

УДК 519.83

СИЛЬНО ДИНАМИЧЕСКИ УСТОЙЧИВОЕ КООПЕРАТИВНОЕ РЕШЕНИЕ В ОДНОЙ ДИФФЕРЕНЦИАЛЬНОЙ ИГРЕ УПРАВЛЕНИЯ ВРЕДНЫМИ ВЫБРОСАМИ¹

Громова Е. В.², Петросян Л. А.³

(Санкт-Петербургский государственный университет)

Рассматривается проблема построения сильно динамически устойчивого (СДУ) кооперативного решения для дифференциальных игр двух лиц. Формулируется подход, позволяющий при достаточно общих предположениях построить кооперативное решение, удовлетворяющее СДУ. Показывается, что полученное решение, кроме того, будет удовлетворять условию защиты от иррационального поведения. Теоретические результаты демонстрируются на примере дифференциальной игры управления вредными выбросами в атмосферу.

Ключевые слова: дифференциальные игры, динамическая устойчивость, сильно динамическая устойчивость, условие Янга, модель управления вредными выбросами.

1. Постановка задачи

Рассматривается кооперативная дифференциальная игра n лиц $\Gamma(x_0, T - t_0)$ с предписанной продолжительностью $T - t_0$, на отрезке времени $t \in [t_0, T]$ из начального состояния $x_0 \in R^n$,

¹ Работа выполнена при финансовой поддержке гранта СПбГУ 9.38.245.2014.

² Екатерина Викторовна Громова, кандидат физико-математических наук, доцент (ekaterina.shevkokplyas@gmail.com).

³ Леон Аганесович Петросян, доктор физико-математических наук, профессор (Санкт-Петербург, Петродворец, Университетский пр., д.35, тел. (812) 428-71-59).

уравнениями движения

$$(1) \quad \dot{x} = f(x, u_1, \dots, u_n), \quad x(t_0) = x_0, \\ x \in R^n, u_i \in U \subset \text{comp} R^k, \text{ и функциями выигрыша}$$

$$K_i(x_0, T - t_0; u_1, \dots, u_n) = \int_{t_0}^T h_i(x, u_1, \dots, u_n) dt,$$

где $x(t)$ — решение системы (1) при управлениях (u_1, \dots, u_n) и $i \in N$, где N — множество игроков, $|N| = n$.

Пусть $S \subset N$ — коалиция в игре $\Gamma(x_0, T - t_0)$. Определим для $S \subset N$ характеристическую функцию $v(x_0, T - t_0; S)$ как нижнее значение антагонистической игры между коалицией S , действующей как игрок I (максимизирующий) и коалицией $N \setminus S$, действующей как игрок II (минимизирующий), где под выигрышем игрока S понимается сумма выигрышей игроков, входящих в S , а под стратегией игрока S — элемент декартова произведения множеств стратегий игроков, входящих в S . В данной формализации мы из соображений простоты под стратегией игрока будем понимать функцию $u_i(x, t)$ со значением в множестве мгновенных допустимых управлений U_i . Нижнее значение игры (supinf) всегда существует и является супераддитивной функцией от коалиции $S \subset N$.

Определение 1. Траекторию $\bar{x}(t)$, $t \in [t_0, T]$, назовем кооперативной траекторией, если имеет место

$$\begin{aligned} & \max_{u_1, \dots, u_n} \sum_{i=1}^n K_i(x_0, T - t_0; u_1, \dots, u_n) = \\ & = \max_{u_1, \dots, u_n} \sum_{i=1}^n \int_{t_0}^T h_i(x, u_1, \dots, u_n) dt = \\ & = \sum_{i=1}^n \int_{t_0}^T h_i(\bar{x}(t), \bar{u}_1(t), \dots, \bar{u}_n(t)) dt. \end{aligned}$$

Мы предполагаем, что такая траектория $\bar{x}(t)$, $t \in [t_0, T]$, существует и единственна. В противном случае дальнейшие выкладки и определения необходимо незначительно изменить.

Рассмотрим подыгры игры $\Gamma(x_0, T - t_0)$, $\Gamma(\bar{x}(t), T - t)$ с начальным условием на кооперативной траектории. В каждой подыгре $\Gamma(\bar{x}(t), T - t)$ можно таким же образом, как и в основной игре $\Gamma(x_0, T - t_0)$, определить характеристическую функцию (см. [1]) $v(\bar{x}(t), T - t; S)$, где $S \subset N$, которая также будет супераддитивной. Определим множество дележей $M(\bar{x}(t), T - t)$ в игре $\Gamma(\bar{x}(t), T - t)$ как

$$M(\bar{x}(t), T - t) = \left\{ \alpha = (\alpha_1, \dots, \alpha_n) : \sum_{i=1}^n \alpha_i = v(\bar{x}(t), T - t; N), \right. \\ \left. \alpha_i \geq v(\bar{x}(t), T - t; \{i\}), i \in N \right\}.$$

Из супераддитивности характеристической функции следует, что множество $M(\bar{x}(t), T - t) \neq \emptyset$, $t \in [t_0, T]$. Определим также ядро $C(\bar{x}(t), T - t) \subset M(\bar{x}(t), T - t)$ в игре $\Gamma(\bar{x}(t), T - t)$ и предположим, что для всех $t \in [t_0, T]$, $C(\bar{x}(t), T - t) \neq \emptyset$.

Напомним, что ядро в игре $\Gamma(\bar{x}(t), T - t)$ — это множество дележей $\alpha^t = (\alpha_1^t, \dots, \alpha_n^t)$, удовлетворяющих неравенствам

$$\sum_{i \in S} \alpha_i^t \geq v(\bar{x}(t), T - t; S)$$

для всех $S \subset N$.

