Прикладная математика и механика. Том 48, вып. 4, 1984

УДК 62—50

АНАЛИЗ ОДНОЙ ИЕРАРХИЧЕСКОЙ ДИФФЕРЕНЦИАЛЬНОЙ ИГРЫ ДВУХ ЛИЦ

Клейменов А. Ф.

Предлагается постановка иерархической дифференциальной игры двух лиц с динамикой, описываемой нелинейным дифференциальным уравнением достаточно общего вида, и с терминальными функциями выигрыша. На основе введенной в [1, 2] формализации антагонистической дифференциальной игры дается определение оптимальных стратегий и выявляется их структура. Описываются оптимальные стратегии, образующие паретовскую точку ([3], с. 57) множества равновесных коалиционных стратегий, неулучшаемую для игрока верхнего уровня. Основные предположения (объявление игроком верхнего уровня стратегии до начала игры, рациональность выбора стратегии игроком нижнего уровня) восходят к работам [4, 5], в которых рассматривались статические модели. Исследования иерархических динамических игр проводились, в частности, в [6—8]. Данная работа примыкает к исследованиям [1, 2, 6, 9—11].

1. Пусть управляемая система описывается уравнением вида

$$x^{\cdot} = f(t, x, u, v), \quad u \in P, \quad v \in Q \tag{1.1}$$

где x — n-мерный фазовый вектор, u и v — векторные управления первого и второго игроков размерности p и q соответственно, P и Q — компакты. Функция $f: G \times P \times Q \mapsto R^n$ непрерывна по совокупности аргументов и удовлетворяет условию Липшица по x. Здесь G — компакт в $R^1 \times R^n$, проекция которого на ось t равна заданному отрезку $[t_0, \vartheta]$; предполагается, что все траектории системы (1.1), начинающиеся в произвольной позиции $(t_*, x_*) \in G$, остаются в G при всех $t \in [t_*, \vartheta]$.

Первый игрок стремится выбрать свое управление u в текущий момент времени t так, чтобы при переводе системы из начальной позиции $(t_*, x_*) \in G$ в состояние $x[\vartheta]$ минимизировать показатель $\sigma_1(x[\vartheta])$; с другой стороны, второй игрок стремится минимизировать показатель $\sigma_2(x[\vartheta])$. Здесь $\sigma_i: R^n \mapsto R^1$ $(i = 1, 2)$ — заданные непрерывные функции. Предполагается, что обоим игрокам известен фазовый вектор системы $x[t]$ в текущий момент времени.

Далее в неантагонистической дифференциальной игре действия игроков будут формализоваться в тех же классах, что и в общей теории позиционных антагонистических дифференциальных игр [1, 2]. Именно, в общем случае, когда не выполнено условие седловой точки в маленькой игре ([1], с. 56), в зависимости от информированности игроков о реализующихся управлениях партнера могут быть следующие пары классов действий игроков: {чистые стратегии первого игрока — контрстратегии второго игрока}, {смешанные стратегии первого игрока — смешанные стратегии второго игрока} и {контрстратегии первого игрока — чистые стратегии второго игрока}. Для простоты изложение ведется в случае, когда для функции f выполнено условие седловой точки в маленькой игре. Тогда действия обоих игроков рассматриваются в классах чистых стратегий. Однако полученные ниже результаты справедливы и в общем случае.

Стратегию первого игрока отождествляем с парой $U = \{u(t, x, \varepsilon), \beta_1(\varepsilon)\}$, где $u(t, x, \varepsilon)$ — произвольная функция, заданная при $(t, x) \in G$, $\varepsilon > 0$ со значениями в P, а $\beta_1(\cdot) \in A(0, \infty)$. Через $A(0, \infty)$ обозначен класс функций β: $(0, \infty) \mapsto (0, \infty)$, непрерывных и монотонных, удовлетворяющих условию $\beta(\varepsilon) \to 0$ при $\varepsilon \to 0$. Стратегия как функция позиции (t, x) и параметра точности ε была введена в [2]. Добавление функции $\beta_1(\cdot)$ носит технический характер; его смысл будет обсужден ниже (конец п. 2). Аналогично, в качестве стратегии второго игрока принимаем пару $V = \{v(t, x, \varepsilon), \beta_2(\varepsilon)\}$, где функция $v(t, x, \varepsilon)$ определена при $(t, x) \in G$, $\varepsilon > 0$ и принимает значения в Q, а $\beta_2(\cdot) \in A(0, \infty)$.

