$
$$\int_{t_0}^{T} g^2\left[s, x^{[2]}(s), \upsilon_1^*(s), \upsilon_2(s), \upsilon_3^*(s), \ldots, \upsilon_n^*(s)\right] ds + q^2\left(x^{[2]}(T)\right),$$
$$\vdots \qquad\qquad \vdots$$
$$\vdots \qquad\qquad \vdots$$
$$\int_{t_0}^{T} g^n[s, x^*(s), \upsilon_1^*(s), \upsilon_2^*(s), \ldots, \upsilon_n^*(s)]\, ds + q^n(x^*(T)) \ge$$
$$\int_{t_0}^{T} g^n\left[s, x^{[n]}(s), \upsilon_1^*(s), \upsilon_2^*(s), \ldots, \upsilon_{n-1}^*(s), \upsilon_n(s)\right] ds + q^n\left(x^{[n]}(T)\right);$$

where on the time interval $[t_0, T]$*:*

$$\begin{aligned}
&\dot{x}^*(s) = f\left[s, x^*(s), \upsilon_1^*(s), \upsilon_2^*(s), \ldots, \upsilon_n^*(s)\right], \quad x^*(t_0) = x_0,\\
&\dot{x}^{[1]}(s) = f\left[s, x^{[1]}(s), \upsilon_1(s), \upsilon_2^*(s), \ldots, \upsilon_n^*(s)\right], \quad x^{[1]}(t_0) = x_0,\\
&\dot{x}^{[2]}(s) = f\left[s, x^{[2]}(s), \upsilon_1^*(s), \upsilon_2(s), \upsilon_3^*(s), \ldots, \upsilon_n^*(s)\right], \quad x^{[2]}(t_0) = x_0,\\
&\qquad\vdots \qquad\qquad \vdots\\
&\qquad\vdots \qquad\qquad \vdots\\
&\dot{x}^{[n]}(s) = f\left[s, x^{[n]}(s), \upsilon_1^*(s), \upsilon_2^*(s), \ldots, \upsilon_{n-1}^*(s), \upsilon_n(s)\right], \quad x^{[n]}(t_0) = x_0.
\end{aligned}$$

The set of strategies $\{\upsilon_1^*(s), \upsilon_2^*(s), \ldots, \upsilon_n^*(s)\}$ *is known as a* Nash equilibria *of the game.*

11.2. Open-loop Nash Equilibria. If the players choose to commit their strategies from the outset, the players' information structure can be seen as an *open-loop* pattern in which $\eta^i(s) = \{x_0\}$, $s \in [t_0, T]$. Their strategies become functions of the initial state x_0 and time s, and can be expressed as $\{u_i(s) = \vartheta_i(s, x_0)\,, \text{ for } i \in N\}$. In particular, an open-loop Nash equilibria for the game (11.1) and (11.2) is characterized as:

Theorem 19. *A set of strategies* $\{u_i^*(s) = \zeta_i^*(s, x_0)\,, \textit{ for } i \in N\}$ *provides an open-loop Nash equilibria solution to the game (11.1)–(11.2) , and* $\{x^*(s)\,, t_0 \le s \le T\}$ *is the corresponding state trajectory, if there exist m costate functions* $\Lambda^i(s) : [t_0, T] \to R^m$*, for* $i \in N$*, such that the following relations are satisfied:*

$$\begin{aligned}
&\zeta_i^*(s, x_0) \equiv u_i^*(s) =\\
&\arg\max_{u_i \in U^i} \left\{ g^i\left[s, x^*(s), u_1^*(s), u_2^*(s), \ldots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \ldots, u_n^*(s)\right] \right.\\
&\left. + \Lambda^i(s) f\left[s, x^*(s), u_1^*(s), u_2^*(s), \ldots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \ldots, u_n^*(s)\right] \right\},\\
&\quad \dot{x}^*(s) = f\left[s, x^*(s), u_1^*(s), u_2^*(s), \ldots, u_n^*(s)\right], \quad x^*(t_0) = x_0,\\
&\quad \dot{\Lambda}^i(s) = -\frac{\partial}{\partial x^*}\left\{ g^i\left[s, x^*(s), u_1^*(s), u_2^*(s), \ldots, u_n^*(s)\right] \right.\\
&\qquad\qquad \left. + \Lambda^i(s) f\left[s, x^*(s), u_1^*(s), u_2^*(s), \ldots, u_n^*(s)\right] \right\},\\
&\quad \Lambda^i(T) = \frac{\partial}{\partial x^*} q^i\left(x^*(T)\right), \quad \textit{for } i \in N.
\end{aligned}$$

Proof. Consider the i^{th} equality in conditions of the theorem, which states that $\upsilon_i^*(s) = u_i^*(s) = \zeta_i^*(s, x_0)$ maximizes

$$\begin{aligned}
&\int_{t_0}^{T} g^i\left[s, x(s), u_1^*(s), u_2^*(s), \ldots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \ldots, u_n^*(s)\right] ds\\
&\qquad + q^i\left(x(T)\right),
\end{aligned}$$

over the choice of $\upsilon_i(s) \in U^i$ subject to the state dynamics

$$\dot{x}(s) = f\left[s, x(s), u_1^*(s), u_2^*(s), \ldots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \ldots, u_n^*(s)\right],$$
$$x(t_0) = x_0, \quad \text{for} \quad i \in N.$$

This is standard optimal control problem for Player i, since $u_j^*(s)$, for $j \in N$ and $j \neq i$, are open-loop controls and hence do not depend on $u_i^*(s)$. These results then follow directly from the maximum principle of Pontryagin as stated in 10.3.

11.3. Closed-loop Nash Equilibria. Under the memoryless perfect state information, the players' information structures follow the pattern $\eta^i(s) = \{x_0, x(s)\}$, $s \in [t_0, T]$. The players' strategies become functions of the initial state x_0, current state $x(s)$ and current time s, and can be expressed as $\{u_i(s) = \vartheta_i(s, x, x_0), \text{ for } i \in N\}$. The following theorem provides a set of necessary conditions for any closed-loop no-memory Nash equilibria solution to satisfy.

Theorem 20. *A set of strategies $\{u_i(s) = \vartheta_i(s, x, x_0), \text{ for } i \in N\}$ provides a closed-loop no memory Nash equilibria solution to the game (11.1)–(11.2), and $\{x^*(s), t_0 \leq s \leq T\}$ is the corresponding state trajectory, if there exist N costate functions $\Lambda^i(s) : [t_0, T] \to R^m$, for $i \in N$, such that the following relations are satisfied:*

$$\vartheta_i^*(s, x^*, x_0) \equiv u_i^*(s) =$$
$$\arg\max_{u_i \in U^i} \left\{ g^i\left[s, x^*(s), u_1^*(s), u_2^*(s), \ldots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \ldots, u_n^*(s)\right] \right.$$
$$\left. + \Lambda^i(s) f\left[s, x^*(s), u_1^*(s), u_2^*(s), \ldots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \ldots, u_n^*(s)\right] \right\},$$
$$\dot{x}^*(s) = f\left[s, x^*(s), u_1^*(s), u_2^*(s), \ldots, u_n^*(s)\right], \quad x^*(t_0) = x_0,$$
$$\dot{\Lambda}^i(s) = -\frac{\partial}{\partial x^*} \left\{ g^i\left[s, x^*(s), \vartheta_1^*(s, x^*, x_0), \vartheta_2^*(s, x^*, x_0), \ldots \right.\right.$$
$$\left. \ldots, \vartheta_{i-1}^*(s, x^*, x_0), u_i^*(s), \vartheta_{i+1}^*(s, x^*, x_0), \ldots, \vartheta_n^*(s, x^*, x_0)\right]$$
$$+ \Lambda^i(s) f\left[s, x^*(s), \vartheta_1^*(s, x^*, x_0), \vartheta_2^*(s, x^*, x_0), \ldots \right.$$
$$\left.\left. \ldots, \vartheta_{i-1}^*(s, x^*, x_0), u_i^*(s), \vartheta_{i+1}^*(s, x^*, x_0), \ldots, \vartheta_n^*(s, x^*, x_0)\right] \right\},$$
$$\Lambda^i(T) = \frac{\partial}{\partial x^*} q^i(x^*(T)); \quad \textit{for} \quad i \in N.$$

Proof. Consider the i^{th} equality in conditions of the theorem, which fixed all players' strategies (except those of the i^{th} player) at $u_j^*(s) = \vartheta_j^*(s, x^*, x_0)$, for $j \neq i$ and $j \in N$, and constitutes an optimal control problem for Player i. Therefore, the above conditions follow from the maximum principle of Pontryagin, and Player i maximizes

$$\int_{t_0}^{T} g^i \left[s, x(s), u_1^*(s), u_2^*(s), \ldots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \ldots, u_n^*(s)\right] ds$$
$$+q^i(x(T)),$$

over the choice of $\upsilon_i(s) \in U^i$ subject to the state dynamics:

$$\dot{x}(s) = f\left[s, x(s), u_1^*(s), u_2^*(s), \ldots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \ldots, u_n^*(s)\right],$$
$$x(t_0) = x_0, \qquad \text{for } i \in N.$$

Note that the partial derivative with respect to x in the costate equations receives contributions from dependence of the other $n-1$ players' strategies on the current value of x. This is a feature absent from the costate equations in 11.1. The set of equations in Theorem (see 11.2) in general admits of an uncountable number of solutions, which correspond to "*informationally nonunique*" Nash equilibria solutions of differential games under memoryless perfect state information pattern.

11.4. Feedback Nash Equilibria. To eliminate information nonuniqueness in the derivation of Nash equilibria, one can constrain the Nash solution further by requiring it to satisfy the feedback Nash equilibrium property. In particular, the players' information structures follow either a *closed-loop perfect state* (CLPS) pattern in which $\eta^i(s) = \{x(s), t_0 \le t \le s\}$ or a *memoryless perfect state* (MPS) pattern in which $\eta^i(s) = \{x_0, x(s)\}$. Moreover, we require the following feedback Nash equilibrium condition to be satisfied.

Definition 15. *For the n-person differential game (10.1) – (10.2) with MPS or CLPS information, an n-tuple of strategies $\left\{u_i^*(s) = \phi_i^*(s,x) \in U^i,\right.$ for $\left. i \in N\right\}$ constitutes a* feedback Nash equilibrium solution *if there exist functionals $V^i(t,x)$ defined on $[t_0, T] \times R^m$ and satisfying the following relations for each $i \in N$:*

$$V^i(T,x) = q^i(x),$$
$$V^i(t,x) =$$
$$\int_t^T g^i\left[s, x^*(s), \phi_1^*(s,\eta_s), \phi_2^*(s,\eta_s), \ldots, \phi_n^*(s,\eta_s)\right] ds + q^i(x^*(T)) \ge$$
$$\int_t^T g^i \Big[s, x^{[i]}(s), \phi_1^*(s,\eta_s), \phi_2^*(s,\eta_s), \ldots$$
$$\ldots, \phi_{i-1}^*(s,\eta_s), \phi_i(s,\eta_s), \phi_{i+1}^*(s,\eta_s), \ldots, \phi_n^*(s,\eta_s)\Big] ds + q^i\left(x^{[i]}(T)\right),$$
$$\forall \phi_i(\cdot,\cdot) \in \Gamma^i, x \in R^n$$

where on the interval $[t_0, T]$,

$$\dot{x}^{[i]}(s) = f\Big[s, x^{[i]}(s), \phi_1^*(s,\eta_s), \phi_2^*(s,\eta_s), \ldots$$
$$\ldots, \phi_{i-1}^*(s,\eta_s), \phi_i(s,\eta_s), \phi_{i+1}^*(s,\eta_s), \ldots, \phi_n^*(s,\eta_s)\Big], \quad x^{[1]}(t) = x;$$
$$\dot{x}^*(s) = f[s, x^*(s), \phi_1^*(s,\eta_s), \phi_2^*(s,\eta_s), \ldots, \phi_n^*(s,\eta_s)], \quad x(s) = x$$

and η_s *stands for either the data set* $\{x(s), x_0\}$ *or* $\{x(\tau), \tau \le s\}$*, depending on whether the information pattern is MPS or CLPS.*