Определение 2. [6]. Функция $\beta_i(\tau)$, $\tau \in [t_0, T]$, $i \in N$, называется процедурой распределения дележа $\alpha \in M(x_0, T - t_0)$, если

$$\alpha_i = \int_{t_0}^T \beta_i(\tau) d\tau, \quad i \in N.$$

Определение 3. [6]. Ядро $C(x_0, T - t_0)$ в игре $\Gamma(x_0, T - t_0)$ называется динамически устойчивым, если для каждого дележа $\alpha \in C(x_0, T - t_0)$ найдется процедура распределения дележа (ПРД) такая, что

$$\int_t^T \beta_i(\tau) d\tau \in C(\bar{x}(t), T - t), \quad t \in [t_0, T], i \in N.$$

Нами было ранее показано (см. [4, 6]), что если $C(\bar{x}(t), T - t) \neq \emptyset$ при $t \in [t_0, T]$ и существует дифференцируемый селектор $\alpha(t) \in C(\bar{x}(t), T - t)$ ($\alpha(t_0) = \alpha$), то ядро является динамически устойчивым и ПРД $\{\beta_i(t)\}$ определяется по формуле

$$\beta_i(\tau) = -\frac{d}{d\tau}\alpha_i(\tau), \quad i = 1, \dots, n.$$

$$\alpha(t_0) = \alpha.$$

Приведем определение сильного динамически устойчивого ядра $C(x_0, T - t_0)$ в игре $\Gamma(x_0, T - t_0)$.

Определение 4. Ядро $C(x_0, T - t_0)$ сильно динамически устойчиво в игре $\Gamma(x_0, T - t_0)$, если

- 1) $C(\bar{x}(t), T - t) \neq \emptyset, t \in [t_0, T]$;
- 2) существует такой дележ $\alpha \in C(x_0, T - t_0)$ и такая ПРД $\beta(\tau) = (\beta_1(\tau), \dots, \beta_n(\tau)), \tau \in [t_0, T]$, что

$$\alpha_i = \int_{t_0}^T \beta_i(\tau) d\tau, \quad \forall i = 1, \dots, n,$$

и

$$C(x_0, T - t_0) \supset \int_{t_0}^t \beta(\tau) d\tau \oplus C(\bar{x}(t), T - t), \quad t \in [t_0, T].$$

Здесь символ \oplus определяется следующим образом. Пусть $a \in R^n, B \subset R^n$, тогда $a \oplus B = \{a + b : b \in B\}$.

Заметим, что определение 4 незначительно отличается от определения сильной динамической устойчивости в работах [3, 14]. Сильная динамическая устойчивость ядра означает, что при развитии игры вдоль кооперативной траектории $\bar{x}(t)$ в ядре существует дележ $\bar{\alpha}$ такой, что однократное отклонение в момент t от этого дележа в пользу другого оптимального дележа (т.е. другого дележа из ядра в подыгре, начинающейся в момент времени t из состояния $\bar{x}(t)$), приведет к суммарным выплатам игрокам во всей игре, которые также содержатся в изначально выбранном принципе оптимальности (ядре).

2. Сильно динамически устойчивое ядро в кооперативной дифференциальной игре с предписанной продолжительностью и с двумя участниками

2.1. СУЩЕСТВОВАНИЕ СИЛЬНО ДИНАМИЧЕСКИ УСТОЙЧИВОГО ЯДРА

Как мы уже отметили и как это следует непосредственно из определения множества дележей и ядра для игры двух лиц, эти множества совпадают, т.е. $M(x_0, T - t_0) = C(x_0, T - t_0)$, и также множества $M(\bar{x}(t), T - t) = C(\bar{x}(t), T - t)$, $t \in [t_0, T]$.

Для игры $\Gamma(x_0, T - t_0)$ с двумя участниками ядро имеет вид

$$\begin{aligned} C(x_0, T - t_0) = \{ \alpha = (\alpha_1, \alpha_2) : \alpha_1 + \alpha_2 = v(x_0, T - t_0; N), \\ \alpha_1 \geq v(x_0, T - t_0; \{1\}), \\ \alpha_2 \geq v(x_0, T - t_0; \{2\}) \} = M(x_0, T - t_0). \end{aligned}$$

Аналогичным образом для подыгр $\Gamma(x(t), T - t)$ мы имеем

$$\begin{aligned} C(\bar{x}(t), T - t) = \{ \alpha^t = (\alpha_1^t, \alpha_2^t) : \alpha_1^t + \alpha_2^t = v(\bar{x}(t), T - t; N), \\ \alpha_1^t \geq v(\bar{x}(t), T - t; \{1\}), \\ \alpha_2^t \geq v(\bar{x}(t), T - t; \{2\}) \} = M(\bar{x}(t), T - t). \end{aligned}$$

Введем величины $A_1(t) \geq 0$ и $A_2(t) \geq 0$ следующим образом:

$$(2) \quad A_1(t) + A_2(t) = v(\bar{x}(t), T - t; N) - \sum_{i=1}^2 v(\bar{x}(t), T - t; \{i\}) \geq 0.$$

Вычислим $\gamma_i(t) = -\frac{d}{dt} A_i(t)$, $i = 1, 2$. Будем также использовать обозначение A_i для $A_i(T)$, $i = 1, 2$. Очевидно, что

$$A_i = \int_{t_0}^T \gamma_i(t) dt.$$

Введем функции (в предположении дифференцируемости функций $v(\bar{x}(t), T - t; \{i\})$)

$$(3) \quad \bar{\beta}_i(\tau) = \gamma_i(\tau) - \frac{d}{d\tau} v(\bar{x}(\tau), T - \tau; \{i\})$$

и потребуем, чтобы при всех $\tau \in [t_0, T]$ имело место

$$(4) \quad \bar{\beta}_1(\tau) + \bar{\beta}_2(\tau) = -\frac{d}{d\tau}v(\bar{x}(\tau), T - \tau; N).$$

Тогда получаем

$$(5) \quad \sum_{i=1}^2 \gamma_i(\tau) = -\frac{d}{d\tau}v(\bar{x}(\tau), T - \tau; N) + \sum_{i=1}^2 \frac{d}{d\tau}v(\bar{x}(\tau), T - \tau; \{i\})$$

и, интегрируя, получаем (2)

$$A_1 + A_2 = -\int_{t_0}^T \frac{d}{d\tau}v(\bar{x}(\tau), T - \tau; N)d\tau + \sum_{i=1}^2 \int_{t_0}^T \frac{d}{d\tau}v(\bar{x}(\tau), T - \tau; \{i\})d\tau.$$

$$A_1 + A_2 = v(\bar{x}(t_0), T - t_0; N) - \sum_{i=1}^2 v(\bar{x}(t_0), T - t_0; \{i\}) \geq 0.$$

То есть если $\gamma_i(\tau)$, $i = 1, 2$ удовлетворяет (5), то условие (2) всегда выполнено.

