

УДК 517.977.8, 519.83

ББК 22.18

# УСТОЙЧИВАЯ КООПЕРАЦИЯ В ДИФФЕРЕНЦИАЛЬНЫХ ИГРАХ СО СЛУЧАЙНОЙ ПРОДОЖИТЕЛЬНОСТЬЮ

ЕКАТЕРИНА В. ШЕВКОПЛЯС

Факультет прикладной математики –  
процессов управления

Санкт-Петербургский государственный университет  
198504, Санкт-Петербург, Университетский пр., 35  
e-mail: [katya\\_shev@mail.ru](mailto:katya_shev@mail.ru)

Работа посвящена изучению проблемы динамической устойчивости кооперативных решений, впервые сформулированной Петросяном Л.А. в 1977 г. для дифференциальных игр с предписанной продолжительностью. В данной работе рассматривается модификация дифференциальной игры с предписанной продолжительностью, а именно, предполагается, что игра заканчивается в некоторый случайный момент времени. Кроме того, в качестве кооперативного решения используется вектор Шепли. Для такой постановки задачи сформулировано понятие процедуры распределения дележа, и получена аналитическая формула для проверки динамической устойчивости вектора Шепли. Также в работе изучается условие защиты от иррационального поведения участников (условие Д.Янга, 2006) и предложен механизм проверки выполнения этого свойства, основанный на процедуре распределения дележа. Теоретические результаты демонстрируются на примере дифференциальной игры разработки невозобновляемых ресурсов.

*Ключевые слова:* динамическая устойчивость, устойчивая кооперация, защита от иррационального поведения, разработка невозобновляемых ресурсов, дифференциальная игра со случайной продолжительностью.

## 1. Введение

Исследование в данном направлении было начато в 1998 г. под руководством Л.А. Петросяна, которым была сформулирована тема работы – «Кооперативные дифференциальные игры со случайной продолжительностью». Л.А. Петросян предложил изучить проблему динамической устойчивости принципов оптимальности, сформулированную им в конце 1970–х годов для дифференциальных игр с предписанной продолжительностью (см. [2], [1]), для нового класса дифференциальных игр, а именно дифференциальных игр со случайной продолжительностью. В 1966 году в работе [4] изучалась антагонистическая игра преследования двух лиц с терминальными выигрышами в последний момент времени, который являлся случайной величиной с известной функцией распределения. Л. А. Петросян предложил автору рассмотреть общую постановку дифференциальной игры, заканчивающейся в случайный момент времени. Кроме того, выигрыши игроков в новой задаче предполагались интегральными.

Данное исследование переросло в диплом, а затем в кандидатскую диссертацию. Кроме проблемы динамической устойчивости, изученной в работе [5], было выведено уравнение типа Беллмана, позволяющее находить управления с обратной связью для задачи со случайной продолжительностью [6].

В данной работе собраны, переработаны и дополнены материалы, касающиеся проблемы устойчивой кооперации в дифференциальных играх со случайной продолжительностью.

В разделе 2 дается определение игры. В разделе 3 на примере вектора Шепли формулируется проблема динамической устойчивости принципов оптимальности для дифференциальных игр со случайной продолжительностью. В разделе 4 формулируется условие защиты от иррационального поведения участников. В разделе 5 приводится пример дифференциальной игры разработки невозобновляемых

ресурсов, причем момент окончания игры является случайной величиной, распределенной по закону Вейбулла.

## 2. Модель игры

В исследованиях в области дифференциальных игр, как правило, изучаются дифференциальные игры с предписанной продолжительностью. Это означает, что игра развивается во времени на фиксированном временном промежутке  $[t_0, T]$ , причем момент окончания игры  $T$  известен заранее.

В данной работе изучается модификация дифференциальной игры  $n$  лиц с предписанной продолжительностью, а именно, предполагается, что игра развивается на промежутке  $[t_0, T]$ , где  $T$  — случайная величина с известной функцией распределения  $F(t)$ ,  $t \in [t_0, \infty)$  [4], [5]. Таким образом, постановка дифференциальной игры со случайной продолжительностью является обобщением постановки дифференциальной игры с предписанной продолжительностью [2].

Итак, рассмотрим дифференциальную игру  $n$  лиц  $\Gamma(x_0, t_0)$  со случайной продолжительностью  $(T - t_0)$  и начальным состоянием  $x_0$  [5]. Динамика игры задается системой обыкновенных дифференциальных уравнений:

$$\begin{aligned} \dot{x} &= g(x, u_1, \dots, u_n), \quad x \in R^m, u_i \in U \subseteq \text{comp } R^l, \\ x(t_0) &= x_0. \end{aligned} \quad (2.1)$$

Предполагаем, что вектор-функция  $g(x, u_1, \dots, u_n)$  непрерывна на  $R^m \times U_1 \times \dots \times U_n$ , удовлетворяет условию Липшица по  $x$  и существует  $\lambda > 0$ , такое что  $\|g(x, u_1, \dots, u_n)\| \leq \lambda(1 + \|x\|)$  для всех  $x \in R^m$ ,  $u_i \in U$  [2].

Игра начинается в момент  $t_0$  из состояния  $x_0$ , однако, момент ее окончания не фиксирован заранее, а является реализацией некоторой случайной величины  $T$ . Будем полагать, что для случайной величины  $T$  задана функция распределения  $F(t)$ , которая определена при  $t \in [t_0, \infty)$  и удовлетворяет условию нормировки:

$$\int_{t_0}^{\infty} dF(t) = 1.$$

Кроме того, далее будем предполагать существование функции плотности  $f(t) = F'(t)$  для случайной величины  $T$ .

Функция «мгновенного» выигрыша игрока  $i$  в момент времени  $\tau$ ,  $\tau \in [t_0, \infty)$  зависит от времени  $\tau$  и фазовой переменной  $x(t_0, x_0, u(\cdot))$ , где  $u(\cdot) = \{u_1(\cdot), \dots, u_n(\cdot)\}$  —  $n$ -набор допустимых программных управлений игроков. Под допустимыми программными управлениями понимаются измеримые по Лебегу программные управления  $u_i(\cdot) : t \rightarrow u_i(t) \in R^l$ , такие что  $u_i(t) \in U$ . Для краткости обозначим мгновенную функцию выигрыша как  $h_i(\tau, x(\tau), u(\tau))$ .

Предполагается, что  $h_i$  являются непрерывными функциями на  $R^m$ . Тогда ожидаемый интегральный выигрыш игрока  $i$  имеет вид:

$$K_i(x_0, t_0, u_1, \dots, u_n) = \int_{t_0}^{\infty} \left[ \int_{t_0}^t h_i(\tau, x(\tau), u(\tau)) d\tau \right] f(t) dt, \quad i = 1, \dots, n. \quad (2.2)$$

Известно, что по теореме Фубини–Тонелли о перестановке интегралов при требовании неотрицательности функций  $h_i(\tau, x(\tau), u(\tau))$ , функционалы вида (2.2) в форме повторных интегралов могут быть сведены к функционалам, имеющим стандартный для динамического программирования вид. Если же нельзя гарантировать неотрицательность функции мгновенного выигрыша  $h_i$ , но при этом выполнено условие абсолютной сходимости кратного интеграла

$$\iint_{[0, +\infty) \times [0, +\infty)} |f(t)h_i(\tau, x(\tau), u(\tau))| dt d\tau < +\infty,$$

то все равно можно использовать теорему Фубини–Тонелли и также изменить порядок интегрирования [6].