One salient feature of the concept introduced above is that if an n-tuple $\{\phi_i^*;\, i \in N\}$ provides a feedback Nash equilibrium solution (FNES) to an N-person differential game with duration $[t_0, T]$, its restriction to the time interval $[t, T]$ provides an FNES to the same differential game defined on the shorter time interval $[t, T]$, with the initial state taken as $x(t)$, and this being so for all $t_0 \le t \le T$. An immediate consequence of this observation is that, under either MPS or CLPS information pattern, feedback Nash equilibrium strategies will depend only on the time variable and the current value of the state, but not on memory (including the initial state x_0). Therefore the players' strategies can be expressed as $\{u_i(s) = \phi_i(s,x),\ \text{for}\ i \in N\}$. The following theorem provides a set of necessary conditions characterizing a feedback Nash equilibrium solution for the game (11.1) and (11.2) is characterized as follows

Theorem 21. *An n-tuple of strategies* $\left\{u_i^*(s) = \phi_i^*(t,x) \in U^i,\ \text{for}\ i \in N\right\}$ *provides a feedback Nash equilibrium solution to the game (11.1)–(11.2) if there exist continuously differentiable functions* $V^i(t,x) : [t_0, T] \times R^m \to R,\ i \in N$*, satisfying the following set of partial differential equations:*

$$\begin{aligned}
-V_t^i(t,x) = \max_{u_i} \big\{ & g^i[t, x, \phi_1^*(t,x), \phi_2^*(t,x), \ldots \\
& \quad \ldots, \phi_{i-1}^*(t,x), u_i(t,x), \phi_{i+1}^*(t,x), \ldots, \phi_n^*(t,x)] \\
& + V_x^i(t,x) f[t, x, \phi_1^*(t,x), \phi_2^*(t,x), \ldots \\
& \quad \ldots, \phi_{i-1}^*(t,x), u_i(t,x), \phi_{i+1}^*(t,x), \ldots, \phi_n^*(t,x)]\big\} \\
= \big\{ & g^i[t, x, \phi_1^*(t,x), \phi_2^*(t,x), \ldots, \phi_n^*(t,x)] \\
& + V_x^i(t,x) f[t, x, \phi_1^*(t,x), \phi_2^*(t,x), \ldots, \phi_n^*(t,x)]\big\}, \\
V^i(T,x) = q^i(x), & \quad i \in N.
\end{aligned}$$

Proof. By (10.2), $V^i(t,x)$ is the value function associated with the optimal control problem of Player i, $i \in N$. Together with the i^{th} expression in Definition, the conditions in the Theorem imply a Nash equilibrium.

Consider the two-person zero-sum version of the game (11.1)–(11.2) in which the payoff of Player 1 is the negative of that of Player 2. Under either MPS or CLPS information pattern, a feedback saddle-point is characterized as follows.

Theorem 22. *A pair of strategies* $\{\phi_i^*(t,x); i=1,2\}$ *provides a feedback saddle-point solution to the zero-sum version of the game (11.1)–(11.2) if there exists a function* $V:[t_0,T]\times R^m \to R$ *satisfying the partial differential equation*

$$\begin{aligned}
-V_t(t,x) &= \min_{u_1\in S^1}\max_{u_2\in S^2}\{g[t,x,u_1(t),u_2(t)]+V_x f[t,x,u_1(t),u_2(t)]\} \\
&= \max_{u_2\in S^2}\min_{u_1\in S^1}\{g[t,x,u_1(t),u_2(t)]+V_x f[t,x,u_1(t),u_2(t)]\} \\
&= \{g[t,x,\phi_1^*(t,x),\phi_2^*(t,x)]+V_x f[t,x,\phi_1^*(t,x),\phi_2^*(t,x)]\}, \\
V(T,x) &= q(x).
\end{aligned}$$

Proof. This result follows as a special case of the previous Theorem by taking $n=2$, $g^1(\cdot)=-g^2(\cdot)\equiv g(\cdot)$, and $q^1(\cdot)=-q^2(\cdot)\equiv q(\cdot)$, in which case $V^1 = -V^2 \equiv V$ and existence of a saddle point is equivalent to interchangeability of the min max operations.

The partial differential equation in this Theorem was first obtained by Isaacs (see [1]), and is therefore called the Isaacs equation.

12 Application of differential games in economics

In this section we consider application of differential games in competitive advertising.

12.1. Open-loop solution in competitive advertising. Consider the competitive dynamic advertising game in [8]. There are two firms in a market and the profit of firm 1 and that of 2 are respectively

$$\int_0^T \left[q_1 x(s) - \frac{c_1}{2}u_1(s)^2\right]\exp(-rs)\,ds + \exp(-rT)\,S_1 x(T)$$

$$\text{and} \tag{12.1}$$

$$\int_0^T \left[q_2(1-x(s)) - \frac{c_2}{2}u_2(s)^2\right]\exp(-rs)\,ds + \exp(-rT)\,S_2[1-x(T)],$$

where r, q_i, c_i, S_i, for $i\in\{1,2\}$, are positive constants, $x(s)$ is the market share of firm 1 at time s, $[1-x(s)]$ is that of firm 2's, $u_i(s)$ is advertising rate for firm $i\in\{1,2\}$.

It is assumed that market potential is constant over time. The only marketing instrument used by the firms is advertising. Advertising has diminishing returns, since there are increasing marginal costs of advertising as reflected through the quadratic cost function. The dynamics of firm 1's market share is governed by

$$\dot{x}(s) = u_1(s)[1-x(s)]^{1/2} - u_2(s)\,x(s)^{1/2}, \qquad x(0)=x_0. \tag{12.2}$$

There are saturation effects, since u_i operates only on the buyer market of the competing firm j.

Consider that the firms would like to seek an open-loop solution. Using open-loop strategies requires the firms to determine their action paths at the outset. This is realistic only if there are restrictive commitments concerning advertising. Invoking 11.2, an open-loop solution to the game (12.1) – (12.2) has to satisfy the following conditions

$$\begin{aligned}
u_1^*(s) = \arg\max_{u_1} \Big\{ &\left[q_1 x^*(s) - \frac{c_1}{2} u_1(s)^2 \right] \exp(-rs) \\
&+ \Lambda^1(s) \left(u_1(s) \left[1 - x^*(s)\right]^{1/2} - u_2(s) x^*(s)^{1/2} \right) \Big\}, \\
u_2^*(s) = \arg\max_{u_2} \Big\{ &\left[q_2 (1 - x^*(s)) - \frac{c_2}{2} u_2(s)^2 \right] \exp(-rs) \\
&+ \Lambda^2(s) \left(u_1(s) \left[1 - x^*(s)\right]^{1/2} - u_2(s) x^*(s)^{1/2} \right) \Big\}, \\
\dot{x}^*(s) &= u_1^*(s) \left[1 - x^*(s)\right]^{1/2} - u_2^*(s) x^*(s)^{1/2}, \qquad x^*(0) = x_0, \\
\dot{\Lambda}^1(s) &= \\
&\left\{ -q_1 \exp(-rs) + \Lambda^1(s) \left(\frac{1}{2} u_1^*(s) \left[1 - x^*(s)\right]^{-1/2} + \frac{1}{2} u_2^*(s) x^*(s)^{-1/2} \right) \right\}, \\
\dot{\Lambda}^2(s) &= \\
&\left\{ q_2 \exp(-rs) + \Lambda^2(s) \left(\frac{1}{2} u_1^*(s) \left[1 - x^*(s)\right]^{-1/2} + \frac{1}{2} u_2^*(s) x^*(s)^{-1/2} \right) \right\}, \\
\Lambda^1(T) &= \exp(-rT) S_1, \\
\Lambda^2(T) &= -\exp(-rT) S_2.
\end{aligned} \tag{12.3}$$

Using (12.3), we obtain

$$u_1^*(s) = \frac{\Lambda^1(s)}{c_1} \left[1 - x^*(s)\right]^{1/2} \exp(rs),$$

and

$$u_2^*(s) = \frac{\Lambda^2(s)}{c_2} \left[x^*(s)\right]^{1/2} \exp(rs).$$

Upon substituting $u_1^*(s)$ and $u_2^*(s)$ into (12.3) yields:

$$\dot{\Lambda}^1(s) = \left\{ -q_1 \exp(-rs) + \left(\frac{\left[\Lambda^1(s)\right]^2}{2c_1} + \frac{\Lambda^1(s) \Lambda^2(s)}{2c_2} \right) \right\}, \tag{12.4}$$

$$\dot{\Lambda}^2(s) = \left\{ q_2 \exp(-rs) + \left(\frac{\left[\Lambda^2(s)\right]^2}{2c_2} + \frac{\Lambda^1(s) \Lambda^2(s)}{2c_1} \right) \right\},$$

with boundary conditions

$$\Lambda^1(T) = \exp(-rT) S_1 \text{ and } \Lambda^2(T) = -\exp(-rT) S_2.$$

The game equilibrium state dynamics becomes

$$\dot{x}^*(s) = \frac{\Lambda^1(s)\exp(rs)}{c_1}[1 - x^*(s)] - \frac{\Lambda^2(s)\exp(rs)}{c_2}x^*(s),$$
$$x^*(0) = x_0. \tag{12.5}$$

Solving the block recursive system of differential equations (12.4)–(12.5) gives the solutions to $x^*(s)$, $\Lambda^1(s)$ and $\Lambda^2(s)$. Upon substituting them into $u_1^*(s)$ and $u_2^*(s)$ yields the open-loop game equilibrium strategies.

12.2. Feedback solution in competitive advertising. A feedback solution which allows the firm to choose their advertising rates contingent upon the state of the game is a realistic approach to the problem (12.1)–(12.2). Invoking 11.4, a feedback Nash equilibrium solution to the game (12.1)–(12.2) has to satisfy the following conditions:

$$\begin{aligned}
-V_t^1(t,x) &= \max_{u_1}\Big\{\Big[q_1 x - \frac{c_1}{2}u_1^2\Big]\exp(-rt) \\
&\quad + V_x^1(t,x)\Big(u_1[1-x]^{1/2} - \phi_2^*(t,x)\,x^{1/2}\Big)\Big\}, \\
-V_t^2(t,x) &= \max_{u_2}\Big\{\Big[q_2(1-x) - \frac{c_2}{2}u_2^2\Big]\exp(-rt) \\
&\quad + V_x^2(t,x)\Big(\phi_1^*(t,x)[1-x]^{1/2} - u_2 x^{1/2}\Big)\Big\}, \\
V^1(T,x) &= \exp(-rT) S_1 x, \\
V^2(T,x) &= \exp(-rT) S_2(1-x).
\end{aligned} \tag{12.6}$$

Performing the indicated maximization in (12.6) yields:

$$\phi_1^*(t,x) = \frac{V_x^1(t,x)}{c_1}[1-x]^{1/2}\exp(rt) \quad \text{and}$$
$$\phi_2^*(t,x) = \frac{V_x^2(t,x)}{c_2}[x]^{1/2}\exp(rt).$$

Upon substituting $\phi_1^*(t,x)$ and $\phi_2^*(t,x)$ into (12.6) and solving (12.6) we obtain the value functions:

$$V^1(t,x) = \exp[-r(t)][A_1(t)x + B_1(t)] \quad \text{and}$$
$$V^2(t,x) = \exp[-r(t)][A_2(t)(1-x) + B_2(t)] \tag{12.7}$$

where $A_1(t)$, $B_1(t)$, $A_2(t)$ and $B_2(t)$ satisfy:

$$\begin{aligned}
\dot{A}_1(t) &= rA_1(t) - q_1 + \frac{A_1(t)^2}{2c_1} + \frac{A_1(t)A_2(t)}{2c_2}, \\
\dot{A}_2(t) &= rA_2(t) - q_2 + \frac{A_2(t)^2}{2c_2} + \frac{A_1(t)A_2(t)}{2c_1}, \\
A_1(T) &= S_1,\ B_1(T) = 0,\ A_2(T) = S_2 \text{ and } B_2(T) = 0.
\end{aligned}$$

Upon substituting the relevant partial derivatives of $V^1(t,x)$ and $V^2(t,x)$ from (12.7) into (12.6) yields the feedback Nash equilibrium strategies

$$\phi_1^*(t,x) = \frac{A_1(t)}{c_1}[1-x]^{1/2} \quad \text{and} \quad \phi_2^*(t,x) = \frac{A_2(t)}{c_2}[x]^{1/2}. \tag{12.8}$$

13 Infinite-horizon differential games

13.1. Consider the infinite-horizon autonomous game problem with constant discounting, in which T approaches infinity and where the objective functions and state dynamics are both autonomous. In particular, the game becomes:

$$\max_{u_i} \int_{t_0}^{\infty} g^i[x(s), u_1(s), u_2(s), \ldots, u_n(s)] \exp[-r(s-t_0)]\, ds,$$

$$\text{for } i \in N, \tag{13.1}$$

subject to the deterministic dynamics

$$\dot{x}(s) = f[x(s), u_1(s), u_2(s), \ldots, u_n(s)], \quad x(t_0) = x_0, \tag{13.2}$$

where r is a constant discount rate.