Далее предполагаем, что начальная позиция $(t_*, x_*) \in G$ фиксирована. Пусть заданы стратегии U и V, а также ε_1 и ε_2 — значения параметра точности ε, выбранные соответственно первым и вторым игроками. Пусть $\Delta_1 = \{\tau_i^{(1)}\}$ и $\Delta_2 = \{\tau_j^{(2)}\}$ — разбиения отрезка $[t_*, \vartheta]$ системами непересекающихся полуинтервалов $[\tau_i^{(1)}, \tau_{i+1}^{(1)})$ и $[\tau_j^{(2)}, \tau_{j+1}^{(2)})$, выбранные первым и вторым игроками соответственно, причем выполнены условия $\delta(\Delta_s) \leqslant \beta_s(\varepsilon_s)$ $(s = 1, 2)$, где $\delta(\Delta_s)$ — шаг разбиения Δ_s, равный $\max_i (\tau_{i+1}^{(s)} - \tau_i^{(s)})$.

Ломаной Эйлера, порожденной стратегиями U и V из начальной позиции (t_*, x_*) при выбранных значениях ε_1 и ε_2 и разбиениях Δ_1 и Δ_2, назовем непрерывную функцию

$$x_\Delta{}^\varepsilon[t] = x_{\Delta_1, \Delta_2}^{\varepsilon_1, \varepsilon_2}[t, t_*, x_*, U, V]$$

удовлетворяющую пошаговому дифференциальному уравнению

$$\dot{x}_\Delta^\varepsilon[t] = f(t, x_\Delta{}^\varepsilon[t], u(\tau_i^{(1)}, x_\Delta{}^\varepsilon[\tau_i^{(1)}], \varepsilon_1), v(\tau_j^{(2)}, x[\tau_j^{(2)}], \varepsilon_2))$$
$$t \in [\tau_i^{(1)}, \tau_{j+1}^{(2)}) \quad (\tau_j^{(2)} \leqslant \tau_i^{(1)} < \tau_{j+1}^{(2)} \leqslant \tau_{i+1}^{(1)})$$

Движением, порожденным стратегиями U и V из начальной позиции (t_*, x_*), назовем непрерывную функцию $x[t] = x[t, t_*, x_*, U, V]$, для которой найдется последовательность ломаных Эйлера $x_{\Delta_1^k, \Delta_2^k}^{\varepsilon_1^k, \varepsilon_2^k}[t, t_*^k, x_*^k, U, V]$, равномерно сходящаяся к $x[t]$ на $[t_*, \vartheta]$ при $k \to \infty$, $\varepsilon_1^k \to 0$, $\varepsilon_2^k \to 0$, $t_*^k \to t_*$, $x_*^k \to x_*$, $\delta(\Delta_1^{(s)}) \leqslant \beta_s(\varepsilon_1^s)$, $s = 1, 2$. Далее, не ограничивая общности, будем считать, что $\varepsilon_2^k \leqslant \varepsilon_1^k$. Действительно, возможности второго игрока при этом не ограничиваются, поскольку «масштаб» его параметра точности может быть соответствующим образом изменен. Пара стратегий порождает, вообще говоря, пучок движений, который обозначим через $X(t_*, x_*, U, V)$. Этот пучок будет компактом в $C[t_*, \vartheta]$.

Считаем, что оба игрока имеют полную информацию о системе, т. е. они знают уравнение (1.1), множества P и Q, целевые функции σ_1 и σ_2.

Введем следующие предположения.

А. Первый игрок выбирает свою стратегию $U^* = \{u^*(t, x, \varepsilon), \beta_1^*(\varepsilon)\}$ до начала игры и сообщает ее второму игроку.

Б. Второй игрок, получив информацию о выбранной первым игроком стратегии U^*, выбирает рациональную стратегию $V^* = \{v^*(t, x, \varepsilon), \beta_2^*(\varepsilon)\}$ из условия

$$\min_V \max_{x[\cdot] \in X(t_*, x_*, U^*, V)} \sigma_2(x[\vartheta, t_*, x_*, U^*, V]) = \tag{1.2}$$
$$= \max_{x[\cdot] \in X(t_*, x_*, U^*, V^*)} \sigma_2(x[\vartheta, t_*, x_*, U^*, V^*]) = \rho^*(t_*, x_*, U^*)$$

При выполнении предположений А, Б первого игрока будем называть игроком верхнего уровня иерархии, второго игрока — игроком нижнего уровня, а рассматриваемую дифференциальную игру — иерархической.