Рассмотрим теперь вектор $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_2)$, где

$$(6) \quad \bar{\alpha}_i = \int_{t_0}^T \bar{\beta}_i(\tau)d\tau.$$

Легко видеть, что $\bar{\alpha} \in C(x_0, T - t_0)$. Действительно,

$$\begin{aligned} \bar{\alpha}_i &= \int_{t_0}^T \bar{\beta}_i(\tau)d\tau = \int_{t_0}^T \gamma_i(\tau)d\tau + \int_{t_0}^T \left(-\frac{d}{d\tau}v(\bar{x}(\tau), T - \tau; \{i\})\right)d\tau = \\ &= A_i + v(x_0, T - t_0; \{i\}), \end{aligned}$$

и согласно (2)

$$\bar{\alpha}_1 + \bar{\alpha}_2 = \sum_{i=1}^2 A_i + v(x_0, T - t_0; \{i\}) = v(x_0, T - t_0; N).$$

А поскольку $A_i > 0$, имеем

$$\bar{\alpha}_i = A_i + v(x_0, T - t_0; \{i\}) \geq v(x_0, T - t_0; \{i\}),$$

т.е. $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_2) \in C(x_0, T - t_0)$.

Условие 2.1. Существуют такие A_1 и A_2 , удовлетворяющие (2), и такие $\gamma_1(\tau)$, $\gamma_2(\tau)$, $\tau \in [t_0, T]$, удовлетворяющие (4), что $\int_{t_0}^t \gamma_i(\tau) d\tau \geq 0$ при $t \in [t_0, T]$.

Теорема 1. При выполнении условия 2.1 ядро игры $\Gamma(x_0, T - t_0)$ сильно динамически устойчиво.

Доказательство. Во-первых, из-за супераддитивности характеристической функции $C(\bar{x}(\tau), T - \tau) \neq 0$, $\tau \in [t_0, T]$. Возьмем теперь дележ $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_2)$, который определяется формулой (6), $\bar{\alpha} \in C(x_0, T - t_0)$.

Покажем, что имеет место сильная динамическая устойчивость, при этом в качестве дележа $\alpha = (\alpha_1, \alpha_2)$, фигурирующего в определении сильной динамической устойчивости, можно взять дележ $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_2)$. Надо показать включение

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

при всех $t \in [t_0, T]$, где $\bar{\beta}(\tau)$ является ПРД для дележа $\bar{\alpha}$.

Действительно, возьмем произвольный дележ $\alpha^t \in C(\bar{x}(t), T - t)$. Обозначим

$$\hat{\alpha}_i = \int_{t_0}^t \bar{\beta}_i(\tau) d\tau + \alpha_i^t,$$

тогда

$$\begin{aligned} \hat{\alpha}_1 + \hat{\alpha}_2 &= \int_{t_0}^t (\bar{\beta}_1(\tau) + \bar{\beta}_2(\tau)) d\tau + \alpha_1^t + \alpha_2^t = \\ &= \int_{t_0}^t \left[-\frac{d}{d\tau} v(\bar{x}(\tau), T - \tau; N) \right] d\tau + v(\bar{x}(t), T - t; N) = \\ &= -v(\bar{x}(t), T - t; N) + v(\bar{x}(t_0), T - t_0; N) + v(\bar{x}(t), T - t; N) = \\ &= v(\bar{x}(t_0), T - t_0; N). \end{aligned}$$

Кроме того,

$$\begin{aligned} \hat{\alpha}_i &= \int_{t_0}^t \bar{\beta}_i(\tau) d\tau + \alpha_i^t \geq \\ &\geq \int_{t_0}^t [\gamma_i(\tau) - \frac{d}{d\tau} v(\bar{x}(\tau), T - \tau; \{i\})] d\tau + v(\bar{x}(t), T - t; \{i\}) = \\ &= \int_{t_0}^t \gamma_i(\tau) d\tau - v(\bar{x}(t), T - t; \{i\}) + v(\bar{x}(t_0), T - t_0; \{i\}) + \\ &\quad + v(\bar{x}(t), T - t; \{i\}) \geq v(\bar{x}(t_0), T - t_0; \{i\}). \end{aligned}$$

Последнее неравенство выполнено из-за условия 2.1 на A_1 , A_2 и $\gamma_1(\tau)$, $\gamma_2(\tau)$ ($\tau \in [t_0, T]$), т.е. ядро $C(x_0, T - t_0)$ сильно динамически устойчиво, так как дележ $\alpha^t \in C(\bar{x}(t), T - t)$ произвольный.

Заметим, что в каждый момент времени $t \in [t_0, T]$ имеет место

$$\sum_{i=1}^2 \beta_i(t) = -\frac{d}{dt} v(\bar{x}(t), T - t; N) = \sum_{i=1}^2 h_i(\bar{x}(t)),$$

таким образом, используя ПРД $\bar{\beta} = (\bar{\beta}_1, \bar{\beta}_2)$ игроки фактически в каждый момент времени перераспределяют мгновенный суммарный доход, т.е. обеспечивается мгновенная трансферабельность выигрышей, что вполне соответствует идеологии теории кооперативных игр с трансферабельными выигрышами.