13.2. Game equilibrium solutions. Now consider the alternative game to (13.1)–(13.2)

$$\max_{u_i} \int_{t}^{\infty} g^i[x(s), u_1(s), u_2(s), \ldots, u_n(s)] \exp[-r(s-t)]\, ds,$$

$$\text{for } i \in N, \tag{13.3}$$

subject to the deterministic dynamics

$$\dot{x}(s) = f[x(s), u_1(s), u_2(s), \ldots, u_n(s)], \quad x(t) = x. \tag{13.4}$$

The infinite-horizon autonomous game (13.3)–(13.4) is independent of the choice of t and dependent only upon the state at the starting time, that is x.

In the infinite-horizon optimization problem in 10.2, the control is shown to be a function of the state variable x only. With the validity of the game equilibrium $\left\{u_i^*(s) = \phi_i^*(x) \in U^i, \text{ for } i \in N\right\}$ to be verified later, we define

Definition 16. *For the n-person differential game (13.1)–(13.2) with MPS or CLPS information, an n-tuple of strategies*

$$\left\{u_i^*(s) = \phi_i^*(x) \in U^i, \text{ for } i \in N\right\}$$

constitutes a feedback Nash equilibrium solution *if there exist functionals $V^i(t,x)$ defined on $[t_0,\infty) \times R^m$ and satisfying the following relations for each $i \in N$:*

$$V^i(t,x) = \int_t^\infty g^i\left[x^*(s), \phi_1^*(\eta_s), \phi_2^*(\eta_s), \ldots, \phi_n^*(\eta_s)\right] \exp\left[-r(s-t_0)\right] ds \geq$$

$$\int_t^\infty g^i\left[x^{[i]}(s), \phi_1^*(\eta_s), \phi_2^*(\eta_s), \ldots, \phi_{i-1}^*(\eta_s)\,\phi_i(\eta_s)\,\phi_{i+1}^*(\eta_s), \ldots, \phi_n^*(\eta_s)\right]$$

$$\times \exp\left[-r(s-t_0)\right] ds$$

$$\forall \phi_i(\cdot,\cdot) \in \Gamma^i, \quad x \in R^n,$$

where on the interval $[t_0,\infty)$,

$$\dot{x}^{[i]}(s) =$$

$$f\left[s^{[i]}(s), \phi_1^*(\eta_s), \phi_2^*(\eta_s), \ldots, \phi_{i-1}^*(\eta_s)\,\phi_i(\eta_s)\,\phi_{i+1}^*(\eta_s), \ldots, \phi_n^*(\eta_s)\right],$$

$$x^{[1]}(t) = x;$$

$$\dot{x}^*(s) = f\left[x^*(s), \phi_1^*(\eta_s), \phi_2^*(\eta_s), \ldots, \phi_n^*(\eta_s)\right], \qquad x^*(s) = x;$$

and η_s stands for either the data set $\{x(s), x_0\}$ or $\{x(\tau),\ \tau \leq s\}$, depending on whether the information pattern in MPS or CLPS.

We can write

$$V^i(t,x) = \exp\left[-r(t-t_0)\right] \int_t^\infty g^i\left[x^*(s), \phi_1^*(\eta_s), \phi_2^*(\eta_s), \ldots, \phi_n^*(\eta_s)\right]$$

$$\times \exp\left[-r(s-t)\right] ds,$$

$$\text{for } \ x(t) = x = x_t^* = x^*(t).$$

Since

$$\int_t^\infty g^i\left[x^*(s), \phi_1^*(\eta_s), \phi_2^*(\eta_s), \ldots, \phi_n^*(\eta_s)\right] \exp\left[-r(s-t)\right] ds$$

depends on the current state x only, we can write:

$$W^i(x) = \int_t^\infty g^i\left[x^*(s), \phi_1^*(\eta_s), \phi_2^*(\eta_s), \ldots, \phi_n^*(\eta_s)\right] \exp\left[-r(s-t)\right] ds.$$

It follows that:

$$V^i(t,x) = \exp\left[-r(t-t_0)\right] W^i(x), \tag{13.5}$$

$$V_t^i(t,x) = -r\exp[-r(t-t_0)]\,W^i(x)\,, \text{ and}$$
$$V_x^i(t,x) = \exp[-r(t-t_0)]\,W_x^i(x)\,, \text{ for } i \notin N.$$

A feedback Nash equilibrium solution for the infinite-horizon autonomous game (13.3) and (13.4) can be characterized as follows:

Theorem 23. *An n-tuple of strategies $\left\{u_i^*(s) = \phi_i^*(\cdot) \in U^i;\ \text{for } i \in N\right\}$ provides a feedback Nash equilibrium solution to the infinite-horizon game (13.3) and (13.4) if there exist continuously differentiable functions $W^i(x) : R^m \to R,\ i \in N$, satisfying the following set of partial differential equations:*

$$\begin{aligned}
&rW^i(x) = \\
&\max_{u_i}\left\{g^i\left[x, \phi_1^*(x), \phi_2^*(x), \ldots, \phi_{i-1}^*(x), u_i, \phi_{i+1}^*(x), \ldots, \phi_n^*(x)\right]\right. \\
&\qquad \left. + W_x^i(x)\, f\left[x, \phi_1^*(x), \phi_2^*(x), \ldots, \phi_{i-1}^*(x), u_i(x), \phi_{i+1}^*(x), \ldots, \phi_n^*(x)\right]\right\} \\
&= \left\{g^i[x, \phi_1^*(x), \phi_2^*(x), \ldots, \phi_n^*(x)]\right. \\
&\qquad \left. + W_x^i(x)\, f[x, \phi_1^*(x), \phi_2^*(x), \ldots, \phi_n^*(x)]\right\}, \qquad \textit{for } i \in N.
\end{aligned}$$

Proof. By 10.2, $W^i(x)$ is the value function associated with the optimal control problem of Player i, $i \in N$. Together with the i^{th} expression in Definition, the conditions in Theorem imply a Nash equilibrium.

Since time s is not explicitly involved the partial differential equations in Theorem, the validity of the feedback Nash equilibrium $\{u_i^* = \phi_i^*(x)$, for $i \in N\}$ are functions independent of time is obtained.

Substituting the game equilibrium strategies in Theorem into (13.2) yields the game equilibrium dynamics of the state path as:

$$\dot{x}(s) = f[x(s), \phi_1^*(x(s)), \phi_2^*(x(s)), \ldots, \phi_n^*(x(s))], \qquad x(t_0) = x_0.$$

Solving the above dynamics yields the optimal state trajectory $\{x^*(t)\}_{t \ge t_0}$ as

$$x^*(t) = x_0 + \int_{t_0}^{t} f[x^*(s), \phi_1^*(x^*(s)), \phi_2^*(x^*(s)), \ldots, \phi_n^*(x^*(s))]\, ds, \quad \text{for } t \ge t_0.$$

We denote term $x^*(t)$ by x_t^*. The feedback Nash equilibrium strategies for the infinite-horizon game (13.1)–(13.2) can be obtained as

$$[\phi_1^*(x_t^*), \phi_2^*(x_t^*), \ldots, \phi_n^*(x_t^*)], \qquad \text{for } t \ge t_0.$$

Following the above analysis and using Theorems in 6.4 and 6.5, we can characterize an open loop equilibrium solution to the infinite-horizon game (13.3) and (13.4) as:

Theorem 24. *A set of strategies* $\{u_i^*(s) = \zeta_i^*(s, x_t)\,,\ \text{for } i \in N\}$ *provides an open-loop Nash equilibrium solution to the infinite-horizon game (13.3) and (13.4), and* $\{x^*(s)\,,\ t \le s \le T\}$ *is the corresponding state trajectory, if there exist* m *costate functions* $\Lambda^i(s) : [t, T] \to R^m$, *for* $i \in N$, *such that the following relations are satisfied:*

$$\zeta_i^*(s, x) \equiv u_i^*(s) =$$
$$\arg\max_{u_i \in U^i} \left\{ g^i\left[x^*(s), u_1^*(s), u_2^*(s), \ldots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \ldots, u_n^*(s)\right] \right.$$
$$\left. +\lambda^i(s) f\left[x^*(s), u_1^*(s), u_2^*(s), \ldots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \ldots, u_n^*(s)\right] \right\},$$
$$\dot{x}^*(s) = f\left[x^*(s), u_1^*(s), u_2^*(s), \ldots, u_n^*(s)\right], \qquad x^*(t) = x_t,$$
$$\dot{\lambda}^i(s) = r\lambda(s) - \frac{\partial}{\partial x^*}\left\{g^i\left[x^*(s), u_1^*(s), u_2^*(s), \ldots\right.\right.$$
$$\left.\left. \ldots, u_n^*(s)\right] + \lambda^i(s) f\left[x^*(s), u_1^*(s), u_2^*(s), \ldots, u_n^*(s)\right]\right\},$$
$$\textit{for } i \in N.$$

Proof. Consider the i^{th} equality in the above Theorem, which states that $\upsilon_i^*(s) = u_i^*(s) = \zeta_i^*(s, x_t)$ maximizes

$$\int_{t_0}^{\infty} g^i\left[x(s), u_1^*(s), u_2^*(s), \ldots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \ldots, u_n^*(s)\right] ds,$$

over the choice of $\upsilon_i(s) \in U^i$ subject to the state dynamics:

$$\dot{x}(s) = f\left[x(s), u_1^*(s), u_2^*(s), \ldots, u_{i-1}^*(s), u_i(s), u_{i+1}^*(s), \ldots, u_n^*(s)\right],$$
$$x(t) = x_t, \qquad \text{for } i \in N.$$

This is the infinite-horizon optimal control problem for Player i, since $u_j^*(s)$, for $j \in N$ and $j \neq i$, are open-loop controls and hence do not depend on $u_i^*(s)$. These results are stated in 10.3.

13.3. Infinite-horizon duopolistic competition. Consider a dynamic duopoly in which there are two publicly listed firms selling a homogeneous good. Since the value of a publicly listed firm is the present value of its discounted expected future earnings, the terminal time of the game, T, may be very far in the future and nobody knows when the firms will be out of business. Therefore setting $T = \infty$ may very well be the best approximation for the true game horizon. Even if the firm's management restricts itself to considering profit maximization over the next year, it should value its asset positions at the end of the year by the earning potential of these assets in the years to come. There is a lag in price adjustment so the evolution of market price over time is assumed to be a function of the current market price and the price specified by the current demand condition. In particular, we follow [11] and assume that

$$\dot{P}(s) = k\left[a - u_1(s) - u_2(s) - P(s)\right], \quad P(t_0) = P_0, \tag{13.6}$$

where $P(s)$ is the market price at time s, $u_i(s)$ is output supplied firm $i \in \{1,2\}$, current demand condition is specified by the instantaneous inverse demand function $P(s) = [a - u_1(s) - u_2(s)]$, and $k > 0$ represents the price adjustment velocity.

The payoff of firm i is given as the present value of the stream of discounted profits

$$\int_{t_0}^{\infty} \left\{ P(s)\, u_i(s) - c u_i(s) - (1/2)\,[u_i(s)]^2 \right\} \exp[-r(s - t_0)]\, ds,$$

$$\text{for } \; i \in \{1,2\}, \tag{13.7}$$

where $c u_i(s) + (1/2)\,[u_i(s)]^2$ is the cost of producing output $u_i(s)$ and r is the interest rate.