Вопрос о существовании стратегии V^* в общем случае здесь обсуждаться не будет. Отметим лишь, что для стратегий U^*, которые будут рассматриваться ниже, рациональные стратегии V^* существуют. Множество рациональных стратегий второго игрока, отвечающих объявленной стратегии U^* первого игрока, обозначим через $K(t_*, x_*, U^*)$.

Справедливо следующее свойство рациональных стратегий: для любой стратегии $V^* \in K(t_*, x_*, U^*)$ можно указать движение $x_*[\cdot] \in \in X(t_*, x_*, U^*, V^*)$, для которого

(1.3) $$\sigma_2(x_*[\vartheta]) \leqslant \gamma_2(t, x_*[t]), \quad \forall t \in [t_*, \vartheta]$$

Здесь $\gamma_2(t, x)$ — непрерывная при $(t, x) \in G$ функция цены антагонистической игры Γ_2, динамика которой описывается уравнением (1.1) и в которой второй игрок, распоряжающийся управлением v, стремится минимизировать показатель $\sigma_2(x[\vartheta])$, а первый игрок ему противодействует. Известно [2], что такая игра имеет универсальную седловую точку

(1.4) $$U_2^\circ = \{u^{(2)\circ}(t, x, \varepsilon),\ \beta_{12}^\circ(\varepsilon)\}$$
$$V_2^\circ = \{v^{(2)\circ}(t, x, \varepsilon),\ \beta_{22}^\circ(\varepsilon)\}$$

Для того чтобы сформулировать условие (1.2) в аппроксимационной форме, сделаем еще одно предположение.

В. Первый игрок одновременно с началом игры выбирает значение ε_1 своего параметра точности и сообщает его второму игроку.

Если предположение В выполнено, то условие (1.2) означает, что для любого $\zeta > 0$ можно указать $\varkappa(\zeta) > 0$, такое, что для любых $\varepsilon_1 \leqslant \leqslant \varkappa(\zeta)$, $\varepsilon_2 \leqslant \varepsilon_1$, любых разбиений Δ_1, Δ_2, $\delta(\Delta_s) \leqslant \beta_s^*(\varepsilon_s)$, $s = 1, 2$ имеет место неравенство

(1.5) $$\sigma_2(x_{\Delta_1, \Delta_2}^{\varepsilon_1, \varepsilon_2}[\vartheta, t^*, x^*, U^*, V^*]) \leqslant \rho^*(t_*, x_*, U^*) + \zeta$$
$$((|t^* - t_*|^2 + \|x^* - x_*\|^2)^{1/2} \leqslant \min(\beta_1^*(\varepsilon_1), \beta_2^*(\varepsilon_2))$$

То обстоятельство, что второму игроку сообщается значение ε_1 лишь одновременно с началом игры, не позволяет ему использовать эту информацию для уточнения своей рациональной стратегии. Информация об ε_1 используется вторым игроком для выбора значения ε_2 своего параметра точности с целью обеспечить неравенство $\varepsilon_2 \leqslant \varepsilon_1$ при построении ломаных Эйлера.

Итак, при объявлении первым игроком стратегии U второй игрок имеет в своем распоряжении множество рациональных стратегий $K(t_*, x_*, U)$, причем любая стратегия $V \in K(t_*, x_*, U)$ гарантирует второму игроку результат, равный $\rho^*(t_*, x_*, U)$. В то же время результат, получаемый первым игроком, зависит, вообще говоря, от того, какую стратегию из множества $K(t_*, x_*, U)$ выбирает второй игрок.

Будем различать два случая.

Случай 1. Второй игрок выбирает рациональную стратегию из множества $K(t_*, x_*, U)$ произвольным образом.