Тогда имеем:

$$\begin{aligned} K_i(x_0, t_0, u_1, \dots, u_n) &= \int_{t_0}^{\infty} \left[ \int_{t_0}^t h_i(\tau, x(\tau), u(\tau)) d\tau \right] f(t) dt = \\ &= \int_{t_0}^{\infty} (1 - F(\tau)) h_i(\tau, x(\tau), u(\tau)) d\tau. \end{aligned} \quad (2.3)$$

При развитии игры во времени в некоторый промежуточный момент  $\vartheta$ ,  $\vartheta \in (t_0; \infty)$ , игроки попадают в подыгру  $\Gamma(x(\vartheta))$  с начальным состоянием  $x(\vartheta) = x$ . Очевидно, что игра может и закончиться до момента  $\vartheta$  с вероятностью  $F(\vartheta)$ , а вероятность продолжить игру

после момента  $\vartheta$  равна  $(1 - F(\vartheta))$ . Тогда под выигрышем в подыгре  $\Gamma(x(\vartheta))$  будем понимать условное математическое ожидание выигрыша, а именно:

$$K_i(x, \vartheta, u_1, \dots, u_n) = \frac{1}{1 - F(\vartheta)} \int_{\vartheta}^{\infty} (1 - F(\tau)) h_i(\tau, x(\tau), u(\tau)) d\tau. \quad (2.4)$$

### 3. Проблема динамической устойчивости кооперативных решений

#### 3.1. Кооперативная игра

Рассмотрим кооперативную форму игры  $\Gamma(x_0, t_0)$ . Перед началом игры игроки договариваются об использовании ими таких допустимых программных управлений, которые будут максимизировать совокупный ожидаемый выигрыш игроков:

$$\begin{aligned} \max_{u_1, \dots, u_n} \sum_{i=1}^n K_i(x_0, t_0, u_1, \dots, u_n) = \\ = \max_{u_1, \dots, u_n} \sum_{i=1}^n \int_{t_0}^{\infty} (1 - F(\tau)) h_i(\tau, x(\tau), u(\tau)) d\tau. \end{aligned} \quad (3.1)$$

Управления  $\{u_1^*(t), \dots, u_n^*(t)\}$ , доставляющие максимум (3.1), будем называть оптимальными, а траекторию  $x^*(t)$ , соответствующую оптимальным управлениям, – условно-оптимальной. Дальнейшее изложение предполагает, что условно-оптимальная траектория единственна.

Важным вопросом, который решается в кооперативной теории игр, является вопрос о выборе конкретного принципа оптимальности как справедливого способа раздела заработанного совместными усилиями выигрыша. Однако мы не будем останавливаться на данном аспекте и для определенности далее будем полагать, что игроки договорились использовать вектор Шепли для раздела суммы (3.1):

$$Sh_i = \sum_{\substack{S \subseteq N \\ i \in S}} \frac{(n-s)!(s-1)!}{n!} [V(x_0, t_0, S) - V(x_0, t_0, S \setminus \{i\})], \quad i = 1, \dots, n. \quad (3.2)$$

Характеристической функцией  $V(x_0, t_0, S)$ ,  $S \subseteq N$ ,  $|N| = n$ , в игре  $\Gamma(x_0, t_0)$  будем называть функцию множества, удовлетворяющую

условиям:

$$V(x_0, t_0, \emptyset) = 0, \quad (3.3)$$

$$V(x_0, t_0, S_1 \cup S_2) \geq V(x_0, t_0, S_1) + V(x_0, t_0, S_2), \quad \forall S_1, S_2 \subset N, S_1 \cap S_2 = \emptyset,$$

где  $V(x_0, t_0, S)$  интерпретируется как максимальное значение математического ожидания выигрыша, которое может обеспечить себе коалиция  $S$  в игре  $\Gamma(x_0, t_0)$ , действуя самостоятельно. Следовательно,  $V(x_0, t_0, N)$  определяется по формуле (3.1).

Аналогичным образом определяется характеристическая функция в подыгре  $\Gamma(x_t, t)$ , начинающейся в момент времени  $t$  из состояния  $x_t$ . Отметим, что под характеристической функцией в подыгре понимается максимальное значение условного математического ожидания выигрыша, где условием является «дожитие» до момента  $t$ . Следовательно,  $V(x_t, t, N)$  определяется как  $\max_u \sum_{i=1}^n K_i(x_t, t, u_1, \dots, u_n)$ .

Не будем подробно останавливаться на способе построения характеристической функции  $V(x_0, t_0, S)$  в дифференциальных играх со случайной продолжительностью. Данный вопрос был подробно исследован в работе [10]. Характеристическая функция может быть построена как стандартным способом — с использованием значения вспомогательной антагонистической игры  $\Gamma_{S, N \setminus S}$ , так и каким-либо другим образом при условии проверки выполнения свойства супераддитивности (3.1). В работе [9] предлагается следующий алгоритм построения характеристической функции  $V(x_0, t_0, S)$  в дифференциальных играх: коалиция  $S$  максимизирует свой выигрыш, а остальные игроки, не входящие в  $S$ , используют равновесные по Нэшу стратегии.

Итак, предположим, что игроки в начальный момент  $t_0$  договорились использовать оптимальные управления  $\{u_1^*, \dots, u_n^*\}$ , чтобы получить ожидаемый выигрыш (3.1), а затем разделить его согласно принципу оптимальности (вектору Шепли). Тогда, как и в любой дифференциальной игре, возникает вопрос о реализуемости вектора Шепли во времени или *проблема динамической устойчивости* выбранного игроками принципа оптимальности [2], [1].

### 3.2. Динамическая устойчивость вектора Шепли

Развитию игры во времени соответствует движение вдоль условно-оптимальной траектории  $x^*(t)$ , на которой по определению игроки получают наибольший ожидаемый дележ. Однако движение вдоль оптимальной траектории еще не обеспечивает сохранение кооперации. Действительно, при движении вдоль  $x^*(t)$  игроки попадают в подыгры с текущими начальными состояниями, в которых один и тот же игрок имеет различные возможности. Следовательно, в некоторый момент  $\vartheta$  может возникнуть ситуация, когда решение текущей игры  $\Gamma(x^*(\vartheta), \vartheta)$  будет неоптимальным в смысле первоначально выбранного принципа оптимальности (в нашем случае — вектора Шепли). Тогда перед игроками встанет вопрос о целесообразности придерживаться далее намеченного перед началом игры соглашения действовать «совместно оптимально». Последнее будет означать динамическую неустойчивость вектора Шепли и, соответственно, самого движения по траектории  $x^*(t)$ .

**Определение 3.1.** Рассмотрим вектор-функцию  $\beta(t) = \{\beta_i(t) \geq 0\}_{i=1, \dots, n}$ , такую что компоненты вектора Шепли  $Sh = \{Sh_i\}_{i=1, \dots, n}$  в игре  $\Gamma(x_0, t_0)$  представимы в виде

$$Sh_i = \int_{t_0}^{\infty} (1 - F(t))\beta_i(t)dt, \quad i = 1, \dots, n. \quad (3.4)$$

Вектор-функцию  $\beta(t) = \{\beta_i(t)\}$  будем называть процедурой распределения дележа (ПРД).

Определение ПРД для игр с фиксированной продолжительностью было введено в работе [1]. В нашей постановке ПРД определяет правило, по которому компоненты ожидаемого дележа распределяются во времени  $[t_0, \infty)$ . Отметим, что позднее в работах Петросяна Л.А. требование неотрицательности компонент  $\beta_i(t)$ ,  $\forall t \geq t_0$ , было отменено (см., например, [9], [3]), однако в данной работе будем придерживаться изначальной формулировки.