2.2. УСЛОВИЕ ЯНГА

Проблема динамической и сильно динамической устойчивости кооперативных решений изучается при предположении о рациональном поведении игроков. В то же время является актуальным вопрос динамической устойчивости кооперативного решения при предположении о возможности нерационального поведения для отдельных игроков. Как известно [15], в данном случае условие для защиты игроков от иррационального поведения

других участников игры выглядит следующим образом:

$$(7) \quad \int_{t_0}^{\vartheta} \beta_i(t) dt + V(\{i\}, \bar{x}(\vartheta), T - \vartheta) \geq V(\{i\}, x_0, T - t_0), \quad i = 1, \dots, n.$$

Это означает, что в случае иррационального разрушения кооперации в момент ϑ выигрыш, который получит игрок i все равно будет не меньше, чем выигрыш в некооперативном варианте игры, т.е. если он с самого начала будет действовать самостоятельно.

Для случая игры двух лиц условие (7) может быть переписано в следующем виде (при помощи первой производной по ϑ):

$$(8) \quad \beta_i(t) \geq -\frac{d}{dt} V(\{i\}, \bar{x}(t), T - t), \quad i = 1, 2.$$

Рассмотрим некоторый селектор $\bar{\alpha}_i(t) = \int_t^T \bar{\beta}_i(\tau) d\tau$ из сильно динамически устойчивого решения (6). Тогда условие (8) может быть представлено в следующем виде:

$$(9) \quad \frac{d}{dt} \bar{\alpha}_i(t) \leq \frac{d}{dt} V(\{i\}, \bar{x}(t), T - t), \quad i = 1, 2.$$

Фактически условие (9) является «уточнением» условия индивидуальной рациональности ($\bar{\alpha}_i(t) \geq V(\{i\}, \bar{x}(t), T - t)$) для производной первого порядка.

Поскольку для игры двух лиц по построению $\alpha_i(t) = A_i(t) + V(\{i\}, \bar{x}(t), T - t)$ где $\frac{d}{dt} A_i(t) = -\gamma_i(t) \leq 0$, $i = 1, 2$, имеем:

$$(10) \quad \begin{aligned} \frac{d}{dt} \bar{\alpha}_i(t) &= -\gamma_i(t) + \frac{d}{dt} V(\{i\}, \bar{x}(t), T - t) \leq \\ &\leq \frac{d}{dt} V(\{i\}, \bar{x}(t), T - t), \quad i = 1, 2. \end{aligned}$$

Таким образом, неравенство (10) выполняется для любого дележа, построенного по описанной выше процедуре при помощи ПРД (3).

Это означает, что предложенная процедура распределения дележа (3) в случае игры двух лиц будет обеспечивать не только выполнение свойства сильно динамической устойчивости ядра, но и гарантировать выполнение условие Янга защиты от иррационального поведения участников.

3. Дифференциальная игра управления вредными выбросами в атмосферу

3.1. МОДЕЛЬ ИГРЫ

В качестве примера рассмотрим теоретико-игровую модель управления вредными выбросами в атмосферу [10, 13]. В игре принимают участие 2 игрока (фирмы, страны), каждый из которых имеет промышленное производство на своей территории. Предполагается, что объём производства прямо пропорционален вредным выбросам u_i . Таким образом, стратегией игрока является выбор объёма вредных выбросов $u_i \in [0; u_i^{max}]$. В данном примере будем искать решение в классе позиционных стратегий $u_i(t, x)$.

Динамика изменения общего уровня загрязнения $x(t)$ задаётся уравнением

$$\dot{x}(t) = \sum_{i=1}^2 u_i(t) - \delta x(t), \quad x(t_0) = x_0,$$

где δ – коэффициент абсорбции, соответствующий естественному очищению атмосферы.

Доход игрока i в момент времени t определяется по формуле

$$R(u_i(t)) = c_i u_i(t) - \frac{1}{2} u_i^2(t).$$

Каждый игрок несет расходы, связанные с устранением загрязнений. Мгновенный выигрыш (полезность) игрока i равен $R(u_i(t)) - k_i x(t)$, $k_i > 0$.

Без ограничения общности будем предполагать, что момент начала игры $t_0 = 0$.

Тогда выигрыш i -го игрока имеет вид

$$(11) \quad K_i(0, x_0, u_1, u_2) = \int_0^T (R_i(u_i(\tau)) - k_i x(\tau)) d\tau.$$

Предположим, что выполняется следующее условие регулярности:

$$(12) \quad \frac{\sum_{j=1}^2 k_j}{\delta} \leq c_i \leq u_i^{max}, \quad i = 1, 2.$$

3.2. КООПЕРАТИВНАЯ ИГРА

Рассмотрим кооперативный вариант игры, в котором игроки максимизируют суммарный выигрыш

$$(13) \quad \sum_{i=1}^2 K_i(0, x_0, u_1, u_2).$$

Для определения управлений $\bar{u} = (\bar{u}_1, \bar{u}_2)$, доставляющих максимум выражению (13), запишем уравнение Гамильтона–Якоби–Беллмана:

$$(14) \quad -V_t^{12} = \max_{u_i} \left\{ -k_{12}x + \sum_{i=1}^2 c_i u_i - \frac{1}{2} \sum_{i=1}^2 u_i^2 + V_x^{12} \sum_{i=1}^2 u_i - \delta x V_x^{12} \right\},$$

где V_t^{12} и V_x^{12} – частные производные функции Беллмана $V^{12}(x, t)$, $k_{12} = \sum_{i=1}^2 k_i$.

Максимизирующее управление найдем из выражения $c_i - u_i + V_x = 0$, откуда следует $\bar{u}_i = c_i + V_x$. После подстановки \bar{u}_i в (14) получаем

$$(15) \quad -V_t^{12} = -k_{12}x + \frac{1}{2} \hat{c}_{12} + c_{12} V_x^{12} + \frac{3}{2} (V_x^{12})^2 - \delta x V_x^{12},$$

где $c_{12} = \sum_{i=1}^2 c_i$ и $\hat{c}_{12} = \sum_{i=1}^2 c_i^2$.