Очевидно, в случае 1 гарантированный результат первого игрока при объявлении им стратегии U будет

(1.6) $$\sup_{V \in K(t_*, x_*, U)} \ \max_{x[\cdot] \in X(t_*, x_*, U, V)} \sigma_1(x[\vartheta, t_*, x_*, U, V]) = \rho^{(1)}(U)$$

Случай 2. Второй игрок проявляет доброжелательность по отношению к первому игроку, выбирая рациональную стратегию V из условия

$$(1.7)\qquad \min_{V_*\in K(t_*,x_*,U)}\ \max_{x[\cdot]\in X(t_*,x_*,U,V_*)} \sigma_1(x[\vartheta,t_*,x_*,U,V_*]) = \max_{x[\cdot]\in X(t_*,x_*,U,V)} \sigma_1(x[\vartheta,t_*,x_*,U,V]) = \rho^{(2)}(U)$$

В (1.7) записываем min, а не inf, имея в виду, что для тех U, с которыми будем иметь дело, минимум на множестве $K(t_*, x_*, U)$ достигается. Величина $\rho^{(2)}(U)$ будет гарантированным результатом первого игрока при объявлении им стратегии в случае 2.

Условимся в дальнейшем говорить просто об иерархической дифференциальной игре, если имеет место случай 1, и об иерархической дифференциальной игре с доброжелательным вторым игроком, если имеет место случай 2.

Задача 1. Найти стратегию первого игрока $U^\circ = \{u^\circ(t, x, \varepsilon), \beta_1^\circ(\varepsilon)\}$, такую, что

$$(1.8)\qquad \rho^{(1)}(U^\circ) = \min_U \rho^{(1)}(U)$$

Задача 2. Найти стратегию первого игрока $U_0 = \{u_0(t, x, \varepsilon), \beta_{10}(\varepsilon)\}$, такую, что

$$(1.9)\qquad \rho^{(2)}(U_0) = \min_U \rho^{(2)}(U)$$

Определение 1. Стратегию U°, являющуюся решением задачи 1, будем называть оптимальной стратегией первого игрока в иерархической дифференциальной игре. Оптимальной стратегией второго игрока будем называть любую стратегию из множества $K(t_*, x_*, U^\circ)$.

Определение 2. Стратегию U_0, являющуюся решением задачи 2, будем называть оптимальной стратегией первого игрока в иерархической дифференциальной игре с доброжелательным вторым игроком. Оптимальной стратегией доброжелательного второго игрока будем называть любую стратегию из множества $K(t_*, x_*, U_0)$, удовлетворяющую условию (1.7).

Сформулируем следующую вспомогательную задачу оптимального управления.

Задача 3. Пусть динамика управляемой системы описывается уравнением (1.1). Найти пару измеримых функций $(u(t), v(t), t_* \leqslant t \leqslant \vartheta)$, доставляющую минимум показателю $\sigma_1(x(\vartheta))$ при условии

$$(1.10)\qquad \sigma_2(x(\vartheta)) \leqslant \gamma_2(t, x(t)),\quad t_* \leqslant t \leqslant \vartheta$$

где $\gamma_2(t, x)$ — функция цены антагонистической игры Γ_2, а $x(t)$, $t_* \leqslant t \leqslant \vartheta$ — траектория системы (1.1), порожденная управлениями $u(\cdot)$ и $v(\cdot)$ из начальной позиции (t_*, x_*).

Траектории системы (1.1), удовлетворяющие неравенству (1.10), назовем допустимыми. Можно показать, что множество допустимых траекторий непусто и является компактным в метрике пространства $C[t_*, \vartheta]$, если дополнительно предположить, что вектограмма

$$(1.11)\qquad Q^*(t, x) = \{q \in R^n : q = f(t, x, u, v),\ u \in P,\ v \in Q\}$$

выпукла. Тогда решение задачи 3 существует.

2. Пусть пара измеримых функций $u^*(\cdot)$, $v^*(\cdot)$ доставляет решение задачи 3; $x^*(\cdot)$ — соответствующая траектория. Используя теорему Лузина, можно найти семейства непрерывных функций $\{u^\varepsilon(\cdot)\}$, $\{v^\varepsilon(\cdot)\}$, таких, что $\| x^\varepsilon(t) - x^*(t) \| \leqslant \varepsilon$ при всех $t \in [t_*, \vartheta]$. Здесь $x^\varepsilon(\cdot)$ — траектория системы (1.1), порожденная управлениями $u^\varepsilon(\cdot)$, $v^\varepsilon(\cdot)$.