**Определение 3.2.** Будем называть вектор Шепли  $\{\bar{Sh}_i\}$  динамически устойчивым вектором Шепли, если существует такая ПРД  $\{\beta_i(t) \geq 0\}$ ,  $t \in [t_0, \infty)$ , что вектор  $\bar{Sh}^\vartheta = \{\bar{Sh}_i^\vartheta\}$ ,  $\forall \vartheta \in [t_0, \infty)$ ,

вычисленный по формуле

$$\bar{S}h_i^\vartheta = \frac{1}{(1 - F(\vartheta))} \int_{\vartheta}^{\infty} (1 - F(t))\beta_i(t)dt, \quad i = 1, \dots, n, \quad (3.5)$$

также является вектором Шепли в соответствующей подыгре  $\Gamma(x^*(\vartheta), \vartheta)$ ,  $\vartheta \in [t_0, \infty)$ .

Определение 3.2 означает, что при распределении дележа  $\{\bar{S}h_i\}$  во времени при помощи выплат согласно ПРД  $\{\beta_i(\tau)\}$ , в каждый текущий момент времени  $\vartheta$ ,  $\vartheta \in [t_0, \infty)$ , ожидаемый дележ  $\{\bar{S}h_i^\vartheta\}$  в оставшейся подыгре  $\Gamma(x^*(\vartheta), \vartheta)$  также является вектором Шепли. Таким образом, игроки не имеют оснований для нарушения соглашения о кооперации, заключенного перед началом игры. Последнее означает *динамическую устойчивость* или, согласно терминологии в англоязычной литературе, *временную состоятельность* выбранного принципа оптимальности (вектора Шепли).

Принимая во внимание (3.5), заметим, что динамически устойчивый вектор Шепли  $\{\bar{S}h_i\}$  в игре  $\Gamma(x_0, t_0)$  может быть представлен в следующем виде:

$$\bar{S}h_i = \int_{t_0}^{\vartheta} (1 - F(\tau))\beta_i(\tau)d\tau + (1 - F(\vartheta))\bar{S}h_i^\vartheta, \quad \forall \vartheta \in [t_0, \infty), \quad i = 1, \dots, n. \quad (3.6)$$

Первое слагаемое в (3.6) соответствует сумме, которую игрок получит при движении вдоль условно-оптимальной траектории  $x^*(t)$  при  $t \in [t_0, \vartheta]$ . Второе слагаемое является математическим ожиданием выигрыша в подыгре  $\Gamma(x^*(\vartheta), \vartheta)$  при условии, что игра не закончилась до момента  $\vartheta$ .

Дифференцируя (3.6) по  $\vartheta$ , получаем аналитическую формулу для вычисления ПРД:

$$\beta_i(\vartheta) = \frac{f(\vartheta)}{(1 - F(\vartheta))} \bar{S}h_i^\vartheta - (\bar{S}h_i^\vartheta)', \quad \vartheta \in [t_0, \infty), \quad i = 1, \dots, n. \quad (3.7)$$

Очевидно, что в игре  $\Gamma(x_0, t_0)$  мы всегда можем распределить во времени вектор Шепли  $\{Sh_i\}$ , используя формулу для выплат (3.7). Однако в общем случае нельзя гарантировать неотрицательности компонент  $\beta_i(\vartheta)$ ,  $\forall \vartheta \in [t_0, \infty)$ . Следовательно, в рамках Определения



3.2 вектор Шепли не является динамически устойчивым в общем случае. Алгоритм проверки динамической устойчивости вектора Шепли является следующим: вычислить компоненты ПРД по формуле (3.7) и проверить выполнение условия  $\{\beta_i(\vartheta) \geq 0\}$ ,  $\forall \vartheta \in [t_0, \infty)$ . Если неотрицательность выполнена, то вектор Шепли  $\{Sh_i\}$ , распределенный во времени в игре  $\Gamma(x_0, t_0)$  согласно (3.7), является динамически устойчивым.

В противном случае вектор Шепли не является динамически устойчивым принципом оптимальности. Тогда, при выполнении свойства неотрицательности функции мгновенного выигрыша  $h_i(\tau, x(\tau), u(\tau)) \geq 0$ ,  $i = 1, \dots, n$ , для получения нового динамически устойчивого (регуляризованного) принципа оптимальности на основе первоначально выбранного игроками динамически неустойчивого принципа оптимальности, может быть использована новая процедура распределения дележа, а именно:

$$\bar{\beta}_i(\vartheta) = \frac{Sh_i^\vartheta \sum_{i=1}^n h_i(\vartheta, x^*(\vartheta), u^*(\vartheta))}{V(x^*(\vartheta), \vartheta, N)}, \quad \vartheta \in [t_0, \infty). \quad (3.8)$$

На основе  $\bar{\beta}_i(\vartheta) \geq 0$ ,  $i = 1, \dots, n$ , можно сформировать так называемый регуляризованный вектор Шепли по формуле (3.4), который будет удовлетворять условию (3.6) (см. [5]).

### 3.3. Формула для вычисления ПРД

Заметим, что множитель  $\frac{f(\vartheta)}{1-F(\vartheta)}$  (функция плотности для случайного момента окончания игры  $T$  при условии, что игра не закончилась до момента  $\vartheta$ ), появившийся в правой части уравнения (3.7), является стандартной для теории надежности функцией интенсивности отказов:

$$\lambda(t) = \frac{f(t)}{(1 - F(t))}. \quad (3.9)$$

Тогда, учитывая обозначение (3.9), выражение для ПРД (3.7) может быть переписано в следующем виде:

$$\beta_i(\vartheta) = \lambda(\vartheta)S\bar{h}_i^\vartheta - (S\bar{h}_i^\vartheta)', \quad \vartheta \in [t_0, \infty), \quad i = 1, \dots, n. \quad (3.10)$$

Кроме того, в данной терминологии  $(1 - F(\vartheta))$  является функцией «дожития» до момента  $\vartheta$ , для которой справедлива формула:

$$1 - F(\vartheta) = e^{-\int_{t_0}^\vartheta \lambda(t)dt}.$$

Тогда справедливо следующее представление для интегрального выигрыша игрока (2.3):

$$\int_{t_0}^{\infty} (1-F(\tau))h_i(\tau, x(\tau), u(\tau))d\tau = \int_{t_0}^{\infty} h_i(\tau, x(\tau), u(\tau))e^{-\lambda(\tau-t_0)}d\tau, \quad (3.11)$$

а динамически устойчивый вектор Шепли (3.6) может быть представлен следующим образом:

$$\bar{S}h_i = \int_{t_0}^{\vartheta} \beta_i(\tau)e^{-\lambda(\tau-t_0)}d\tau + e^{-\lambda(\vartheta-t_0)}\bar{S}h_i^{\vartheta}. \quad (3.12)$$

Следовательно, задача со случайной продолжительностью является не только обобщением задачи с предписанной продолжительностью, но и обобщением задачи на бесконечном временном промежутке с дисконтированием мгновенных выигрышей игроков (подробнее см. [6]).

Очевидно, что последнее утверждение распространяется и на результаты относительно вычисления ПРД в играх с предписанной продолжительностью [1], [2] и в играх с дисконтированием с бесконечным временным горизонтом [9]. При  $f(\vartheta) = 0$  ( $\lambda(\vartheta) = 0$ ) фактически рассматривается детерминированный случай, однако необходимо дополнительно требовать сходимость несобственных интегралов, соответствующих интегральным выигрышам игроков. Тогда из (3.10) получаем следующее выражение для ПРД:

$$\beta_i(\vartheta) = -(\bar{S}h_i^{\vartheta})', \quad i = 1, \dots, n,$$

которое было получено в работе [1].