Once again, we consider the alternative game

$$\max_{u_i} \int_{t_0}^{\infty} \left\{ P(s)\, u_i(s) - c u_i(s) - (1/2)\,[u_i(s)]^2 \right\} \exp[-r(s - t)]\, ds,$$

$$\text{for } \; i \in \{1,2\}, \tag{13.8}$$

subject to

$$\dot{P}(s) = k\,[a - u_1(s) - u_2(s) - P(s)], \quad P(t) = P. \tag{13.9}$$

The infinite-horizon game (13.8)–(13.9) has autonomous structures and a constant rate. Therefore, we can apply the second theorem from 13.2 to characterize a feedback Nash equilibrium solution as

$$rW^i(P) = \max_{u_i} \left\{ \left[P u_i - c u_i - (1/2)\,(u_i)^2 \right] \right. \tag{13.10}$$

$$\left. + W_P^i \left[k \left(a - u_i - \phi_j^*(P) - P \right) \right] \right\}, \quad \text{for } \; i \in \{1,2\}.$$

Performing the indicated maximization in (13.10), we obtain:

$$\phi_i^*(P) = P - c - k W_P^i(P), \quad \text{for } \; i \in \{1,2\}. \tag{13.11}$$

Substituting the results from (13.11) into (13.10), and upon solving (13.10) yields:

$$W^i(P) = \frac{1}{2} A P^2 - B P + C, \tag{13.12}$$

where

$$A = \frac{r + 6k - \sqrt{(r+6k)^2 - 12k^2}}{6k^2},$$

$$B = \frac{-akA + c - 2kcA}{r - 3k^2 A + 3k}, \quad \text{and}$$

$$C = \frac{c^2 + 3k^2 B^2 - 2kB\,(2c + a)}{2r}.$$

Again, one can readily verify that $W^i(P)$ in (13.12) indeed solves (13.10) by substituting $W^i(P)$ and its derivative into (13.10) and (13.11).

The game equilibrium strategy can then be expressed as:

$$\phi_i^*(P) = P - c - k(AP - B), \quad \text{for } i \in \{1,2\}.$$

Substituting the game equilibrium strategies above into (13.6) yields the game equilibrium state dynamics of the game (13.6)–(13.7) as:

$$\dot{P}(s) = k[a - 2(c + kB) - (3 - kA)P(s)], \quad P(t_0) = P_0.$$

Solving the above dynamics yields the optimal state trajectory as

$$P^*(t) = \left[P_0 - \frac{k[a - 2(c + kB)]}{k(3 - kA)}\right] \exp[-k(3 - kA)t] + \frac{k[a - 2(c + kB)]}{k(3 - kA)}.$$

We denote term $P^*(t)$ by P_t^*. The feedback Nash equilibrium strategies for the infinite-horizon game (13.6)–(13.7) can be obtained as

$$\phi_i^*(P_t^*) = P_t^* - c - k(AP_t^* - B), \quad \text{for } \{1,2\}.$$

14 Cooperative differential games in characteristic function form

We begin with the basic formulation of cooperative differential games in characteristic function form and the solution imputations.

14.1. Game formulation. Consider a general N-person differential game in which the state dynamics has the form

$$\dot{x}(s) = f[s, x(s), u_1(s), u_2(s), \ldots, u_n(s)], \qquad x(t_0) = x_0, \tag{14.1}$$

The payoff of Player i is:

$$\int_{t_0}^{T} g^i[s, x(s), u_1(s), u_2(s), \ldots, u_n(s)]\, ds + q^i(x(T)),$$
$$\text{for } i \in N = \{1, 2, \ldots, n\}, \tag{14.2}$$

where $x(s) \in X \subset R^m$ denotes the state variables of game, and $u_i \in U^i$ is the control of Player i, for $i \in N$. In particular, the players' payoffs are transferable. Invoking 11.4, a feedback Nash equilibrium solution can be characterized if the players play noncooperatively.

Now consider the case when the players agree to cooperate. Let $\Gamma_c(x_0, T - t_0)$ denote a cooperative game with the game structure of $\Gamma(x_0, T - t_0)$ in which the players agree to act according to an agreed upon optimality principle. The agreement on how to act cooperatively and allocate cooperative payoff constitutes the solution optimality principle of a cooperative scheme. In particular, the solution optimality principle for a cooperative game $\Gamma_c(x_0, T - t_0)$ includes

(i) an agreement on a set of cooperative strategies/controls, and

(ii) a mechanism to distribute total payoff among players.

The solution optimality principle will remain in effect along the cooperative state trajectory path $\{x_s^*\}_{s=t_0}^{T}$. Moreover, group rationality requires the players to seek a set of cooperative strategies/controls that yields a pareto-optimal solution. In addition, the allocation principle has to satisfy individual rationality in the sense that neither player would be no worse off than before under cooperation.

To fulfill group rationality in the case of transferable payoffs, the players have to maximize the sum of their payoffs

$$\sum_{j=1}^{N} \left\{ \int_{t_0}^{T} g^j \left[s, x\left(s\right), u_1\left(s\right), u_2\left(s\right), \ldots, u_n\left(s\right) \right] ds + q^j \left(x\left(T\right) \right) \right\}, \tag{14.3}$$

subject to (14.1).

Invoking Pontryagin's Maximum Principle, a set of optimal controls $u^*\left(s\right) = \left[u_1^*\left(s\right), u_2^*\left(s\right), \ldots, u_n^*\left(s\right)\right]$ can be characterized as in 10.3. Substituting this set of optimal controls into (14.1) yields the optimal trajectory $\{x^*\left(t\right)\}_{t=t_0}^{T}$, where

$$x^*\left(t\right) = x_0 + \int_{t_0}^{t} f\left[s, x^*\left(s\right) u^*\left(s\right)\right] ds, \qquad \text{for } \; t \in \left[t_0, T\right]. \tag{14.4}$$

For notational convenience in subsequent exposition, we use $x^*\left(t\right)$ and x_t^* interchangeably.

We denote

$$\sum_{j=1}^{n} \left\{ \int_{t_0}^{T} g^j \left[s, x^*\left(s\right), u^*\left(s\right) \right] ds + q^j \left(x^*\left(T\right) \right) \right\}$$

by $v\left(N; x_0, T - t_0\right)$. Let $S \subseteq N$ and $v\left(S; x_0, T - t_0\right)$ stands for a characteristic function reflecting the payoff of coalition S. The quantity $v\left(S; x_0, T - t_0\right)$ yields the maximized payoff to coalition S as the rest of the players form a coalition $N \backslash S$ to play against S. Calling on the super-additivity property of characteristic functions, $v\left(S; x_0, T - t_0\right) \geq v\left(S'; x_0, T - t_0\right)$ for $S' \subset S \subseteq N$. Hence, it is advantageous for the players to form a maximal coalition N and obtain a maximal total payoff $v\left(N; x_0, T - t_0\right)$ that is possible in the game.

14.2. Solution imputation. One of the integral parts of cooperative game is to explore the possibility of forming coalitions and offer an "agreeable" distribution of the total cooperative payoff among players. In fact, the characteristic function framework displays the possibilities of coalitions in an effective manner and establishes a basis for formulating distribution schemes of the total payoffs that are acceptable to participating players.

We can use $\Gamma_v(x_0, T - t_0)$ to denote a *cooperative differential game in characteristic function form*. The optimality principle for a cooperative game in characteristic function form includes

(i) an agreement on a set of cooperative strategies/controls

$$u^*(s) = [u_1^*(s), u_2^*(s), \ldots, u_n^*(s)], \quad \text{for } s \in [t_0, T], \text{ and}$$

(ii) a mechanism to distribute total payoff among players.

A set of payoff distributions satisfying the optimality principle is called a solution imputation to the cooperative game. We will now examine the solutions to $\Gamma_v(x_0, T - t_0)$. Denote by $\xi_i(x_0, T - t_0)$ the share of the player $i \in N$ from the total payoff $v(N; x_0, T - t_0)$.

Definition 17. *A vector*

$$\xi(x_0, T - t_0) = [\xi_1(x_0, T - t_0), \xi_2(x_0, T - t_0), \ldots, \xi_n(x_0, T - t_0)]$$

that satisfies the conditions:

(i)
$$\xi_i(x_0, T - t_0) \geq v(\{i\}; x_0, T - t_0), \quad \textit{for } i \in N, \textit{ and}$$

(ii)
$$\sum_{j \in N} \xi_j(x_0, T - t_0) = v(N; x_0, T - t_0)$$

is called an imputation of the game $\Gamma_v(x_0, T - t_0)$.

Part (i) of Definition guarantees individual rationality in the sense that each player receives at least the payoff he or she will get if play against the rest of the players. Part (ii) ensures Pareto optimality and hence group rationality.

Theorem 25. *Suppose the function $w : 2^n \times R^m \times R^1 \to R^1$ is additive in $S \in 2^n$, that is for any $S,\ A \in 2^n,\ S \cap A = \emptyset$ we have $w(S \cup A;\ x_0, T - t_0) = w(S; x_0, T - t_0) + w(A; x_0, T - t_0)$. Then in the game $\Gamma_w(x_0, T - t_0)$ there is a unique imputation $\xi_i(x_0, T - t_0) = w(\{i\}; x_0, T - t_0)$, for $i \in N$. Proof. From the additivity of w we immediately obtain*

$$\begin{aligned} &w(N; x_0, T - t_0) = \\ &\quad w(\{1\}; x_0, T - t_0) + w(\{2\}; x_0, T - t_0) + \cdots + w(\{n\}; x_0, T - t_0), \end{aligned}$$

Hence the Theorem follows.

Games with additive characteristic functions are called inessential and games with superadditive characteristic functions are called essential. In an essential game $\Gamma_v(x_0, T - t_0)$ there is an infinite set of imputations. Indeed, any vector of the form

$$[v(\{1\}; x_0, T - t_0) + \alpha_1, v(\{2\}; x_0, T - t_0) + \alpha_2, \ldots$$

$$\ldots, v\left(\{n\}; x_0, T - t_0\right) + \alpha_n\right], \tag{14.5}$$

$$\text{for } \alpha_i \geq 0, \quad i \in N \text{ and}$$

$$\sum_{i \in N} \alpha_i = v\left(N; x_0, T - t_0\right) - \sum_{i \in N} v\left(\{i\}; x_0, T - t_0\right),$$

is an imputation of the game $\Gamma_v(x_0, T - t_0)$. We denote the imputation set of $\Gamma_v(x_0, T - t_0)$ by $E_v(x_0, T - t_0)$.

Definition 18. *The imputation $\xi(x_0, T - t_0)$ dominates the imputation $\eta(x_0, T - t_0)$ in the coalition S, or $\xi(x_0, T - t_0) \overset{S}{\succ} \eta(x_0, T - t_0)$, if*

(i)
$$\xi_i\left(x_0, T - t_0\right) \geq \eta_i\left(x_0, T - t_0\right), \quad i \in S; \text{ and}$$

(ii)
$$\sum_{i \in S} \xi_i\left(x_0, T - t_0\right) \leq v\left(S; x_0, T - t_0\right).$$

The imputation $\xi(x_0, T - t_0)$ is said to dominate the imputation $\eta(x_0, T - t_0)$, or $\xi(x_0, T - t_0) \succ \eta(x_0, T - t_0)$, if there does not exist any $S \subset N$ such that $\eta(x_0, T - t_0) \overset{S}{\succ} \xi(x_0, T - t_0)$ but there exists coalition $S \subset N$ such that $\xi(x_0, T - t_0) \overset{S}{\succ} \eta(x_0, T - t_0)$. It follows from the definition of imputation that domination in single-element coalition and coalition N, is not possible.

Definition 19. *The set of undominated imputations is called the core of the game $\Gamma_v(x_0, T - t_0)$ and is denoted by $C_v(x_0, T - t_0)$.*

Definition 20. *The set $L_v(x_0, T - t_0) \subset E_v(x_0, T - t_0)$ is called the Neumann-Morgenstern solution (the NM-solution) of the game $\Gamma_v(x_0, T - t_0)$ if*

(i)
$$\begin{aligned} &\xi\left(x_0, T - t_0\right), \eta\left(x_0, T - t_0\right) \in L_v\left(x_0, T - t_0\right), \text{ implies} \\ &\xi\left(x_0, T - t_0\right) \not\succ \eta\left(x_0, T - t_0\right), \end{aligned}$$

(ii)
$$\begin{aligned} &\text{for } \eta\left(x_0, T - t_0\right) \notin L_v\left(x_0, T - t_0\right) \text{ there exists} \\ &\xi\left(x_0, T - t_0\right) \in L_v\left(x_0, T - t_0\right) \text{ such that} \\ &\xi\left(x_0, T - t_0\right) \succ \eta\left(x_0, T - t_0\right). \end{aligned}$$

Note that the NM-solution always contains the core.