Предположим, что функция Беллмана [9, 11] имеет вид

$$(16) \quad V^{12}(x, t) = A(t)x + B(t),$$

тогда соответствующие частные производные будут записываться следующим образом: $V_t^{12} = \dot{A}(t)x + \dot{B}(t)$ и $V_x^{12} = A(t)$. Подставляя эти выражения в (15) и группируя подобные слагаемые, получаем систему двух дифференциальных уравнений:

$$(17) \quad \begin{cases} \dot{A}(t) = \delta A(t) + k_{12}, \\ \dot{B}(t) = -\frac{1}{2} \hat{c}_{12} - c_{12} A(t) - \frac{3}{2} A^2(t), \end{cases}$$

с краевыми условиями $A(T) = B(T) = 0$. Решение системы (17) имеет следующий вид:

$$\begin{cases} A(t) = -\frac{k_{12}}{\delta}[1 + e^{-\delta T} e^{\delta t}], \\ B(t) = \frac{\alpha_0(T-t) - \alpha_1(e^{-2\delta(T-t)} - 1) + \alpha_2(e^{-\delta(T-t)} - 1)}{4\delta^3}, \end{cases}$$

где $\alpha_0 = 2\delta(\hat{c}_{12}\delta^2 - 2c_{12}\delta k_{12} + 3k_{12}^2)$, $\alpha_1 = -3k_{12}^2$, $\alpha_2 = 4k_{12}(3k_{12} - c_{12}\delta)$.

Таким образом, оптимальные управления \bar{u}_i имеют вид $\bar{u}_i = c_i - \frac{k_{12}}{\delta}[1 - e^{-\delta T} e^{\delta t}]$, $i = 1, 2$. Отметим, что выполнение условия (12) гарантирует принадлежность оптимального управления \bar{u}_i интервалу $[0, u_i^{max}]$.

Оптимальная траектория $\bar{x}(t)$, соответствующая оптимальным управлениям (\bar{u}_1, \bar{u}_2) , имеет вид

$$(18) \quad \bar{x}(t) = \frac{c_{12}}{\delta} + e^{-\delta t} \left(x_0 - \frac{c_{12}}{\delta} + \frac{3e^{-T\delta}(2e^{T\delta} - 1)k_{12}}{2\delta^2} \right) - \frac{3e^{-\delta t}e^{-\delta(T-2t)}(2e^{\delta(T-t)} - 1)k_{12}}{2\delta^2}.$$

Далее, найдем выражение для значений характеристической функции $V(\{1\}, x_0, T - t_0)$ и $V(\{2\}, x_0, T - t_0)$. Найдем значение характеристической функции для игрока 2, действующего самостоятельно, т.е.

$$V(\{2\}, x_0, T - t_0) = \max_{u_2} \min_{u_1} K_2(x_0, u_1, u_2).$$

Для удобства будем использовать следующее сокращенное обозначение: $V(\{2\}, x(t), T - t) = V(\{2\})$. Запишем соответствующее уравнение Гамильтона–Якоби–Беллмана:

$$(19) \quad \begin{aligned} V(\{2\})_t &= \max_{u_2} \min_{u_1} \left\{ -k_2x + c_2u_2 - \frac{1}{2}u_2^2 + \right. \\ &\quad \left. + V(\{2\})_x \sum_{i=1}^2 u_i - \delta x V(\{2\})_x, \right. \end{aligned}$$

где $V(\{2\})_t$ и $V(\{2\})_x$ – частные производные функции Беллмана $V(\{2\}, x, T - t)$.

Максимизирующие управления имеют вид $u_i^o = c_i + V_x$, $i = 2$, а минимизирующее управление определяется как

$$u_1^o = \begin{cases} 0, & V(\{2\})_x \geq 0, \\ u_1^{max}, & V(\{2\})_x < 0. \end{cases}$$

После подстановки найденных управлений в (19) получаем

$$(20) \quad V(\{2\})_t = -k_2x + \frac{1}{2}\hat{c}_2 + c_2V(\{2\})_x + V(\{2\})_x^2 - \delta xV(\{2\})_x + V(\{2\})_x u_1^o,$$

где $\hat{c}_2 = c_2^2$.

Выберем функцию Беллмана в следующем виде [9]:

$$(21) \quad V(\{2\}, x, T-t) = A\{2\}(t)x + B\{2\}(t).$$

Тогда $V(\{2\})_t = \dot{A}\{2\}(t)x + \dot{B}\{2\}(t)$ и $V(\{2\})_x = A\{2\}(t)$. После подстановки частных производных в (20) и приведения подобных слагаемых получается система двух дифференциальных уравнений:

$$(22) \quad \begin{cases} \dot{A}\{2\}(t) = \delta A\{2\}(t) + k_2, \\ \dot{B}\{2\}(t) = -\frac{1}{2}\hat{c}_2 - c_2A\{2\}(t) - (A\{2\}(t))^2 - A\{2\}(t)u_1^o, \end{cases}$$

с краевыми условиями $A\{2\}(T) = B\{2\}(T) = 0$. Анализируя решение первого дифференциального уравнения заключаем, что $A(t) \leq 0$ для всех $t \in [t_0, T]$ и, следовательно, $u_1^o = u_1^{max}$. Можно легко показать, что выполнение условия (12) гарантирует выполнение требования $\bar{u}_i \in [0, u_i^{max}]$, $i = 2$.