Кроме того, функция интенсивности отказов  $\lambda(t)$  является константой тогда и только тогда, когда случайная величина  $T$  распределена по экспоненциальному закону:

$$f(t) = \lambda e^{-\lambda(t-t_0)}; \quad F(t) = 1 - e^{-\lambda(t-t_0)}, \quad \forall t \geq t_0;$$

$$\frac{f(t)}{1 - F(t)} = \lambda.$$

Тогда интегральный выигрыш (3.11) в точности совпадает с интегральным выигрышем для постановки задачи с бесконечным временным горизонтом и дисконтированием мгновенных выигрышей экспоненциальной функцией с дискаунт-фактором  $\lambda$ :

$$\int_{t_0}^{\infty} h_i(\tau, x(\tau), u(\tau))e^{-\lambda(\tau-t_0)}d\tau,$$

а формула для ПРД (3.10) в точности совпадает с формулой, полученной в работе [9]:

$$\beta_i(\vartheta) = \lambda Sh_i^\vartheta - (Sh_i^\vartheta)'. \quad (3.13)$$

Отметим, что этот результат был получен еще в работе [5].

#### 4. Защита от иррационального поведения игроков

Проблема динамической устойчивости принципов оптимальности, которая была изучена выше на примере проблемы динамической устойчивости вектора Шепли, предполагал рациональное поведение всех участников дифференциальной игры. В этом случае, используя механизм ПРД, можно было добиться того, чтобы у *рациональных* участников не возникло мотивации нарушить соглашение о кооперации.

Однако в настоящее время в теории игр начало уделяться внимание и моделированию *иррациональных* поступков игроков.

Предположим, что в некоторый момент времени  $\vartheta$  игрок иррационально нарушает соглашение о кооперации, что приводит к распаду большой коалиции  $N$ . Для определенности будем полагать, что изначально игроки договорились разделить заработанный совместными усилиями максимальный ожидаемый выигрыш согласно вектору Шепли. Распад коалиции будет означать, что вектор Шепли не реализуем во времени. Следовательно, требуется выполнение некоторого условия для защиты от иррационального поведения игроков. Это условие впервые было сформулировано в работе [11] для игр с предписанной продолжительностью.

Итак, для кооперативной дифференциальной игры со случайной продолжительностью условие защиты от иррационального поведения может быть сформулировано следующим образом:

$$V(x_0, t_0, \{i\}) \leq \int_{t_0}^{\vartheta} (1 - F(\tau)) \beta_i(\tau) d\tau + (1 - F(\vartheta)) V(x^*(\vartheta), \vartheta, \{i\}),$$

$$i = 1, \dots, n, \quad \theta \in [t_0; \infty). \quad (4.1)$$

Условие (4) означает, что даже если в некоторый момент времени  $\vartheta$  игрок (либо группа игроков) иррационально нарушил соглашение

действовать совместно оптимально, то выбором ПРД можно гарантировать, что ожидаемый выигрыш игрока  $i$  во всей игре все равно будет не меньше, чем выигрыш в случае, если бы игрок с самого начала действовал самостоятельно и получил гарантированный выигрыш  $V(x_0, t_0, \{i\})$ .

Дифференцируя (4), при предположении дифференцируемости функции  $V(x^*(\vartheta), \vartheta, \{i\})$  по  $\vartheta$ , получаем условие на ПРД, обеспечивающее защиту от иррационального поведения:

$$\beta_i(t) \geq \lambda(t)V(\bar{x}(t), t, \{i\}) - \frac{d}{dt}V(\bar{x}(t), t, \{i\}), \quad i = 1, \dots, n, \quad (4.2)$$

где  $\lambda(t)$  определяется по формуле (3.9).

Очевидно, что для случая  $\lambda(t) = 0$  мы фактически имеем полностью детерминированную задачу. Тогда при дополнительном требовании существования всех несобственных интегралов типа (3.11), из условия (4.2) имеем неравенства, полученные в работе [3] для дифференциальных игр с предписанной продолжительностью:

$$\beta_i(t) \geq -\frac{d}{dt}V(\bar{x}(t), t, \{i\}), \quad i = 1, \dots, n, \quad (4.3)$$

Отметим, что свойство динамической устойчивости принципа оптимальности никак не связано с выполнением условия защиты от иррационального поведения. В следующем разделе приведен пример, когда вектор Шепли является динамически устойчивым, но условие защиты от иррационального поведения не выполнено, и наоборот. Однако оба этих условия являются важными аспектами кооперации в динамических играх [3]. В том случае, когда ПРД удовлетворяет и уравнению (3.10), и неравенству (4.2), будем говорить, что вектор Шепли, распределенный во времени согласно ПРД (3.4), является *устойчивым принципом кооперации*.

В данной работе мы не затрагиваем такой аспект устойчивости кооперативного соглашения, как стратегическая поддержка (см. [3]). Однако отметим, что стратегическая поддержка в кооперативной дифференциальной игре со случайной продолжительностью, т.е. существование специально сконструированного равновесия по Нэшу, может быть конструктивно доказана для независимых движений игроков в (2.1).

Продолжим рассматривать полностью детерминированную задачу, т.е. пусть  $\lambda(t) = 0$ . В том случае, когда вектор Шепли является динамически устойчивым принципом оптимальности, т.е. существует  $\{\beta_i(t) \geq 0\}$ , такая что  $\beta_i(t) = -(Sh_i^t)'$ , и при этом выполнено условие защиты от иррационального поведения (4.3), получаем следующее условие устойчивости:

$$(Sh_i^t)' \leq \frac{d}{dt}V(x^*(t), t, \{i\}), \quad i = 1, \dots, n. \quad (4.4)$$

Напомним, что классическое условие индивидуальной рациональности имеет вид:

$$Sh_i^t \geq V(x^*(t), t, \{i\}), \quad i = 1, \dots, n. \quad (4.5)$$

Таким образом, условие (4.4) накладывает ограничения на первые производные для величин, присутствующих в неравенствах (4.5).

Теперь рассмотрим дифференциальную игру со случайной продолжительностью ( $\lambda(t) \neq 0$ ). Тогда одновременное выполнение условия динамической устойчивости вектора Шепли и защиты от иррационального поведения участников означает выполнение следующих неравенств:

$$\lambda(t)[Sh_i^t - V(\bar{x}(t), t, \{i\})] \geq [(Sh_i^t)' - \frac{d}{dt}V(\bar{x}(t), t, \{i\})], \quad (4.6)$$

$$\forall t \in [t_0, \infty), \quad i = 1, \dots, n.$$

Очевидно, что из (4.6) следует выполнение (4.4) при  $\lambda(t) = 0$ . Таким образом, результат, полученный для игр со случайной продолжительностью, покрывает результат, полученный для детерминированных игр.

## 5. Пример

В качестве примера рассмотрим теоретико-игровую модель [7] разработки невозобновляемых ресурсов (в частности, нефти) симметричными игроками. Особо отметим, что спецификой добычи нефти, особенно на континентальном шельфе, является прямая зависимость убытков от аварийности данного предприятия. Аварии на скважинах приводят к простоему производства во время замены и ремонта оборудования, а также к тяжелым экологическим последствиям, затраты

на устранение которых часто приводят к колоссальным убыткам. В большинстве известных теоретико-игровых моделей, описывающих динамический процесс добычи нефти несколькими игроками, также как и в работе [7] предполагается, что игра развивается на бесконечном промежутке времени с постоянным дисконтированием мгновенных выигрышей. В данной работе будем предполагать, что мгновенные выигрыши игроков не дисконтируются, но игра заканчивается в случайный момент времени  $T$ , распределенный по закону Вейбулла. Выбор распределения Вейбулла, как одного из основных распределений, описывающих жизненный цикл работы технических систем, подробно обоснован в работе [6].