Definition 21. *The vector*

$$\Phi^v\left(x_0, T - t_0\right) = \left\{\Phi_i^v\left(x_0, T - t_0\right), \; i = 1, \ldots, n\right\}$$

is called the Shapley value if it satisfies the following conditions:

$$\begin{aligned} &\Phi_i^v\left(x_0, T - t_0\right) = \\ &\sum_{S \subset N (S \ni i)} \frac{(n-s)!\,(s-1)!}{n!} \left[v\left(S; x_0, T - t_0\right) - v\left(S \backslash i; x_0, T - t_0\right)\right]; \\ &i = 1, \ldots, n. \end{aligned}$$

The components of the Shapley value are the players' shares of the cooperative payoff. The Shapley value is unique and is an imputation (see [12]). Unlike the core and NM-solution, the Shapley value represents an optimal distribution principle of the total gain $v(N; x_0, T-t_0)$ and is defined without using the concept of domination.

15 Imputation in a dynamic context

Sec. 14.2 characterizes the solution imputation at the outset of the game. In dynamic games, the solution imputation along the cooperative trajectory $\{x^*(t)\}_{t=t_0}^{T}$ would be of concern to the players. In this section, we focus our attention on the dynamic imputation brought about by the solution optimality principle.

Let an optimality principle be chosen in the game $\Gamma_v(x_0, T-t_0)$. The solution of this game constructed in the initial state $x(t_0) = x_0$ based on the chosen principle of optimality contains the solution imputation set $W_v(x_0, T-t_0) \subseteq E_v(x_0, T-t_0)$ and the conditionally optimal trajectory $\{x^*(t)\}_{t=t_0}^{T}$ which maximizes

$$\sum_{j=1}^{n}\left\{\int_{t_0}^{T} g^j[s, x^*(s), u^*(s)]\,ds + q^j(x^*(T))\right\}.$$

Assume that $W_v(x_0, T-t_0) \neq \emptyset$.

Definition 22. *Any trajectory $\{x^*(t)\}_{t=t_0}^{T}$ of the system (14.1) such that*

$$\sum_{j=1}^{n}\left\{\int_{t_0}^{T} g^j[s, x^*(s), u^*(s)]\,ds + q^j(x^*(T))\right\} = v(N; x_0, T-t_0)$$

is called a conditionally optimal trajectory in the game $\Gamma_v(x_0, T-t_0)$.

Definition suggests that along the conditionally optimal trajectory the players obtain the largest total payoff. For exposition sake, we assume that such a trajectory exits. Now we consider the behavior of the set $W_v(x_0, T-t_0)$ along the conditionally optimal trajectory $\{x^*(t)\}_{t=t_0}^{T}$. Towards this end, in each current state $x^*(t) \equiv x_t^*$ the current subgame $\Gamma_v(x_t^*, T-t)$ is defined as follows. At time t with state $x^*(t)$, we define the characteristic function

$$v(S; x_t^*, T-t) = \begin{cases} 0, & S = \emptyset \\ Val\ \Gamma_S(x_t^*, T-t), \quad \text{if} & S \subset N \\ K_N(x^*(t), u^*(\cdot), T-t) & S = N \end{cases} \tag{15.1}$$

where

$$K_N(x_t^*, u^*(\cdot), T-t) = \sum_{j=1}^{n}\left\{\int_{t}^{T} g^j[s, x^*(s), u^*(s)]\,ds + q^j(x^*(T))\right\}$$

is the total payoff of the players over the time interval $[t, T]$ along the conditionally optimal trajectory $\{x^*(s)\}_{s=t}^{T}$; and $Val\ \Gamma_S(x_t^*, T-t)$ is the value of the zero-sum differential game $\Gamma_S(x_t^*, T-t)$ between coalitions S and $N \backslash S$ with initial state $x^*(t) \equiv x_t^*$, duration $T-t$ and the S coalition being the maximizer.

The imputation set in the game $\Gamma_v(x_t^*, T-t)$ is of the form:

$$E_v(x_t^*, T-t) = \left\{ \xi \in R^n \middle| \xi_i \geq v(\{i\}; x_t^*, T-t),\ i = 1, 2, \ldots, n; \right.$$
$$\left. \sum_{i \in N} \xi_i = v(N; x_t^*, T-t) \right\}, \tag{15.2}$$

where

$$v(N; x_t^*, T-t) = v(N; x_0, T-t_0) - \sum_{j=1}^{n} \left\{ \int_{t_0}^{t} g^j[s, x^*(s), u^*(s)]\, ds + q^j(x^*(T)) \right\}.$$

The quantity

$$\sum_{j=1}^{n} \left\{ \int_{t_0}^{t} g^j[s, x^*(s), u^*(s)]\, ds + q^j(x^*(T)) \right\}$$

denoted the cooperative payoff of the players over the time interval $[t_0, t]$ along the trajectory $\{x^*(s)\}_{s=t_0}^{T}$.

Consider the family of current games

$$\{\Gamma_v(x_t^*, T-t),\ t_0 \leq t \leq T\},$$

and their solutions $W_v(x_t^*, T-t) \subset E_v(x_t^*, T-t)$ generated by the same principle of optimality that yields the initially solution $W_v(x_0, T-t_0)$.

Lemma 9. *The set $W_v(x_T^*, 0)$ is a solution of the current game $\Gamma_v(x_T^*, 0)$ at time T and is composed of the only imputation*

$$\begin{aligned} q(x^*(T)) &= \left\{q^1(x^*(T)), q^2(x^*(T)), \ldots, q^n(x^*(T))\right\} \\ &= \left\{q^1(x_T^*), q^2(x_T^*), \ldots, q^n(x_T^*)\right\}. \end{aligned}$$

Proof. Since the game $\Gamma_v(x_T^*, 0)$ is of zero-duration, then for all $i \in N$, $v(\{i\}; x_T^*, 0) = q^i(x_T^*)$. Hence

$$\sum_{i \in N} v(\{i\}; x_T^*, 0) = \sum_{i \in N} q^i(x_T^*) = v(N; x_T^*, 0),$$

and the characteristic function of the game $\Gamma_v(x_T^*, 0)$ is additive for S and, by Theorem,

$$E_v(x_T^*, 0) = q(x_T^*) = W_v(x_T^*, 0).$$

This completes the proof of Lemma.

16 Principle of dynamic stability

Formulation of optimal behaviors for players is a fundamental element in the theory of cooperative games. The players' behaviors satisfying some specific optimality principles constitute a solution of the game. In other words, the solution of a cooperative game is generated by a set of optimality principles (for instance, in the [12]), the von Neumann Morgenstern solution [3] and the Nash bargaining solution [9]). For dynamic games, an additional stringent condition on their solutions is required: the specific optimality principle must remain optimal at any instant of time throughout the game along the optimal state trajectory chosen at the outset. This condition is known as *dynamic stability or time consistency.* Assume that at the start of the game the players adopt an optimality principle (which includes the consent to maximize the joint payoff and an agreed upon payoff distribution principle). When the game proceeds along the "optimal" trajectory, the state of the game changes and the optimality principle may not be feasible or remain optimal to all players. Then, some of the players will have an incentive to deviate from the initially chosen trajectory. If this happens, instability arises. In particular, the dynamic stability of a solution of a cooperative differential game is the property that, when the game proceeds along an "optimal" trajectory, at each instant of time the players are guided by the same optimality principles, and yet do not have any ground for deviation from the previously adopted "optimal" behavior throughout the game.

The question of dynamic stability in differential games has been explored rigorously in the past three decades. Haurie [10] discussed the problem of instability in extending the Nash bargaining solution to differential games. Petrosyan [13] formalized mathematically the notion of dynamic stability in solutions of differential games. Petrosyan and Danilov [14] introduced the notion of "imputation distribution procedure" for cooperative solution. In particular, the method of regularization was introduced to construct time consistent solutions. Yeung and Petrosyan [15] designed a time consistent solution in differential games and characterized the conditions that the allocation distribution procedure must satisfy.

17 Dynamic stable solutions

Let there exist solutions $W_v\left(x_t^*, T-t\right) \neq \emptyset$, $t_0 \leq t \leq T$ along the conditionally optimal trajectory $\{x^*(t)\}_{t=t_0}^{T}$. If this condition is not satisfied, it is impossible for the players to adhere to the chosen principle of optimality, since at the very first instant t, when $W_v\left(x_t^*, T-t\right) \neq \emptyset$, the players have no possibility to follow this principle. Assume that at time t_0 when the initial state x_0 is the players agree on the imputation

$$\xi\left(x_0, T-t_0\right) = \\ \left[\xi_1\left(x_0, T-t_0\right), \xi_2\left(x_0, T-t_0\right), \ldots, \xi_n\left(x_0, T-t_0\right)\right] \in W_v\left(x_0, T-t_0\right).$$

This means that the players agree on an imputation of the gain in such a way that the share of the i^{th} player over the time interval $[t_0, T]$ is equal to $\xi_i(x_0, T-t_0)$. If according to $\xi(x_0, T-t_0)$ Player i is supposed to receive a payoff equaling $\varpi_i[\xi(x_0, T-t_0)\ ; x^*(\cdot), t-t_0]$ over the time interval $[t_0, t]$, then over the remaining time interval $[t, T]$ according to the $\xi(x_0, T-t_0)$ Player i is supposed to receive

$$\eta_i[\xi(x_0, T-t_0); x^*(t), T-t] = \tag{17.1}$$

$$\xi_i(x_0, T-t_0) - \varpi_i[\xi(x_0, T-t_0); x^*(\cdot), t-t_0].$$

Theorem 26. *Let $\eta[\xi(x_0, T-t_0); x^*(t), T-t]$ be the vector containing*

$$\eta_i[\xi(x_0, T-t_0); x^*(t), T-t], \quad \textit{for } i \in \{1, 2, \ldots, n\}.$$

For the original imputation agreement (that is the imputation $\xi(x_0, T-t_0)$) to remain in force at the instant t, it is essential that the vector

$$\eta[\xi(x_0, T-t_0); x^*(t), T-t] \in W_v(x_t^*, T-t), \tag{17.2}$$

and $\eta[\xi(x_0, T-t_0); x^*(t), T-t]$ is indeed a solution of the current game $\Gamma_v(x_t^*, T-t)$. If such a condition is satisfied at each instant of time $t \in [t_0, T]$ along the trajectory $\{x^*(t)\}_{t=t_0}^{T}$, then the imputation $\xi(x_0, T-t_0)$ is dynamical stable.

Along the trajectory $x^*(t)$ over the time interval $[t, T]$, $t_0 \le t \le T$, the coalition N obtains the payoffs

$$v(N; x^*(t), T-t) = \sum_{j=1}^{n} \left\{ \int_t^T g^j[s, x^*(s), u^*(s)]\, ds + q^j(x^*(T)) \right\}. \tag{17.3}$$

Then the difference

$$v(N; x_0, T-t_0) - v(N; x^*(t), T-t) = \sum_{j=1}^{n} \left\{ \int_{t_0}^{t} g^j[s, x^*(s), u^*(s)]\, ds \right\}$$

is the payoff the coalition N obtains on the time interval $[t_0, t]$.

Dynamic stability or time consistency of the solution imputation $\xi(x_0, T-t_0)$ guarantees that the extension of the solution policy to a situation with a later starting time and along the optimal trajectory remains optimal. Moreover, group and individual rationalities are satisfied throughout the entire game interval. A payment mechanism leading to the realization of this imputation scheme must be formulated. This will be done in the next section.

18 Payoff distribution procedure

A payoff distribution procedure (PDP) proposed by Petrosyan [13] will be formulated so that the agreed upon dynamically stable imputations can be realized.