Таким образом, характеристическая функция $V(\{2\})$ имеет вид $V(\{2\}, x, T-t) = A\{2\}(t)x + B\{2\}(t)$, где $A\{2\}(t)$, $B\{2\}(t)$ вычисляются следующим образом:

$$A\{2\}(t) = -\frac{k_2(1 - e^{-\delta(T-t)})}{\delta},$$

$$B\{2\}(t) = \frac{Tc_2^2\delta^3 - 2Tc_2\delta^2k_2 + 2c_2\delta k_2 - 2T u_{max}\delta^2k_2 + 2T\delta k_2^2 + 2u_{max}\delta k_2 - 3k_2^2}{2\delta^3} - \frac{t(c_2^2\delta^2 - 2c_2\delta k_2 - 2u_{max}\delta k_2 + 2k_2^2)}{2\delta^2} - \frac{k_2^2 e^{-2\delta(T-t)}}{2\delta^3} - \frac{k_2 e^{-\delta(T-t)}(c_2\delta - 2k_2 + \delta u_{max})}{\delta^3}.$$

Аналогично находится характеристическая функция для коалиции $S_1 = \{1\}$.

Очевидно, что характеристическая функция (16, 21) является супераддитивной по построению. Следовательно, вдоль оптимальной траектории $\bar{x}(t)$ (18) в любой момент времени $t \in [0, T]$ выполняется следующее неравенство:

$$(23) \quad V(\{1, 2\}, \bar{x}(t), T - t) \geq V(\{1\}, \bar{x}(t), T - t) + V(\{2\}, \bar{x}(t), T - t).$$

3.3. ПОСТРОЕНИЕ СИЛЬНО ДИНАМИЧЕСКИ УСТОЙЧИВОГО ПРИНЦИПА ОПТИМАЛЬНОСТИ

Построим сильно динамический принцип оптимальности. Для этого найдем разность значений характеристической функции $V(\{1, 2\}, \bar{x}(t), T - t)$ и $V(\{1\}, \bar{x}(t), T - t) + V(\{2\}, \bar{x}(t), T - t)$. По определению, эта разность соответствует $A_1(t) + A_2(t)$ (2), а по построению $(A_1(t) + A_2(t))$ является неотрицательной величиной:

$$(24) \quad A_1(t) + A_2(t) = V(\{1, 2\}, \bar{x}(t), T - t) - V(\{1\}, \bar{x}(t), T - t) - V(\{2\}, \bar{x}(t), T - t) \geq 0.$$

Получаем:

$$(25) \quad A_1(t) + A_2(t) = \frac{-k_1 k_2 (e^{-\delta(T-t)} - 1)^2 + (2k_1 k_2 - \delta\xi)(e^{-\delta(T-t)} - 1)}{\delta^3} + \frac{(2\delta k_1 k_2 - \delta^2 \xi)(T - t)}{\delta^3} + \frac{k_{12}(e^{-\delta(T-t)} - 1) + \delta k_{12}(T - t)}{\delta^2} u_{max},$$

где

$$(26) \quad \xi = c_1 k_2 + c_2 k_1.$$

Найдем выражение для $\gamma_1(t) + \gamma_2(t)$. Из определения (5) следует, что

$$(27) \quad \sum_{i=1}^2 \gamma_i(t) = -\frac{d}{dt} \sum_{i=1}^2 A_i(t) = -\frac{d}{dt} V(\{1, 2\}, \bar{x}(t), T - t) + \frac{d}{dt} \sum_{i=1}^2 V(\{i\}, \bar{x}(t), T - t).$$

Заметим, что функция $V(\{1, 2\}, \bar{x}(t), T - t) - V(\{1\}, \bar{x}(t), T - t) - V(\{2\}, \bar{x}(t), T - t)$ представляет собой разность неотрицательных убывающих функций, однако в общем случае мы не можем

гарантировать неотрицательность ее первой производной (так как разность убывающих функций не является убывающей функцией в общем случае).

Однако в нашем примере данное условие неотрицательности суммы $\gamma_1(t) + \gamma_2(t)$ (или, что эквивалентно, неположительность производной $\frac{d}{dt} \sum_{i=1}^2 A_i(t)$) легко гарантировать при помощи дополнительных ограничений на параметры модели.

Имеем:

$$\begin{aligned} \sum_{i=1}^2 \gamma_i(t) &= -\frac{d}{dt} \sum_{i=1}^2 A_i(t) = \\ &= \frac{2k_1k_2}{\delta^2} (e^{-\delta(T-t)} - 1)^2 + \delta \left(1 - e^{-\delta(T-t)}\right) (k_{12}u_{max} - \xi). \end{aligned}$$

Очевидно, что при $k_{12}u_{max} - \xi \geq 0$ мы гарантируем неотрицательность $\gamma_1(t) + \gamma_2(t)$. Поскольку ξ имеет вид (26), имеем

$$(28) \quad (k_1 + k_2)u_{max} \geq k_1c_2 + k_2c_1.$$

Теперь определим величины $A_1(t)$, $A_2(t)$ следующим образом:

$$(29) \quad \begin{aligned} A_i(t) &= \frac{-k_1k_2(e^{-\delta(T-t)} - 1)^2 + (2k_1k_2 - \delta\xi)(e^{-\delta(T-t)} - 1)}{2\delta^3} + \\ &+ \frac{(2\delta k_1k_2 - \delta^2\xi)(T-t)}{2\delta^3} + \frac{k_{12}(e^{-\delta(T-t)} - 1) + \delta k_{12}(T-t)}{2\delta^2} u_{max}, \\ & \quad i = 1, 2. \end{aligned}$$

Очевидно, что $A_i(t) \geq 0$, $i = 1, 2$, в силу (24). Тогда $\gamma_i(t) = -\frac{d}{dt} A_i(t) = \frac{k_1k_2}{\delta^2} (e^{-\delta(T-t)} - 1)^2 + \delta \left(1 - e^{-\delta(T-t)}\right) \frac{(k_{12}u_{max} - \xi)}{2}$, причем $\gamma_i(t) \geq 0$, $i = 1, 2$, в силу (28).