Распределение Вейбулла имеет функцию интенсивности отказов следующего вида:

$$\begin{aligned} \lambda(t) &= \lambda \delta t^{\delta-1}; \\ t &\geq 0; \lambda > 0; \delta > 0. \end{aligned} \quad (5.1)$$

Здесь  $\lambda$  и  $\delta$  – параметры, определяющие данное распределение.  $\lambda$  – это параметр масштаба, а параметр формы  $\delta$  соответствует одной из трех фаз, в которой может находиться система. Значение  $\delta < 1$  соответствует «новорожденному» сценарию игры (период приработки). Здесь функция интенсивности отказов  $\lambda(t)$  является убывающей функцией. При  $\delta = 1$  система находится в режиме нормальной эксплуатации,  $\lambda(t)$  равна константе  $\lambda$ . Отметим, что при  $\delta = 1$  распределение Вейбулла соответствует экспоненциальному распределению. При  $\delta > 1$  система находится в состоянии износа,  $\lambda(t)$  является возрастающей функцией.

Итак, согласно модели [7], в игре участвуют  $n$  игроков – фирмы или страны, которые разрабатывают некоторый невозобновляемый природный ресурс, например, нефть. Множество всех игроков обозначим как  $N = \{1, 2, \dots, n\}$ . Пусть  $x(t)$  – это поток невозобновляемого ресурса. Управлениями игроков являются темпы разработки ресурса, которые обозначим как  $\{u_i(t)\}$ . Динамика изменений потока ресурса  $x(t)$  описывается следующим дифференциальным уравнением:

$$\dot{x}(t) = - \sum_{i=1}^n u_i(t); \quad u_i \geq 0, \quad x(t_0) = x_0, \quad x_0 > 0. \quad (5.2)$$

Ожидаемый выигрыш (2.2) игрока  $i$ ,  $i = 1, \dots, n$ , при условии, что момент окончания игры описывается законом Вейбулла, принимает вид:

$$K_i(x_0, t_0, u_1, \dots, u_n) = \int_0^\infty h_i(t, x(t), u(t))e^{-\lambda t^\delta} dt. \quad (5.3)$$

В данном примере каждый игрок  $i$  имеет функцию полезности  $h_i(t, x, u)$  (функцию мгновенного выигрыша) в виде  $h_i = h(u_i)$ , определенную для всех  $u_i > 0$ , которая зависит от маргинальной полезности  $\eta$ :

$$h(u_i) = A \ln(u_i) + B, \quad \eta = 1; \quad (5.4)$$

$$h(u_i) = A \frac{u_i^{1-\eta}}{1-\eta} + B, \quad \eta \neq 1. \quad (5.5)$$

При  $u_i = 0$  по определению полагаем  $h(u_i) = 0$ . Не умаляя общности, далее будем считать, что  $A = 1$ ,  $B = 0$ . Кроме того, положим  $t_0 = 0$ .

Тогда общий ожидаемый выигрыш игроков вычисляется по формуле

$$\sum_{i=1}^n K_i(x_0, t_0, u_1, \dots, u_n) = \int_0^\infty \sum_{i=1}^n h(u_i) e^{-\lambda t^\delta} dt. \quad (5.6)$$

Очевидно, что в данной модели рассматриваются только симметричные игроки, поэтому положим  $u_i = u_j = u$ .

Задача максимизации общего ожидаемого выигрыша (5.6) при условии (5.2) может быть решена при помощи уравнения типа Беллмана, выведенного в работе [6]:

$$\lambda(t)W(x, t) = \frac{\partial W(x, t)}{\partial t} + \max_u \left( \sum_{i=1}^n h_i(x, u, t) + \frac{\partial W(x, t)}{\partial x} g(x, u) \right). \quad (5.7)$$

Кроме того, это уравнение будет использоваться и для вычисления значений характеристической функции  $V(x, t, S)$ . Подробное построение характеристической функции описано в работе [10].

### 5.1. Логарифмическая функция полезности

Рассмотрим функцию полезности вида  $h(u_i) = \ln(u_i)$ . Будем искать функцию Беллмана в виде  $W(x, t) = A(t)\ln x + B(t)$ ,

$\lim_{t \rightarrow \infty} W(x, t) = 0$ . Тогда частные производные  $W(x, t)$  вычисляются по формуле

$$\frac{\partial W(x, t)}{\partial x} = \frac{A(t)}{x}; \quad \frac{\partial W(x, t)}{\partial t} = \dot{A}(t)\ln(x) + \dot{B}(t). \quad (5.8)$$

Учитывая (5.8), из условия максимизации правой части уравнения (5.7) следует, что оптимальные управления имеют вид  $u = \frac{x}{A(t)}$ . Применяя метод неопределенных коэффициентов в уравнении (5.7), получаем следующую систему уравнений для коэффициентов  $A(t), B(t)$ :

$$\dot{A}(t) - \lambda(t)A(t) + n = 0;$$

$$\dot{B}(t) - \lambda(t)B(t) - n\ln(A(t)) - n = 0,$$

и краевыми ограничениями

$$\lim_{t \rightarrow \infty} A(t) = 0, \quad \lim_{t \rightarrow \infty} B(t) = 0.$$

Окончательно, получаем следующие оптимальные управления для задачи разработки невозобновляемых ресурсов со случайной продолжительностью:

$$u_i^* = u^* = \frac{x \cdot e^{-\lambda(t)t}}{n \int_t^\infty e^{-\lambda(s)s} ds}, \quad (5.9)$$

где  $\lambda(t)$  удовлетворяет (5.1). Тогда при  $\delta = 1$ , соответствующем экспоненциальному распределению момента окончания игры, фактически рассматривается уже изученная модель с дисконтированными выигрышами на бесконечном временном промежутке [7]. Непосредственно из (5.9) следует, что при  $\delta = 1$  оптимальными стратегиями игроков являются

$$u_i^* = u^* = \frac{\lambda}{n}x, \quad i = 1, \dots, n.$$

Тогда оптимальные управления и траектория вычисляются по формуле

$$x^*(t) = x_0 \cdot e^{-\lambda t}; \quad u_i^*(t) = \frac{\lambda}{n}x_0 \cdot e^{-\lambda t}.$$



Этот результат совпадает с результатом, полученным в работе Докнера и др.[7] для случая дисконтированных выигрышей на бесконечном временном промежутке для единичной эластичности маргинальной полезности игроков, причем выполнено следующее условие:

$$\lim_{t \rightarrow \infty} x(t) = 0.$$

Отметим, что траектория  $x^*(t)$  удовлетворяет условию устойчивости по Ляпунову.

Значению характеристической функции  $V(x^*(\vartheta), \vartheta, N)$  соответствует значение функции Беллмана  $W(x^*(\vartheta), \vartheta)$ :

$$\begin{aligned} V(x^*(\vartheta), \vartheta, N) &= W(x^*(\vartheta), \vartheta) = \frac{n}{\lambda} \ln(x^*) - \frac{n}{\lambda} - \frac{n \ln(n)}{\lambda} + \frac{n \ln(\lambda)}{\lambda} = \\ &= \frac{n}{\lambda} \ln(x_0) - n(\vartheta) - \frac{n}{\lambda} - \frac{n \ln(n)}{\lambda} + \frac{n \ln(\lambda)}{\lambda}. \end{aligned}$$

Положим  $\vartheta = 0$ . Тогда

$$V(x_0, 0, N) = W(x_0, 0) = \frac{n}{\lambda} \ln(x_0) - \frac{n}{\lambda} - \frac{n \ln(n)}{\lambda} + \frac{n \ln(\lambda)}{\lambda}. \quad (5.10)$$

Далее, для  $\delta = 2$ , которое соответствует распределению Рэля для стареющей системы, из (5.9) получаем

$$u_i^* = \frac{x \cdot e^{-2\lambda t^2}}{n \int_t^\infty e^{-2\lambda s^2} ds}.$$

Тогда оптимальный способ поведения при разработке ресурса должен определяться согласно следующей формуле

$$u_i^* = \frac{2\sqrt{2}\sqrt{\lambda} \cdot e^{-2\lambda t^2}}{n(1 - \operatorname{erf}(\sqrt{2\lambda}t))} x = \frac{2\sqrt{2}\sqrt{\lambda} \cdot e^{-2\lambda t^2}}{n(1 - 2\Phi_0(2\sqrt{\lambda}t))} x,$$

где  $\operatorname{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-s^2} ds$ ,  $\Phi_0(t)$  — интегральная функция Лапласа.