Let the payoff Player i receive over the time interval $[t_0, t]$ be expressed as

$$\varpi_i \left[\xi\left(x_0\left(\cdot\right), T - t_0\right); x^*\left(\cdot\right), t - t_0\right] = \int_{t_0}^{t} B_i\left(s\right) ds, \tag{18.1}$$

where

$$\sum_{j \in N} B_j\left(s\right) = \sum_{j \in N} g^j\left[s, x^*\left(s\right), u^*\left(s\right)\right], \quad \text{for } t_0 \le s \le t \le T.$$

From (18.1) we get

$$\frac{d\varpi_i}{dt} = B_i\left(t\right). \tag{18.2}$$

This quantity may be interpreted as the instantaneous payoff of the Player i at the moment t. Hence it is clear the vector $B\left(t\right) = \left[B_1\left(t\right), B_2\left(t\right), \ldots, B_n\left(t\right)\right]$ prescribes distribution of the total gain among the members of the coalition N. By properly choosing $B\left(t\right)$, the players can ensure the desirable outcome that at each instant $t \in [t_0, T]$ there will be no objection against realization of the original agreement (the imputation $\xi\left(x_0, T - t_0\right)$).

Definition 23. *The imputation $\xi\left(x_0, T - t_0\right) \in W_v\left(x_0, T - t_0\right)$ is dynamically stable in the game $\Gamma_v\left(x_0, T - t_0\right)$ if the following conditions are satisfied:*

1. *there exists a conditionally optimal trajectory $\left\{x^*\left(t\right)\right\}_{t=t_0}^{T}$ along which $W_v\left(x^*\left(t\right), T - t\right) \ne \emptyset$, $t_0 \le t \le T$,*
2. *there exists function $B\left(t\right) = \left[B_1\left(t\right), B_2\left(t\right), \ldots, B_n\left(t\right)\right]$ integrable along $[t_0, T]$ such that*

$$\sum_{j \in N} B_j\left(t\right) = \sum_{j \in N} g^j\left[t, x^*\left(t\right), u^*\left(t\right)\right] \quad \textit{for } t_0 \le s \le t \le T, \quad \textit{and}$$

$$\xi\left(x_0, T - t_0\right) \in \bigcap_{t_0 \le t \le T} \left(\varpi\left[\xi\left(x_0\left(\cdot\right), T - t_0\right); x^*\left(\cdot\right), t - t_0\right] \oplus W_v\left(x^*\left(t\right), T - t\right)\right)$$

where $\varpi\left[\xi\left(x_0\left(\cdot\right), T - t_0\right); x^*\left(\cdot\right), t - t_0\right]$ is the vector of

$$\varpi_i\left[\xi\left(x_0\left(\cdot\right), T - t_0\right); x^*\left(\cdot\right), t - t_0\right], \quad \textit{for } i \in N;$$

and $W_v\left(x^*\left(t\right), T - t\right)$ is a solution of the current game $\Gamma_v\left(x^*\left(t\right), T - t\right)$; and the operator $\oplus$ defines the operation: for $\eta \in R^n$ and $A \subset R^n$, $\eta \oplus A = \left\{\pi + a \mid a \in A\right\}$.

The cooperative differential game $\Gamma_v\left(x_0, T - t_0\right)$ has a dynamic stable solution $W_v\left(x_0, T - t_0\right)$ if all of the imputations $\xi\left(x_0, T - t_0\right) \in W_v\left(x_0, T - t_0\right)$ are dynamically stable. The conditionally optimal trajectory along which there exists a dynamically stable solution of the game $\Gamma_v\left(x_0, T - t_0\right)$ is called an *optimal trajectory*.

From Definition we have

$$\xi\left(x_0, T-t_0\right) \in \left(\varpi\left[\xi\left(x_0\left(\cdot\right), T-t_0\right); x^*\left(\cdot\right), T-t_0\right] \oplus W_v\left(x^*\left(T\right), 0\right)\right)$$

where $W_v\left(x^*\left(T\right), 0\right) = q\left(x^*\left(T\right)\right)$ is a solution of the game $\Gamma_v\left(x^*\left(T\right), 0\right)$. Therefore, we can write

$$\xi\left(x_0, T-t_0\right) = \int_{t_0}^{T} B\left(s\right) ds + q\left(x^*\left(T\right)\right).$$

The dynamically stable imputation $\xi\left(x_0, T-t_0\right) \in W_v\left(x_0, T-t_0\right)$ may be realized as follows. From Definition at any instant $t_0 \leq t \leq T$ we have

$$\xi\left(x_0, T-t_0\right) \in \left(\varpi\left[\xi\left(x_0\left(\cdot\right), T-t_0\right); x^*\left(\cdot\right), t-t_0\right] \oplus W_v\left(x^*\left(t\right), T-t\right)\right), \tag{18.3}$$

where $\varpi\left[\xi\left(x_0\left(\cdot\right), T-t_0\right); x^*\left(\cdot\right), t-t_0\right] = \int_{t_0}^{t} B\left(s\right) ds$ is the payoff vector on the time interval $[t_0, t]$.

Player i's payoff over the same interval can be expressed as:

$$\varpi_i\left[\xi\left(x_0\left(\cdot\right), T-t_0\right); x^*\left(\cdot\right), t-t_0\right] = \int_{t_0}^{t} B\left(s\right) ds.$$

When the game proceeds along the optimal trajectory, over the time interval $[t_0, t]$ the players share the total gain $\int_{t_0}^{t} \sum_{j\in N} g^j\left[s, x^*\left(s\right), u^*\left(s\right)\right] ds$ so that the inclusion

$$\xi\left(x_0, T-t_0\right) - \varpi\left[\xi\left(x_0\left(\cdot\right), T-t_0\right); x^*\left(\cdot\right), t-t_0\right] \in W_v\left(x^*\left(t\right), T-t\right) \tag{18.4}$$

is satisfied. Condition (18.4) implies the existence of a vector $\xi\left(x_t^*, T-t\right) \in W_v\left(x^*\left(t\right), T-t\right)$ satisfying

$$\xi\left(x_0, T-t_0\right) = \varpi\left[\xi\left(x_0\left(\cdot\right), T-t_0\right); x^*\left(\cdot\right), t-t_0\right] + \xi\left(x_t^*, T-t\right).$$

Thus in the process of choosing $B\left(s\right)$, the vector of the gains to be obtained by the players at the remaining game interval $[t, T]$ has to satisfy:

$$\begin{aligned} \xi\left(x_t^*, T-t\right) &= \xi\left(x_0, T-t_0\right) - \varpi\left[\xi\left(x_0\left(\cdot\right), T-t_0\right); x^*\left(\cdot\right), t-t_0\right] \\ &= \int_{t}^{T} B\left(s\right) ds + q\left(x^*\left(T\right)\right), \end{aligned}$$

where

$$\sum_{j\in N} B_j\left(s\right) = \sum_{j\in N} g^j\left[s, x^*\left(s\right), u^*\left(s\right)\right] ds \quad \text{for} \quad t \leq s \leq T, \quad \text{and}$$

$$\xi\left(x_t^*, T-t\right) \in W_v\left(x^*\left(t\right), T-t\right).$$

By varying the vector $\varpi\left[\xi\left(x_0\left(\cdot\right), T-t_0\right); x^*\left(\cdot\right), t-t_0\right]$ restricted by the condition

$$\sum_{j\in N} \varpi_j\left[\xi\left(x_0\left(\cdot\right), T-t_0\right); x^*\left(\cdot\right), t-t_0\right] = \int_{t_0}^{t} \sum_{j\in N} g^j\left[s, x^*\left(s\right), u^*\left(s\right)\right] ds$$

the players ensure displacement of the set

$$\left(\varpi\left[\xi\left(x_0\left(\cdot\right), T-t_0\right); x^*\left(\cdot\right), t-t_0\right] \oplus W_v\left(x^*\left(t\right), T-t\right)\right)$$

in such a way that (18.3) is satisfied.

For any vector $B(\tau)$ satisfying condition (18.3) and (18.4) at each time instant $t_0 \leq t \leq T$ the players are guided by the same optimality principle that leads to the imputation $\xi\left(x_t^*, T-t\right) \in W_v\left(x^*\left(t\right), T-t\right)$ throughout the game, and hence the players have no reason to dispute the previous agreement. In general, it is fairly easy to see that the vectors $B(\tau)$ satisfying conditions (18.3) and (18.4) may not be unique. Thus there exist multiple sharing methods satisfying the condition of dynamic stability.

Dynamic instability of the solution of the cooperative differential game leads to abandonment of the original optimality principle generating this solution, because none of the imputations from the solution set $W_v(x_0, T-t_0)$ remain optimal until the game terminates. Therefore, the set $W_v(x_0, T-t_0)$ may be called a solution to the game $\Gamma_v(x_0, T-t_0)$ only if it is dynamically stable. Otherwise the game $\Gamma_v(x_0, T-t_0)$ is assumed to have no solution.

19 An analysis in pollution control

19.1. Consider the pollution model in Petrosyan and Zaccour [16]. Let N denote the set of countries involved in the game of emission reduction. Emission of player $i \in \{1, 2, \ldots, n\} = N$ at time $t\,(t \in [0, \infty))$ is denoted by $m_i(t)$. Let $x(t)$ denote the stock of accumulated pollution by time t. The evolution of this stock is governed by the following differential equation:

$$\frac{dx(t)}{dt} = \dot{x}(t) = \sum_{i \in I} m_i(t) - \delta x(t), \qquad \text{given } \ x(0) = x_0, \tag{19.1}$$

where δ denotes the natural rate of pollution absorption.

Each player seeks to minimize a stream of discounted sum of emission reduction cost and damage cost. The latter depends on the stock of pollution. We omit from now on the time argument when no ambiguity may arise, $C_i(m_i)$ denotes the emission reduction cost incurred by country i when limiting its emission to level m_i, and $D_i(x)$ its damage cost. We assume that both functions are continuously differentiable and convex, with $C'(m_i) < 0$ and $D'(x) > 0$. The optimization problem of country i is

$$\min_{m_i} J^i(m, x) = \int_0^\infty \exp(-rs)\left\{C_i(m_i(s)) + D_i(x(s))\right\} ds \tag{19.2}$$

subject to (19.1), where $m = (m_1, m_2, \ldots, m_n)$ and r is the common social discount rate.

This formulation is chosen due to the following motivations. First, a simple formulation of the ecological economics problem is chosen to put the emphasis on the cost sharing issue and the mechanism for allocating the total cost over time in a desirable manner. Second, obviously this model still captures one of the main ingredients of the problem, that is, each player's cost depends on total emissions and on inherited stock of pollution. Third, the convexity assumption and the sign assumption for the first derivatives of the two cost functions seem natural. Indeed, the convexity of $C_i(e_i)$ means that the marginal cost of reduction emissions is higher for lower levels of emissions. Fourth, for the sake of mathematical tractability it is assumed that all countries discount their costs at the same rate. Finally, again for tractability, it is worth noticing that this formulation implies that reduction of pollution can only be achieved by reducing emissions.

19.2. Decomposition over time of the Shapley value. A cooperative game methodology to deal with the problem of sharing the cost of emissions reduction is adopted. The steps are as follows:

(i) Computation of the characteristic function values of the cooperative game.

(ii) Allocation among the players of the total cooperative cost based on the Shapley value.

(iii) Allocation over time of each player's Shapley value having the property of being time-consistent.

The Shapley value is adopted as a solution concept for its fairness and uniqueness merits. The first two steps are classical and provide the individual total cost for each player as a lump sum. The third step aims to allocate over time this total cost in a time-consistent way. The definition and the computation of a time-consistent distribution scheme are dealt with below after introducing some notation.

Let the state of the game be defined by the pair (t, x) and denote by $\Gamma_v(x, t)$ the subgame starting at date t with stock of pollution x. Denote by $x^N(t)$ the trajectory of accumulated pollution under full cooperation (grand coalition N). In the sequel, we use $x^N(t)$ and x_t^N interchangeably. Let $\Gamma_v\left(x_t^N, t\right)$ denote a subgame that starts along the cooperative trajectory of the state. The characteristic function value for a coalition $K \subseteq N$ in subgame $\Gamma_v(x, t)$ is defined to be its minimal cost and is denoted $v(K; x, t)$. With this notation, the total cooperative cost to be allocated among the players is then $v(N; x, 0)$ which is the minimal cost for the grand coalition N and its characteristic function value in the game $\Gamma_v(x, 0)$. Let $\Phi^v(x, t) = [\Phi_1^v(x, t), \Phi_2^v(x, t), \ldots, \Phi_n^v(x, t)]$ denote the Shapley value in subgame $\Gamma_v(x, t)$. Finally, denote by $\beta_i(t)$ the cost to be allocated to Player i at instant of time t and $\beta(t) = (\beta_1(t), \ldots, \beta_n(t))$.