Введем функции ПРД $\bar{\beta}_i(t)$, $i = 1, 2$, по формуле (3). Тогда имеем

$$(30) \quad \begin{aligned} \beta_2(t) &= \gamma_2(t) - \frac{d}{dt} V(\{2\}, \bar{x}(t), T-t) = \\ &= -\frac{1}{4\delta^3} \left(\alpha_0(t) + \alpha_1(t)e^{-\delta(T-t)} + \alpha_2e^{-2\delta(T-t)} - \alpha_3(t)e^{-2T\delta} \right), \end{aligned}$$

где

$$\alpha_0(t) = -(5k_{12}^2 - k_1^2)\delta + 2(k_{12}u_1^{max} - k_1u_2^{max} + (T-t)k_1k_{12} + c_1k_2 + c_2k_{12})\delta^2 - 2((T-t)^2k_1^2 + (T-t)c_2k_1 + (T-t)k_1u_1^{max} + c_2^2)\delta^3,$$

$$\alpha_1(t) = (10k_{12}^2 - 2k_1^2)\delta - 2c_1\delta^2k_2 - 4c_2\delta^2k_{12} - 2\delta^2(k_{12}u_1^{max} - k_1u_2^{max}) - 4(T-t)\delta^2k_1k_{12},$$

$$\alpha_2 = -(5k_{12}^2 - k_1^2)\delta,$$

$$\alpha_3(t) = ((2k_{12} - k_1)(1 - e^{T\delta - \delta t}) + (T-t)\delta k_1 e^{T\delta - \delta t})(3k_{12}(1 + e^{2\delta t}) - 2(3k_{12} - c_{12}\delta)e^{T\delta} - 2\delta^2x_0e^{T\delta}).$$

Компонента ПРД $\beta_1(t)$ вычисляется аналогичным образом. Очевидно, что условие 2.1 выполнено.

Теперь рассмотрим вектор $\bar{\alpha} = (\alpha_1, \alpha_2)$, построенный по формуле (6), а именно $\bar{\alpha}_i(t) = \int_0^T \bar{\beta}_i(t)dt$, $i = 1, 2$. Как было доказано, он является дележом из c -ядра, кроме того, построенный таким образом дележ обеспечивает сильную динамическую устойчивость c -ядра.

Окончательно имеем следующий вид для дележа $\bar{\alpha} = (\alpha_1, \alpha_2)$:

$$(31) \quad \alpha_1 = A_1 + V(\{1\}, x_0, T - t_0);$$

$$(32) \quad \alpha_2 = A_2 + V(\{2\}, x_0, T - t_0) = \frac{1}{24\delta^3}(\eta_0(t)x_0 + \eta_1 + \eta_2(t)e^{-\delta(T-t)} + \eta_3(e^{-\delta(T-t)} - 1) + \eta_4e^{-2\delta(T-t)} + \eta_5(t)),$$

где

$$\eta_0(t) = -24\delta^2k_2 - 12\delta^3k_1t + 12\delta^2(k_1 + 2k_2)e^{-\delta(T-t)} + 12T\delta^2(\delta - 1)k_1,$$

$$\eta_1 = -66k_1k_2 - 45k_2^2 - 12k_1^2,$$

$$\eta_2(t) = 12(5k_{12}^2 - 3k_1^2 - 2k_1k_2) - 24(T-t)\delta k_1k_{12},$$

$$\eta_3 = -12\delta(c_1k_2 + 2c_2k_{12}) - 12\delta(k_{12}u_1^{max} - k_1u_2^{max}),$$

$$\eta_4 = 3k_1^2 - 15k_{12}^2,$$

$$\eta_5(t) = 30(T-t)\delta k_{12}^2 - 12T\hat{c}_2\delta^3(t-1) - 12(T-t)^2\delta^2k_1k_{12} + 4(T-t)^3\delta^3k_1^2 - 6(T-t)\delta k_1^2 -$$

$$-12(T-t)\delta^2(c_1k_2 + 2c_2k_{12}) + 6(T-t)^2c_2\delta^3 - 12(T-t)\delta^2(k_{12}u_1^{max} - k_1u_2^{max}) + 6(T-t)^2\delta^3k_1u_1^{max}.$$

Заметим, что в данном случае при выборе функций $A_1(t)$, $A_2(t)$ мы разделили между игроками сумму $A_1(t) + A_2(t)$ попо-

лам. Таким образом, для нового дележа имеем:

$$\begin{aligned}
 \bar{\alpha}_1 &= A_1 + V(\{1\}, x_0, T - t_0) = \\
 &= \frac{1}{2}(A_1 + A_2) + V(\{1\}, x_0, T - t_0) = \\
 &= \frac{1}{2}V(\{1, 2\}, x_0, T - t_0) - V(\{1\}, x_0, T - t_0) - \\
 &\quad - V(\{2\}, x_0, T - t_0) + V(\{1\}, x_0, T - t_0) = \\
 &= \frac{1}{2}(V(\{1, 2\}, x_0, T - t_0) - V(\{2\}, x_0, T - t_0)) + \\
 &\quad + \frac{1}{2}(V(\{1\}, x_0, T - t_0)), \\
 \bar{\alpha}_2 &= A_2 + V(\{2\}, x_0, T - t_0) = \\
 &= \frac{1}{2}(A_1 + A_2) + V(\{2\}, x_0, T - t_0) = \\
 &= \frac{1}{2}V(\{1, 2\}, x_0, T - t_0) - V(\{1\}, x_0, T - t_0) - \\
 &\quad - V(\{2\}, x_0, T - t_0) + V(\{2\}, x_0, T - t_0) = \\
 &= \frac{1}{2}(V(\{1, 2\}, x_0, T - t_0) - V(\{1\}, x_0, T - t_0)) + \\
 &\quad + \frac{1}{2}(V(\{2\}, x_0, T - t_0)),
 \end{aligned}$$

т.е. вектор $\bar{\alpha}$ является вектором Шепли в игре с характеристической функцией $V(\{2\}, x_0, T - t_0)$.

Очевидно, что при выполнении условия $A_1(t) + A_2(t) \geq 0$ предложенный выбор функций $A_1(t) \geq 0$, $A_2(t) \geq 0$ не является единственно возможным. Величина $A_1(t) + A_2(t)$ могла быть разделена между игроками в любых пропорциях, однако в таком случае дележ $\bar{\alpha}$ не являлся бы вектором Шепли.