Для периода приработки (раннего периода) возьмем  $\delta = \frac{1}{2}$ . Тогда из уравнения (5.9) получаем

$$u_i^* = \frac{x \cdot e^{-\frac{\lambda}{2}t^{1/2}}}{n \int_t^\infty e^{-\frac{\lambda}{2}s^{1/2}} ds}.$$

Следовательно, получаем оптимальные стратегии в управлениях с обратной связью:

$$u_i^* = \frac{\lambda^2}{4n(\lambda\sqrt{t} + 2)}x.$$

Таким образом, для модели разработки невозобновляемых ресурсов удалось получить оптимальные решения для всех трех сценариев игры. Графическое изображение  $u_i^*(x, t)$  при фиксированном параметре масштаба  $\lambda = 1$  и параметрах формы  $\delta = 1/2; 1; 2$  приведено на рис.1.



Рисунок 1. Оптимальная скорость разработки  $u_i^*$  для трех сценариев игры

Интересно, что в рамках нашей модели мы получили, что оптимальное поведение игроков коренным образом отличается для различных сценариев игры. Для фазы приработки, т.е. когда оборудование и общая координация еще не налажены, скорость разработки должна быть наименьшей, что соответствует осторожности игроков. В режиме нормальной эксплуатации игроки должны «копать» с постоянной скоростью. В режиме износа оборудования (впрочем, это также касается и «износа» не только технических элементов), когда функция интенсивности отказов возрастает, необходимо увеличить темпы разработки месторождений.

Кроме того, было численно проанализировано поведение условно - оптимальных траекторий для всех трех фаз игры. Качественное различие убывания невозобновляемых ресурсов представлено на Рис. 2 и Рис. 3.

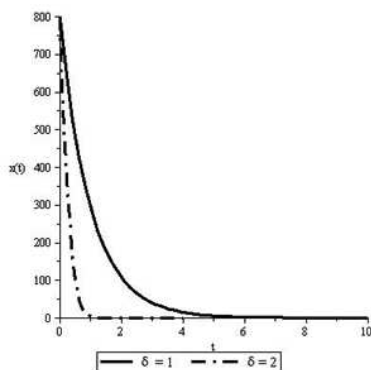


Рисунок 2. Оптимальная траектория  $x^*(t)$  для режима нормальной эксплуатации и фазы износа

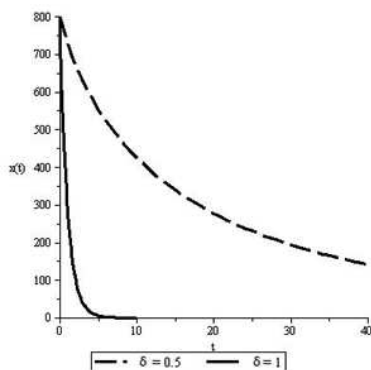


Рисунок 3. Оптимальная траектория  $x^*(t)$  для фазы приработки и режима нормальной эксплуатации

Заметим, что быстрее всего убывает ресурс при  $\delta = 2$ , что соответствует интенсивным разработкам в состоянии износа. Медленнее всего убывает ресурс при  $\delta = 1/2$ , что соответствуем осторожным действиям игроков в состоянии приработки.

Далее, заменив  $\sum_{i=1}^n h_i$  в уравнении (5.7) на  $h_i$ , можно найти равновесные по Нэшу управления  $\{u_i^{nc}\}$  в классе управлений с обратной связью, линейных по фазовой переменной (см. [10]). В данной работе не обсуждается вопрос существования и единственности решения

уравнения (5.7). Проводя аналогичные вычисления, получаем

$$u_i^{nc} = \frac{e^{-\lambda(t)t}}{\int_t^\infty e^{-\lambda(s)s} ds} x, \quad i = 1, \dots, n \quad (5.11)$$

Далее в данном разделе во избежание излишнего нагромождения формул будут представлены результаты только для режима нормальной эксплуатации месторождения, т.е. для  $\delta = 1$ .

Итак, при  $\delta = 1$  выполнено  $\lambda(t) = \lambda$ . Получаем управления, равновесные по Нэшу, а также соответствующие им траекторию и значение характеристической функции.

$$u_i^{nc} = \lambda x, \quad i = 1, \dots, n; \quad (5.12)$$

$$\begin{aligned} x^{nc}(t) &= x^*(\vartheta) e^{-n\lambda(t-\vartheta)}; \\ u_i^{nc}(t) &= \lambda x^*(\vartheta) e^{-n\lambda(t-\vartheta)}; \end{aligned}$$

$$V(x^*(\vartheta), \vartheta, \{i\}) = W_i(x^*(\vartheta)) = \frac{\ln(x^*(\vartheta))}{\lambda} - \frac{n}{\lambda} + \frac{\ln(\lambda)}{\lambda}. \quad (5.13)$$

Положим  $\vartheta = 0$ . Тогда

$$V(x_0, 0, \{i\}) = W_i(x_0, 0) = \frac{\ln(x_0)}{\lambda} - \frac{n}{\lambda} + \frac{\ln(\lambda)}{\lambda}. \quad (5.14)$$

Для построения характеристической функции  $V(x, t, S)$ ,  $S \subseteq N$ , используем подход, предложенный в работе [9]. Будем предполагать, что если  $s$  игроков объединяются в коалицию  $S$ , то оставшиеся игроки  $N \setminus S$  не образуют антикоалицию с целью минимизации совместного выигрыша игроков из  $S$ , а используют равновесные по Нэшу стратегии  $u_j^{nc}$ ,  $j \in N \setminus S$ . Тогда, применяя полученный выше результат (5.12) для равновесия по Нэшу, и используя в уравнении (5.7)  $\sum_{i \in S} h_i$

вместо  $\sum_{i=1}^n h_i$ , получаем следующие результаты [10]. Кооперативная траектория, управления и характеристическая функция имеют следующий вид:

$$\begin{aligned} x^S(t) &= x^*(\vartheta) e^{-(n-s+1)\lambda(t-\vartheta)}; \\ u_i^S(t) &= \frac{\lambda}{s} x^*(\vartheta) e^{-(n-s+1)\lambda(t-\vartheta)}, \quad i \in S; \\ u_j^S(t) &= u_j^{nc}(t), \quad i \in N \setminus S, \end{aligned}$$

$$V(x^*(\vartheta), \vartheta, S) = W_S(x^*(\vartheta), \vartheta) = \frac{s}{\lambda} \ln(x^*(\vartheta)) - \frac{s}{\lambda} - \frac{k(n-s)}{\lambda} - \frac{s}{\lambda} \ln(s) + \frac{s \ln(\lambda)}{\lambda}.$$

Положим  $\vartheta = 0$ . Тогда

$$V(x_0, 0, S) = W_S(x_0, 0) = \frac{s}{\lambda} \ln(x_0) - \frac{s}{\lambda} - \frac{s(n-s)}{\lambda} - \frac{s}{\lambda} \ln(s) + \frac{s \ln(\lambda)}{\lambda}. \quad (5.15)$$

Таким образом, мы построили характеристическую функцию  $V(x_0, 0, S)$ ,  $S \subseteq N$  (см. (5.10), (5.15)), используя подход, описанный в работе [9]. Справедливо следующее утверждение.