Let the vector $B(t) = [B_1(t), B_2(t), \ldots, B_n(t)]$ denote an imputation distribution procedure (IDP) so that

$$\Phi_1^v(x, 0) = \int_0^\infty \exp(-rt)\, \beta_i(t)\, dt, \qquad i = 1, \ldots, n. \tag{19.3}$$

The interpretation of this definition is obvious: a time function $B_i(t)$ qualifies as an IDP if it decomposes over time the total cost of Player i as given by the Shapley value component for the whole game $\Gamma_v(x,0)$, i.e. the sum of discounted instantaneous costs is equal to $\Phi_i^v(x,0)$.

The vector $B(t)$ is a time-consistent IDP if at $\left(x_t^N,t\right)$, $\forall t\in[0,\infty)$ the following condition holds

$$\Phi_i^v(x_0,0)=\int_0^t \exp(-r\tau)\,\beta_i(\tau)\,d\tau+\exp(-rt)\,\Phi_i^v\left(x_t^N,t\right). \tag{19.4}$$

To interpret condition (19.4), assume that the players wish to renegotiate the initial agreement reached in the game $\Gamma_v(x,0)$ at (any) intermediate instant of time t. At this moment, the state of the system is $x^N(t)$, meaning that cooperation has prevailed from initial time until t, and that each Player i would have been allocated a sum of discounted stream of monetary amounts given by the first right-hand side term. Now, if the subgame $\Gamma_v\left(x_t^N,t\right)$, starting with initial condition $x(t)=x^N(t)$, is played cooperatively, then Player i will get his Shapley value component in this game given by the second right-hand side term of (19.4). If what he has been allocated until t and what he will be allocated from this date onward sum up to his cost in the original agreement, i.e. his Shapley value $\Phi_i^v(x_0,0)$, then this renegotiation would leave the original agreement unaltered. If one can find an IDP $B(t)=[B_1(t),B_2(t),\ldots,B_n(t)]$ such that (19.4) holds true, then this IDP is time-consistent. An "algorithm" to build a time-consistent IDP in the case where the Shapley value is differentiable over time is suggested below. One can provide in such a case a simple expression for $B(t)=[B_1(t),B_2(t),\ldots,B_n(t)]$ having an economically appealing interpretation.

19.3. A solution algorithm. The first three steps compute the necessary elements to define the characteristic function given in the fourth step. In the next two, the Shapley value and the functions $B_i(t)$, $i=1,2,\ldots,n$, are computed.

Step 1: Minimize the total cost of the grand coalition.

The grand coalition solves a standard dynamic programming problem consisting of minimizing the sum of all players' costs subject to pollution accumulation dynamics, that is:

$$\min_{m_1,m_2,\ldots,m_n}\sum_{i\in N}\int_t^\infty \exp[-r(\tau-t)]\left\{C_i(m_i(\tau))+D_i(x(\tau))\right\}d\tau$$

$$\text{s.t. } \dot{x}(s)=\sum_{i\in N}m_i(s)-\delta x(s), \qquad x(t)=x^N(t).$$

Denote by $W(N,x,t)$ the (Bellman) value function of this problem, where the first entry refers to the coalition for which the optimization has been performed, here the grand coalition N. The outcome of this optimization is a vector of emission strategy $m^N\left(x^N(\tau)\right)=\left[m_1^N\left(x^N(\tau)\right),\ldots,m_n^N\left(x^N(\tau)\right)\right]$, where $x^N(\tau)$ refers to the accumulated pollution under the scenario of full cooperation (grand coalition).

Step 2. Compute a feedback Nash equilibrium.

Since the game is played over an infinite time horizon, stationary strategies will be sought. To obtain a feedback Nash equilibrium, assuming differentiability of the value function, the following Isaacs-Bellman equations (see 13.2) must be satisfied

$$r\overline{V}^i(x) = \min_{m_i}\left\{C_i(m_i) + D_i(x) + \overline{V}^i_x(x)\left[\sum_{i\in I} m_i - \delta x\right]\right\}, \qquad i \in N.$$

Denote by $m^*(x) = [m_1^*(x), m_2^*(x), \ldots, m_n^*(x)]$ any feedback Nash equilibrium of this noncooperative game. This vector can be interpreted as a business-as-usual emission scenario in the absence of an international agreement. It will be fully determined later on with special functional forms. For the moment, we need to stress that from this computation we can get the player's Nash outcome (cost-to-go) in game $\Gamma_v(x_0, 0)$, that we denote $V^i(0, x_0) = \overline{V}^i(x_0)$, and his outcome in subgame $\Gamma_v\left(x_t^N, t\right)$, that we denote $V^i\left(t, x_t^N\right) = \overline{V}^i\left(x_t^N\right)$.

Step 3: Compute outcomes for all remaining possible coalitions.

The optimal total cost for any possible subset of players containing more than one player and excluding the grand coalition (there will be $2^n - n - 2$ subsets) is obtained in the following way. The objective function is the sum of objectives of players in the subset (coalition) considered. In the objective and in the constraints of this optimization problem, we insert for the left-out players the values of their decision variables (strategies) obtained at Step 2, that is their Nash values. Denote by $W(K, x, t)$ the value function for coalition K. This value is formally obtained as follows

$$W(K, x, t) = \min_{m_i, i\in K} \sum_{i\in K}\left\{\int_t^\infty \exp\left[-r(\tau - t)\right]\left\{C_i(m_i(\tau)) + D_i(x(\tau))\right\} d\tau\right\}$$

$$\text{s.t. } \dot{x}(s) = \sum_{i\in N} m_i(s) - \delta x(s), \qquad x(t) = x^N(t).$$

$$m_j = m_j^N \qquad \text{for } j \in I\backslash K.$$

Step 4: Define the characteristic function.

The characteristic function $v(K; x, t)$ of the cooperative game is defined as follows:

$$v(\{i\}; x, t) = V^i(x, t) = \overline{V}^i(x), \qquad i = 1, \ldots, n;$$
$$v(K; x, t) = W(K; x, t), \qquad K \subseteq I.$$

Step 5: Compute the Shapley value.

Denote by $\Phi^v(x, t) = [\Phi_1^v(x, t), \Phi_2^v(x, t), \ldots, \Phi_n^v(x, t)]$ the Shapley value in game $\Gamma_v(x, t)$. Component i is given by

$$\Phi_i^v(x, t) = \sum_{K\ni i} \frac{(n-k)!\,(k-1)!}{n!}\left[W(K; x, t) - W(K\backslash\{i\}; x, t)\right],$$

where k denotes the number of players in coalition K. In particular, if cooperation is in force for the whole duration of the game, then the total cost of Player i would be given by his Shapley value in the game $\Gamma_v(x_0, 0)$, that is

$$\Phi_i^v(x_0, 0) = \sum_{K \ni i} \frac{(n-k)!\,(k-1)!}{n!} \left[W(K; x_0, 0) - W(K \setminus \{i\}; x_0, 0)\right].$$

Justification for the use of this nonstandard definition of the characteristic function is provided in 19.4.

Step 6: Define a time consistent IDP.

Allocate to Player i, $i = 1, \ldots, n$, at instant of time $t \in [0, \infty)$, the following amount:

$$B_i(t) = \Phi_i^v\left(x_t^N, t\right) - \frac{d}{dt}\Phi_i^v\left(x_t^N, t\right). \tag{19.5}$$

The formula (19.5) allocates at instant of time t to Player i a cost corresponding to the interest payment (interest rate times his cost-to-go under cooperation given by his Shapley value) minus the variation over time of this cost-to-go.

The following proposition shows that $\beta(t) = (\beta_1(t), \ldots, \beta_n(t))$, as given by (19.5) , is indeed a time-consistent IDP.

Proposition 1. *The vector $\beta(t) = (\beta_1(t), \ldots, \beta_n(t))$ where $\beta_i(t)$ is given by (19.5) is a time-consistent IDP.*

Proof. We first show that it is an IDP, that is

$$\int_0^\infty \exp(-rt)\,\beta_i(t)\,dt = \Phi_i^v(x_0, 0).$$

Multiply (19.5) by the discount factor $\exp(-rt)$ and integrate

$$\int_0^\infty \exp(-rt)\,\beta_i(t)\,dt = \int_0^\infty \exp(-rt)\left[r\Phi_i^v\left(x^N, t\right) - \frac{d}{dt}\Phi_i^v\left(x^N, t\right)\right]dt$$
$$= -exp(-rt)\,\Phi_i^v\left(x^N, t\right)\Big|_0^\infty = \Phi_i^v(x_0, 0).$$

Repeating the above integration for $\Phi_i^v\left(x_t^N, t\right)$, one can readily obtain

$$\Phi_i^v(x_0, 0) = \int_0^t \exp(-r\tau)\,\beta_i(\tau)\,d\tau + \exp(-rt)\,\Phi_i^v\left(x_t^N, t\right),$$

which satisfies the time-consistent property.

19.4. Rationale for the algorithm and the special characteristic function.

This section discusses the rationale for the algorithm proposed in 19.3. and the nonstandard definition of the characteristic function adopted.

As Petorsyan and Zaccour pointed out while formulating the solution algorithm, a central element in formal negotiation theories is the status quo, which gives what a player would obtain if negotiation fails. It is a measure of the strategic force of a player when acting alone. The same idea can be extended to subsets of players. To measure the strategic force of a subset of players (coalition), one may call upon the concept of characteristic function, a mathematical tool that is precisely intended to provide such measure. All classical cooperative game solutions (core, the Shapley value, etc.) use the characteristic function to select a subset of imputations that satisfy the requirements embodied in the solution concept adopted. For instance, the core selects the imputations that cannot be blocked by any coalition whereas Shapley value selects one imputation satisfying some axioms, among them fairness. If the set of imputations is not a singleton, the players may negotiate to select one of them. In a dynamic game, the computed imputations usually correspond to the payoffs (here the sum of discounted costs) for the whole duration of the game. In this case, an interesting problem emerges which is how to allocate these amounts over time. One basic requirement is that the distribution over time is feasible, that is, the amounts allocated to each player sum up to his entitled total share (see the definition of an IDP). Obviously, one may construct an infinite number of intertemporal allocations that satisfy this requirement, but not all these streams are conceptually and intuitively appealing. The approach pursued here is to decompose the total individual cost over time so that if the players renegotiate the agreement at any intermediate instant of time along the cooperative state trajectory, then they would obtain the same outcomes.

It is also noted that the computation of the characteristic function values is not standard. The assumption that left-out players ($I \backslash K$) stick to their feedback Nash strategies when the characteristic function value is computed for coalition K is made. Basically, there are few current options in game theory literature regarding this issue.

A one option is the one offered by Von Neumann and Morgenstern [3] where they assumed that the left-out players maximize the cost of the considered coalition. This approach, which gives the minimum guaranteed cost, does not seem to be the best one in our context. Indeed, it is unlikely that if a subset of countries form a coalition to tackle an environmental problem, then the remaining countries would form an anti-coalition to harm their efforts. For instance, the Kyoto Protocol permits that countries fulfill jointly their abatement obligation (this is called “joint implementation” in the Kyoto Protocol). The question is then why, e.g., the European Union would wish to maximize abatement cost of say a collation formed of Canada and USA if they wish to take the joint implementation option? We believe that the Von Neumann-Morgenstern determination of the characteristic function has a great historic value and the advantage of being easy to compute but would not suit well the setting we are dealing with.

A next option is to assume that the characteristic function value for a coalition is its Nash equilibrium total cost in the noncooperative game between this coalition and the other players acting individually or forming an anti-coalition.

One “problem” with this approach is computation. Indeed, this approach requires solving $2^n - 2$ dynamic equilibrium problems (that is, as many as the number of nonvoid coalitions and excluding the grand one). Here we solve only one equilibrium problem, all others being standard dynamic optimization problems. Therefore, the computational burden is not at all of the same magnitude since solving a dynamic feedback equilibrium problem is much harder than dealing with a dynamic optimization one.