Используя ПРД (30) при выполнении дополнительного ограничения на параметры модели (28), мы обеспечиваем динамическую и сильную динамическую устойчивость ядра на основе вектора Шепли (31). Сильная динамическая устойчивость ядра как кооперативного решения в данном примере соответствует тому, что однократное отклонение от соглашения о совместном контроле объемов вредных выбросов не приводит к нереализуемости соглашения действовать совместно оптимально на протяжении всей игры, поскольку при выборе другого дележа из ядра в подыгре, начинающейся в момент нарушения соглашения, суммарные выплаты игрокам все равно принадлежат изначально выбранному принципу оптимальности (ядру). Таким образом, кооперативное соглашение по контролю объемов вредных выбросов не нарушается.

Кроме того, обеспечивается защита от иррационального по-

ведения участников, поскольку ранее было доказано, что в случае игры двух лиц выбор ПРД $\beta_i(t)$ по формуле (30) гарантирует выполнение условия Янга (9).

4. Заключение

В работе предложен конструктивный метод построения сильно динамически устойчивого кооперативного решения для дифференциальной игры двух лиц, а именно, сформулированы достаточно общие условия, при которых в C -ядре можно выделить дележ, однократное отклонение от которого в пользу другого дележа из C -ядра в подыгре, начинающейся в момент отклонения от первоначально выбранного дележа, не приводит к нереализуемости первоначально выбранного соглашения о разделе суммарного выигрыша согласно C -ядру. Кроме того, доказано, что в играх двух лиц данное кооперативное решение будет также защищено от иррационально поведения участников, т.е. от однократного отклонения от кооперативного соглашения по некоторым иррациональным причинам.

Полученные теоретические результаты проиллюстрированы на примере дифференциальной игры управления вредными выбросами с двумя участниками, а именно, в данной задаче получено аналитическое решение, гарантирующее сохранение кооперации в сильно динамическом смысле.

Литература

1. ВОРОБЬЕВ Н.Н. *Теория игр для экономистов-кибернетиков*. – М.: Наука, 1985. – 272 с.
2. КРАСОВСКИЙ Н.Н., КОТЕЛЬНИКОВА А.Н. *О дифференциальной игре на перехват* // Труды математического института им. В.А. Стеклова. – 2010. – Т. 268. – С. 168–214.
3. ПЕТРОСЯН Л.А. *О новых сильно динамически устойчивых принципах оптимальности в кооперативных дифференциальных играх* // Труды математического института

- им. Стеклова «Оптимальное управление и дифференциальные уравнения». – 1995. – Т. 211. – С. 370–376.
4. ПЕТРОСЯН Л.А. *Характеристические функции в кооперативных дифференциальных играх* // Вестник СПбГУ, сер. 1: Математика, механика, астрономия. – 1995. – №1. – С. 48–52.
 5. ПЕТРОСЯН Л.А. *Устойчивость решений в дифференциальных играх со многими участниками* // Вестник ЛГУ. – 1977. – №4. – С. 46–52.
 6. ПЕТРОСЯН Л.А., ДАНИЛОВ Н.Н. *Устойчивые решения неантагонистических дифференциальных игр с транзитивными выигрышами* // Вестник ЛГУ. – 1979. – №1. – С. 46–54.
 7. ПЕТРОСЯН Л.А., ДАНИЛОВ Н.Н. *Кооперативные дифференциальные игры и их приложения*. – Томск: Изд-во Томского университета, 1985. – 273 с.
 8. ПЕТРОСЯН Л.А., ГРОМОВА Е.В. *Двухуровневая кооперация в коалиционных дифференциальных играх*. // Тр. ИММ УрО РАН. – 2014. – №3. – С. 193–203.
 9. BASAR T., OLSDER G.J. *Dynamic Noncooperative Game Theory*. – 2nd Edition, SIAM, 1999. – 511 p.
 10. BRETON M., ZACCOUR G., ZAHAF M. *A differential game of joint implementation of environmental projects* // Automatica. – 2005. – Vol.41(10). – P. 1737–1749.
 11. ENGWERDA J. *LQ Dynamic Optimization and Differential Games* – Wiley, 2005. – 510 p.
 12. NEUMANN J., MORGENSTERN O. *Theory of Games and Economic Behavior*. – Princeton, 1947. – 610 p.
 13. PETROSYAN L., ZACCOUR G. *Time-consistent Shapley value allocation of pollution cost reduction* // Journal of Economic Dynamics and Control. – 2003. – Vol. 27. – P. 381–398.
 14. YEUNG D.W.K., PETROSJAN L.A. *Cooperative Stochastic Differential Games*. – New-York, Heidelberg, London: Springer, 2006. – 242 p.

15. YEUNG D.W.K. *An irrational-behavior-proof condition in cooperative differential games* // International Game Theory Review (IGTR). – 2006. – Vol. 08, Is. 04. – P. 739–744.

STRONGLY TIME-CONSISTENT COOPERATIVE SOLUTION FOR A DIFFERENTIAL GAME OF POLLUTION CONTROL

Ekaterina Gromova, St. Petersburg State University, Cand.Sc., assistant professor (ekaterina.shevkoplyas@gmail.com).

Leon Petrosyan, St. Petersburg State University, Doctor of Science, professor (Universitetskii prospekt 35, Petergof, Saint-Petersburg, Russia 198504, tel. +7 (812) 428-71-59).

Abstract: We find a strongly time-consistent cooperative solution for two-payer differential games. The cooperative solution found satisfies strong time-consistency property under rather general conditions. It is shown that the obtained solution also satisfies the condition of irrational-behavior proofness. The obtained game-theoretical results are illustrated by a differential game of pollution control.

Keywords: differential games, time-consistency, strongly time-consistency, Yeung's condition, model of pollution control.

Статья представлена к публикации членом редакционной коллегии Д.А. Новиковым

Поступила в редакцию 07.03.2015.

Дата опубликования 31.05.2015.