**Утверждение 5.1.** Пусть функция  $V(x_0, 0, S)$ ,  $S \subseteq N$  определяется по формулам (5.10), (5.15). Тогда  $V(x_0, 0, S)$  удовлетворяет свойству супераддитивности (3.1).

Для доказательства этого утверждения будем использовать следующую лемму.

**Лемма 5.1.** Пусть  $s_1 \geq 1, s_2 \geq 1$ . Тогда

$$s_1 \ln(s_1) + s_2 \ln(s_2) + 2s_1 s_2 \geq (s_1 + s_2) \ln(s_1 + s_2). \quad (5.16)$$

Данная лемма доказывается стандартными методами математического анализа. Нетрудно проверить, что левая часть неравенства растет быстрее, чем правая.

Доказательство Утверждения 5.1 непосредственно следует из Леммы 5.1 [10].

Используя построенную характеристическую функцию, получаем значение вектора Шепли в подыгре  $\Gamma(x^*(t), t)$  и всей игре  $\Gamma(x_0, t_0)$ :

$$\begin{aligned} Sh_i(x^*(t)) &= \frac{V(x^*(t), t, N)}{n} = \frac{\ln(x^*(t))}{\lambda} - \frac{1}{\lambda} - \frac{\ln(n)}{\lambda} + \frac{\ln(\lambda)}{\lambda} = \\ &= \frac{\ln(x_0)}{\lambda} - (t - t_0) - \frac{1}{\lambda} - \frac{\ln(n)}{\lambda} + \frac{\ln(\lambda)}{\lambda}, \\ Sh_i(x_0) &= \frac{V(x_0, 0, N)}{n} = \frac{\ln(x_0)}{\lambda} - \frac{1}{\lambda} - \frac{\ln(n)}{\lambda} + \frac{\ln(\lambda)}{\lambda}. \end{aligned} \quad (5.17)$$

Вычислим значения ПРД по формуле (3.10), которая для случая  $\delta = 1$  в точности совпадает с формулой (3.13). Получаем

$$\beta_i(\vartheta) = \ln(x_0) - \lambda\vartheta + \ln(\lambda) - \ln(n). \quad (5.18)$$

Справедливо следующее утверждение.

**Утверждение 5.2.** Вектор Шепли  $\{Sh_i(x_0)\}$  (5.17) не является динамически устойчивым принципом оптимальности.

Доказательство Утверждения 5.2 основано на Определении 3.2. Очевидно, что в (5.18) нельзя гарантировать неотрицательность компонент ПРД.

Отметим, что, поскольку  $h_i = \ln u_i$  не является неотрицательной функцией, мы не можем воспользоваться регуляризацией вектора Шепли (3.8).

Проверим выполнение условия защиты от иррационального поведения участников (4), используя неравенство для ПРД (4.2). В данном примере с логарифмической функцией мгновенного выигрыша для параметра  $\delta = 1$ , получено значение характеристической функции  $V(x^*(\vartheta), \vartheta, \{i\})$  (5.13). Справедливо следующее утверждение.

**Утверждение 5.3.** Для ПРД (5.18) и характеристической функции  $V(x^*(\vartheta), \vartheta, \{i\})$  (5.13) выполнено условие защиты от иррационального поведения (4.2).

Утверждение 5.3 доказывается непосредственной проверкой неравенства (4.2) при  $\lambda(t) = \lambda$ .

Таким образом, в данном примере с логарифмической функцией полезности было показано, что при экспоненциальном распределении момента окончания игры ( $\delta = 1$ ,  $\lambda(t) = \lambda$ ) вектор Шепли не является динамически устойчивым принципом оптимальности, а условие защиты от иррационального поведения участников выполнено.

## 5.2. Изоэластичная функция полезности

Рассмотрим тот же самый пример игры, только с изоэластичной функцией полезности  $h_i$ , т.е.

$$h(u_i) = \frac{u_i^{1-\eta}}{1-\eta}, \quad \eta \neq 1.$$

Все вычисления проводятся аналогично тому, как это было сделано в разделе 5.1 для логарифмической функции полезности. Главным отличием является вид функции Беллмана, которая ищется как  $W(x, t) = A(t)x^{1-\eta} + B(t)$  [10]. Получаем оптимальную скорость разработки ресурсов для произвольной функции интенсивности отказов

$\lambda(t)$  (см. [8]):

$$u_i^* = \frac{e^{-\frac{\lambda(t)t}{\eta}}}{n \int_t^{\infty} e^{-\frac{\lambda(s)s}{\eta}} ds} x.$$

Тогда для экспоненциального распределения для случайной величины  $T$ , которому соответствует  $\lambda(t) = \lambda$ , имеем:

$$u_i^* = \frac{\lambda}{\eta n} x, \quad i = 1, \dots, n.$$

Следовательно, оптимальная траектория и оптимальные управления вычисляются согласно следующим формулам:

$$\begin{aligned} x^*(t) &= x_0 e^{-\frac{\lambda t}{\eta}}; \\ u_i^*(t) &= \frac{x_0 \lambda}{n \eta} e^{-\frac{\lambda t}{\eta}}. \end{aligned}$$

Отметим, что траектория  $x^*(t)$  удовлетворяет условию устойчивости по Ляпунову.

Далее получаем выражение для характеристической функции для максимальной коалиции  $N$ :

$$V(x^*(\vartheta), \vartheta, N) = \left(\frac{n\eta}{\lambda}\right)^\eta \frac{1}{1-\eta} x^*(\vartheta)^{1-\eta} = \left(\frac{n\eta}{\lambda}\right)^\eta \frac{1}{1-\eta} x_0^{1-\eta} e^{-\frac{\lambda(1-\eta)\vartheta}{\eta}}. \quad (5.19)$$

Для равновесия по Нэшу имеем следующие результаты:

$$u_i^{nc} = \frac{\lambda x}{(1 - n + n\eta)}, \quad i = 1, \dots, n; \quad (5.20)$$

$$\begin{aligned} x^{nc}(t) &= x^*(\vartheta) e^{-\frac{n\lambda}{(1-n+n\eta)}(t-\vartheta)}; \\ u_i^{nc}(t) &= \frac{\lambda x^*(\vartheta)}{(1-n+n\eta)} e^{-\frac{n\lambda}{(1-n+n\eta)}(t-\vartheta)}. \end{aligned}$$

Очевидно, что требование неотрицательности скорости разработки  $u_i^{nc} \geq 0$  выполнено только при  $\eta > (1 - 1/n)$ . В противном случае равновесия по Нэшу не существует.