Now, assume that Nash equilirbia exist for all partitions of the set of players (clearly this is far from being automatically guaranteed). First, recall that we aim to compute the Shapley value for Player i. This latter involves his marginal contributions, which are differences between values of the characteristic function of the form $v\left(K,S,t\right)-v\left(K\backslash\left\{i\right\},S,t\right)$. In the equilibrium approach, these values correspond to Nash outcomes of a game between players in coalition K and the remaining players in $I\backslash K$ (acting individually or collectively is not an issue at this level). If in any of the 2^n-2 equilibrium problems that must be solved the equilibrium is not unique, then we face an extremely hard selection problem. In our approach, the coalition computes its value with the assumption that left-out players will continue to adopt a non-cooperative emission strategies. In the event that our Step 2 gives multiple equilibria, we could still compute the Shapley value for each of them without having to face a selection problem.

Finally global environmental problems involve by definition all countries around the globe. Although few of them are heavy weights in the sense that their environmental policies can make a difference on the level of pollution accumulation, many countries can be seen as nonatomistic players. It is intuitive to assume that probably these countries will follow their business-as-usual strategy, i.e., by sticking to their Nash emissions, even when some (possibly far away) countries are joining effort.

20 Illustration with specific functional forms

Consider the following specification of (19.2) in which

$$
\begin{aligned}
&C_i\left(m_i\right)=\frac{\gamma}{2}\left[m_i-\overline{m}_i\right]^2, \;\; 0\leq m_i\leq\overline{m}_i, \;\; \gamma>0 \;\text{ and }\; i\in\left\{1,2,3\right\};\\
&D_i\left(x\right)=\pi x, \qquad \pi>0.
\end{aligned}
$$

Computation of optimal cost of grand coalition (Step 1).
The value function $W\left(N,x,t\right)$ must satisfy the Bellman equation

$$
rW\left(N,x,t\right)= \tag{20.1}
$$

$$
\min_{m_1,m_2,m_3}\left\{\sum_{i=1}^{3}\left(\frac{\gamma}{2}\left[m_i-\overline{m}_i\right]^2+\pi x\right)+W_x\left(N,x,t\right)\left[\sum_{i=1}^{3}m_i-\delta x\right]\right\}.
$$

Performing the indicated minimization in (20.1) yields

$$
m_i^N=\overline{m}_i-\frac{1}{\gamma}W_x\left(N,x,t\right), \qquad \text{for }\; i\in\left\{1,2,3\right\}.
$$

Substituting m_i^N in (20.1) and upon solving yields

$$W(N,x,t) = \overline{W}(N,x) = \frac{3\pi}{r(r+\delta)}\left\{\left[\sum_{i=1}^{3}\overline{m}_i - \frac{3^2\pi}{2\gamma(r+\delta)}\right] + rx\right\}, \text{ and} \tag{20.2}$$

$$m_i^N = \overline{m}_i - \frac{3\pi}{\gamma(r+\delta)}, \text{ for } i \in \{1,2,3\}. \tag{20.3}$$

The optimal trajectory of the stock of pollution can be obtained as

$$x^N(t) = \exp(-\delta t)\,x(0) + \frac{1}{\delta}\left\{\left[\sum_{i=1}^{3} m_i^N\right][1-\exp(-\delta t)]\right\}. \tag{20.4}$$

Computation of feedback Nash equilibrium (Step 2).

To solve a feedback Nash equilibrium for the noncooperative game (19.1)–(19.2), we follow 13.2 and obtain the Bellman equation

$$r\overline{V}^i(x) = \min_{m_i}\left\{\frac{\gamma}{2}[m_i - \overline{m}_i]^2 + \pi x + \overline{V}^i_x(x)\left[\sum_{\substack{j\in[1,2,3]\\ i\neq j}} m_j^* + m_i - \delta x\right]\right\},$$
$$\text{for } i \in \{1,2,3\}. \tag{20.5}$$

Performing the indicated minimization yields

$$m_i^* = \overline{m}_i - \frac{1}{\gamma}\overline{V}^i_x(x), \qquad \text{for } i \in \{1,2,3\}. \tag{20.6}$$

Substituting (20.6) into (20.5) and upon solving yield

$$\overline{V}^i(x) = \frac{\pi}{r(r+\delta)}\left\{\frac{\pi}{2\gamma(r+\delta)} + \sum_{i=1}^{3}\overline{m}_i - \frac{3\pi}{\gamma(r+\delta)} + rx\right\},$$
$$\text{for } i \in \{1,2,3\}. \tag{20.7}$$

The Nash equilibrium feedback level of emission can then be obtained as:

$$m_i^* = \overline{m}_i^* - \frac{\pi}{\gamma(r+\delta)}, \qquad \text{for } i \in \{1,2,3\}. \tag{20.8}$$

The difference between Nash equilibrium emissions and those obtained for the grand coalition is that player takes into account the sum of marginal damage costs of all coalition members and not only his own one.

Computation of optimal cost for intermediate coalitions (Step 3).

The value function $W(K,x,t)$ for any coalition K of two players must satisfy the following Bellman equation

$$rW(K,x,t) = \tag{20.9}$$

$$\min_{m_1, i \in K} \left\{ \sum_{i \in K} \left(\frac{\gamma}{2} \left[m_i - \overline{m}_i \right]^2 + \pi x \right) + W_x (K, x, t) \left[\sum_{i \in K} m_i + m_j^* - \delta x \right] \right\},$$

where $j \notin K$.

Following similar procedure adopted for solving for the grand coalition, one can obtain:

$$W(K, x, t) = \overline{W}(K, x) = \tag{20.10}$$

$$\frac{2\pi}{r(r+\delta)} \left\{ \sum_{i \in K} \overline{m}_i - \frac{4\pi}{2\gamma(r+\delta)} - \frac{\pi}{\gamma(r+\delta)} + rx \right\}.$$

The corresponding emission of coalition K is:

$$m_i^K = \overline{m}_i - \frac{2\pi}{\gamma(r+\delta)}, \qquad i \in K. \tag{20.11}$$

Definition of the characteristic function (Step 4).

The characteristic function values are given by

$$v(\{i\}; x, t) = V^i(x, t) = \overline{V}^i(x) =$$

$$\frac{\pi}{r(r+\delta)} \left\{ \frac{\pi}{2\gamma(r+\delta)} + \sum_{i=1}^{3} \overline{m}_i - \frac{3\pi}{\gamma(r+\delta)} + rx \right\}, \quad i = 1, 2, 3;$$

$$v(K; x, t) = W(K, x, t) = \overline{W}(K, x) =$$

$$\frac{2\pi}{r(r+\delta)} \left\{ \sum_{i \in K} \overline{m}_i - \frac{4\pi}{2\gamma(r+\delta)} - \frac{\pi}{\gamma(r+\delta)} + rx \right\},$$

$$K \subseteq \{1, 2, 3\}.$$

Computation of the Shapley value (Step 5).

Assuming symmetric $\overline{m}_i$, the Shapley value of the game can be expressed as

$$\Phi_i^v(x, t) = \sum_{K \ni i} \frac{(n-k)!\,(k-1)!}{n!} \left[v(K; x, t) - v(K \backslash \{i\}; x, t) \right] \tag{20.12}$$

$$= \frac{1}{2r(r+\delta)} \left\{ 2\pi \left(\sum_{i=1}^{3} \overline{m}_i + \rho S \right) - \frac{9\pi^2}{\gamma(r+\delta)} \right\},$$

$$i = 1, 2, 3.$$

Computation of IDP functions (Step 6).

To provide a allocation that sustains the Shapley value $\Phi_i^v(x, t)$ over time along the optimal trajectory $x^N(t)$ in (20.4), we recall from (19.5) that the IDP functions are given by

$$B_i(t) = \Phi_i^v\left(x_t^N, t\right) - \frac{d}{dt} \Phi_i^v\left(x_t^N, t\right).$$

Straightforward calculations lead to

$$B_i(t) = \pi x^N(t) + \frac{9\pi^2}{2\gamma(r+\delta)^2}, \qquad i = 1,2,3. \tag{20.13}$$

To verify that $B_i(t)$ indeed brings about the Shapley value of Player i we note that

$$\Phi_i^v\left(x^N, 0\right) = \frac{1}{2r(r+\delta)}\left\{2\pi\left(\sum_{i=1}^{3}\overline{m}_i + rx(0)\right) - \frac{9\pi^2}{\gamma(r+\delta)}\right\},$$
$$i = 1,2,3. \tag{20.14}$$

Multiply both sides of (20.13) by the discount factor and integrate

$$\int_0^\infty \exp(-rt)\, B_i(t)\, dt = \int_0^\infty \exp(-rt)\left[\pi x^N(t) + \frac{9\pi^2}{2\gamma(r+\delta)^2}\right] dt,$$
$$i = 1,2,3, \tag{20.15}$$

where from (20.3)–(20.4)

$$x^N(t) = \exp(-\delta t)\, x(0) + \frac{1}{\delta}\left\{\left[\sum_{j=1}^{3}\left(\overline{m}_j - \frac{3\pi}{2\gamma(r+\delta)}\right)\right][1-\exp(-\delta t)]\right\}.$$

Substituting $x^N(t)$ into (20.15) yields:

$$\begin{aligned}
&\int_0^\infty \exp(-rt)\,\beta_i(t)\,dt = \\
&\quad \int_0^\infty \exp[-(r+\delta)t]\,\pi x_0 dt + \int_0^\infty \exp(-rt)\,\frac{\pi}{\delta}\left(\sum_{i=1}^{3}\overline{m}_i - \frac{9\pi}{\gamma(r+\delta)}\right) dt \\
&\quad + \int_0^\infty \exp[-(r+\delta)t]\,\frac{\pi}{\delta}\left(\sum_{i=1}^{3}\overline{m}_i - \frac{9\pi}{\gamma(r+\delta)}\right) dt \\
&\quad + \int_0^\infty \exp(-rt)\,\frac{9\pi^2}{2\gamma(r+\delta)^2} dt.
\end{aligned}$$

Upon integrating

$$\begin{aligned}
&\int_0^\infty \exp(-rt)\,\beta_i(t)\,dt = \\
&\quad \frac{1}{2r(r+\delta)}\left\{2\pi\left(\sum_{i=1}^{3}\overline{m}_i + rx(0)\right) - \frac{9\pi^2}{\gamma(r+\delta)}\right\} = \Phi_i^v(x_0, 0), \\
&\quad \text{for } \; i = 1,2,3.
\end{aligned}$$

References

[1] Isaacs, R. *Differential Games*, p. 384. Wiley, N.Y., 1965.

[2] Petrosjan, L. A. *Differential games of pursuit*, p. 325. World Scientific, Singapore, 1993.

[3] von Neumann, J. and O. Morgenstern. *Theory of Games and Economic Behavior*, p. 625. Wiley, N.Y., 1944.

[4] Bellman, R. *Dynamic Programming.* Princeton Univ. Press, NJ, 1957.

[5] Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mishchenko. *The Mathematical Theory of Optimal Processes*, Interscience Publishers, New York, NY, 1962.

[6] Basar, T. and I. Olsder. *Dynamic Noncooperative Game Theory*, p. 233. Acad. Press, London, 1984.

[7] Yeung, D. W. K. and L. A. Petrosyan. *Cooperative Stochastic Differential Games.* Springer, New York, 2006.

[8] Jorgensen, S. and G. Sorger. *Journal of Optimization Theory and Applications*, 64, 1990.

[9] Nash, J. F. *Econometrica*, 21, 1953.

[10] Haurie, A. *Journal of Optimization Theory and Application*, 18, 1976.

[11] Tsytsui, S. and K. J. Mino. *Econom. Theory,*, 52, 1990.

[12] Shapley, L. S. In H. Kuhn and A. Tucker, editors, *Contributions to the Theory of Games, II.* Princeton Univ. Press, Princeton, 1953.

[13] Petrosjan, L. A. *Vestnik of the Leningrad State University*, 13, 1977.

[14] Petrosjan, L. A. and N. N. Danilov. *Vestnik of the Leningrad State University*, 1, 1979.

[15] Yeung, D. W. K. and L. A. Petrosyan. In *Yanovskaya E.B. (ed) International Conference on Logic, Game Theory and Social Choice. St Petersburg State Unviersity*, 2001.

[16] Petrosjan, L. A. and G. Zaccour. *J.Econ, Dyn. Control*, 27(3), 2003.