Итак, вычисляем значение характеристической функции  $V(x^*(\vartheta), \vartheta, \{i\})$ :

$$\begin{aligned} V(x^*(\vartheta), \vartheta, \{i\}) &= \left(\frac{(1-n+n\eta)}{\lambda}\right)^\eta \frac{1}{1-\eta} x^*(\vartheta)^{1-\eta} = \\ &= \left(\frac{(1-n+n\eta)}{\lambda}\right)^\eta \frac{1}{1-\eta} x_0^{1-\eta} e^{-\frac{\lambda(1-\eta)}{\eta}\vartheta}. \end{aligned} \quad (5.21)$$

Описанным в предыдущем разделе способом, получаем выражение для характеристической функции  $V(x^*(\vartheta), \vartheta, S)$ ,  $S \subset N$ .

$$\begin{aligned} u_i^S &= \frac{\lambda(1-s+s\eta)x}{s\eta(1-n+n\eta)}, \quad i \in S; \\ u_j^S(t) &= u_j^{nc}(t), \quad i \in N \setminus S, \\ V(x^*(\vartheta), \vartheta, S) &= \left( \frac{s\eta(1-n+n\eta)}{\lambda(1-s+s\eta)} \right)^\eta \frac{1}{1-\eta} x(\vartheta)^{1-\eta}; \end{aligned} \quad (5.22)$$

Справедливо следующее утверждение.

**Утверждение 5.4.** Пусть функция  $V(x_0, 0, S)$ ,  $S \subseteq N$  определяется по формулам (5.19), (5.22). Тогда  $V(x_0, 0, S)$  удовлетворяет свойству супераддитивности (3.1).

Доказательство Утверждения 5.4 основано на следующей лемме.

**Лемма 5.2.** Пусть  $s_1 \geq 1, s_2 \geq 1, \eta \in (0, 1)$ . Тогда

$$\frac{(s_1 + s_2)^\eta}{(1 - (s_1 + s_2) + (s_1 + s_2)\eta)^\eta} \geq \frac{s_1^\eta}{(1 - s_1 + s_1\eta)^\eta} + \frac{s_2^\eta}{(1 - s_2 + s_2\eta)^\eta}. \quad (5.23)$$

Нетрудно проверить, что левая часть неравенства растет быстрее, чем правая.

Окончательно получаем следующее выражение для вектора Шепли:

$$Sh_i(x(t)) = \frac{1}{1-\eta} \left( \frac{x(t)}{n} \right)^{1-\eta} \left( \frac{\lambda}{\eta} \right)^{-\eta}; \quad (5.24)$$

$$Sh_i(x_0) = \frac{1}{1-\eta} \left( \frac{x_0}{n} \right)^{1-\eta} \left( \frac{\lambda}{\eta} \right)^{-\eta}. \quad (5.25)$$

Вычислим ПРД согласно (3.13). Тогда

$$\beta_i(\vartheta) = \lambda \left( \frac{n\eta}{\lambda} \right)^\eta x_0^{1-\eta} e^{-\frac{\lambda\vartheta}{\eta}} \frac{1}{\eta(1-\eta)}. \quad (5.26)$$

Следующие утверждения доказываются аналогично тому, как это было сделано выше в разделе для логарифмической функции полезности. Однако результат проверки динамической устойчивости и условия защиты от иррационального поведения оказывается противоположным предыдущему.



**Утверждение 5.5.** Вектор Шепли  $\{Sh_i(x_0)\}$  (5.24) является динамически устойчивым принципом оптимальности при маргинальной полезности  $\eta \in (0; 1)$ .

Доказательство Утверждения 5.5 основано на Определении 3.2. Очевидно, что в (5.26) можно гарантировать неотрицательность компонент ПРД при  $\eta \in (0; 1)$ .

**Утверждение 5.6.** Для ПРД (5.26) и характеристической функции  $V(x^*(\vartheta), \vartheta, \{i\})$  (5.21) не выполнено условие защиты от иррационального поведения (4.2).

Утверждение 5.6 доказывается непосредственной проверкой неравенства (4.2) при  $\lambda(t) = \lambda$ .

Таким образом, в данном примере с изоэластичной функцией полезности было показано, что при экспоненциальном распределении момента окончания игры ( $\delta = 1$ ,  $\lambda(t) = \lambda$ ) вектор Шепли является динамически устойчивым принципом оптимальности при маргинальной полезности  $\eta \in (0; 1)$ , а условие защиты от иррационального поведения участников не выполнено.

Следовательно, согласно результатам разделов 5.1, 5.2, построенный вектор Шепли не является устойчивым кооперативным соглашением ни для случая логарифмической функции полезности, ни для случая изоэластичной функции полезности игроков. При этом в обоих случаях условно-оптимальная траектория игры удовлетворяет условию устойчивости по Ляпунову.

## СПИСОК ЛИТЕРАТУРЫ

1. Петросян Л.А. *Сильно динамически устойчивые принципы оптимальности* // Вестн. С.-Петербург. ун-та. Сер. 1. 1993. Вып. 4. № 22. С. 35–40.
2. Петросян Л.А., Данилов Н.А. *Устойчивые решения неантагонистических дифференциальных игр с транзитивными выигрышами* // Вестник ЛГУ. 1979. № 1. С. 46–54.
3. Петросян Л.А., Зенкевич Н.А. *Принципы устойчивой кооперации* // Математическая теория игр и ее приложения. 2009. Т.

1. Вып. 1. С. 102–117.
4. Петросян Л.А., Мурзов Н.В. *Теоретико-игровые проблемы в механике* // Литовский математический сборник. 1966. VI-3. С. 423–433.
5. Петросян Л.А., Шевкопляс Е.В. *Кооперативные дифференциальные игры со случайной продолжительностью* // Вестник СПбГУ. 2000. Сер. 1. Вып. 4. С. 18–23.
6. Шевкопляс Е.В. *Уравнение Гамильтона-Якоби-Беллмана в дифференциальных играх со случайной продолжительностью* // Математическая теория игр и ее приложения. 2009. Т.1. Вып. 2. С. 98–118.
7. Dockner E.J., Jorgensen S., N. van Long and Sorger G. *Differential games in economics and management science*. Cambridge University Press, 2000.
8. Marin-Solano J., Shevkoplyas E.V. *Non-constant discounting in differential games with random duration* // Contributions to Game Theory and Management, collected papers of the Third International Conference «Game Theory and Management 2009». St. Petersburg. 2010. P. 267–280.
9. Petrosyan L. A., Zaccour G. *Time-consistent Shapley value allocation of pollution cost reduction*. // Journal of Economic Dynamics and Control. 2003. V. 27. P. 381–398.
10. Shevkoplyas E.V. *The Shapley value in cooperative differential games with random duration* // Advances in Dynamic Games. 2011. V. 11. part 4. Edt. by M.Breton and K. Szajowski, Springer's imprint Birkhauser. Boston. P. 359–373. (<http://www.springerlink.com/content/978-0-8176-8088-6/#section=815141&page=1>)
11. Yeung D.W.K. *An irrational-behavior-proofness condition in cooperative differential games* // Int. J. of Game Theory Rew. 2007. V. 9. N 1. P. 256–273.

STABLE COOPERATION IN DIFFERENTIAL GAMES  
WITH RANDOM DURATION

**Ekaterina V. Shevkoplyas**, Faculty of Applied Mathematics and Control Processes, St.Petersburg State University, Saint-Petersburg, Cand.Sc. (katya\_shev@mail.ru).

*Abstract:* The problem of time-consistency of cooperative solutions is investigated in the paper. This problem was stated by Petrosyan L.A. in 1977 for differential games with finite time horizon. In the paper the modification of the game with finite time horizon is considered in the sense that the game has random time horizon. The Shapley value is used as an optimality principle under cooperative behavior of the players. For this formulation the definition of the imputation distribution procedure (IDP) is given and the analytic formula for IDP is derived. Moreover in the paper the irrational behavior proofness condition by D.W.K. Yeung (2006) is modified for problem with random duration. The tool is based on using IDP. Theoretical results are illustrated by an example of differential game of non-renewable resource extraction.

*Keywords:* time-consistency, stable cooperation, irrational behavior proofness, non-renewable resource extraction, differential game with random duration.