Рассмотрим стратегии первого и второго игроков $U^* = \{u^*(t, x, \varepsilon), \beta_1^*(\varepsilon)\}$ и $V^* = \{v^*(t, x, \varepsilon), \beta_1^*(\varepsilon)\}$, где

$$u^*(t, x, \varepsilon) = \begin{cases} u^\varepsilon(t), & \| x - x^\varepsilon(t) \| \leqslant \varepsilon \\ u^{(2)\circ}(t, x, \varepsilon), & \| x - x^\varepsilon(t) \| > \varepsilon \end{cases} \tag{2.1}$$

$$v^*(t, x, \varepsilon) = \begin{cases} v^\varepsilon(t), & \| x - x^\varepsilon(t) \| \leqslant \varepsilon \\ v^{(2)\circ}(t, x, \varepsilon), & \| x - x^\varepsilon(t) \| > \varepsilon \end{cases} \tag{2.2}$$

$$\forall t \in [t_*, \vartheta], \quad \varepsilon > 0$$

а мажоранта $\beta_1^*(\varepsilon)$, общая для обеих стратегий, выбирается так, чтобы обеспечить выполнение неравенства

$$\| x^{\varepsilon, \varepsilon_2}_{\Delta_1, \Delta_2}[t, t_*, x_*, U^*, V^*] - x^\varepsilon(t) \| < \varepsilon, \quad \forall t \in [t_*, \vartheta] \tag{2.3}$$

для ломаных Эйлера при $\varepsilon_2 \leqslant \varepsilon$, $\delta(\Delta_1) \leqslant \beta_1^*(\varepsilon)$, $\delta(\Delta_2) \leqslant \beta_1^*(\varepsilon_2)$. Функции $u^{(2)\circ}(t, x, \varepsilon)$, $v^{(2)\circ}(t, x, \varepsilon)$ определены в (1.4).

Можно проверить, что $V^* \in K(t_*, x_*, U^*)$. Заметим, что пучок $X(t_*, x_*, U^*, V^*)$ состоит из единственной траектории $x^*(\cdot)$.

Обозначим $\tau_* = \min\{\tau \in [t_*, \vartheta] : \sigma_2(x^*(\vartheta)) = \gamma_2(\tau, x^*(\tau))\}$

Могут представиться два случая.

а) $\tau_* = \vartheta$. Содержательно это означает, что второй игрок, получив информацию об объявленной стратегии первого игрока U^*, заинтересован в отслеживании траектории $x^*(\cdot)$ вплоть до окончания игры. Отметим, что рассматривавшаяся в [11] задача относится именно к этому случаю.

Теорема 1. Пусть выполнены предположения А, Б. Пусть пара измеримых функций $u^*(\cdot)$, $v^*(\cdot)$ решает задачу 3, причем для соответствующей траектории выполнено условие $\tau_* = \vartheta$. Тогда стратегии U^*, V^* (2.1)—(2.3) являются оптимальными в иерархической дифференциальной игре.

Доказательство. Как уже отмечалось, $V^* \in K(t_*, x_*, U^*)$. Любая другая стратегия из множества $K(t_*, x_*, U^*)$ также обеспечивает отслеживание траектории $x^*(\cdot)$. Следовательно, объявляя U^*, первый игрок гарантирует себе результат $\rho^{(1)}(U^*) = \sigma_1(x^*(\vartheta))$. Покажем, что этот результат первый игрок не может улучшить. Допустим, что это не так и существуют стратегия первого игрока U^+ и число $\mu > 0$, такие, что для любой стратегии второго игрока $V^+ \in K(t_*, x_*, U^+)$, любого движения $x[\cdot] \in X(t_*, x_*, U^+, V^+)$ выполнено неравенство

$$\sigma_1(x[\vartheta]) \leqslant \rho^{(1)}(U^*) - \mu \tag{2.4}$$

В частности, (2.4) выполняется для движения $x^+[\cdot] \in X(t_*, x_*, U^+, V^+)$, для которого имеет место неравенство (1.3), т. е. $\sigma_2(x^+[\vartheta]) \leqslant \gamma_2(t, x^+[t])$ при всех $t \in [t_*, \vartheta]$.

Учитывая предположение о выпуклости вектограммы $Q^*(t, x)$ (1.9), можно заключить, что существуют измеримые функции $u_*(t)$, $v_*(t)$, $t_* \leqslant t \leqslant \vartheta$, порождающие из начальной позиции (t_*, x_*) траекторию $x_*(\cdot)$, которая совпадает с траекторией $x^+[\cdot]$. Отсюда следует, что траектория $x_*(\cdot)$ допустима в задаче 3. Но, поскольку для нее выполняется неравенство (2.4), это противоречит тому, что $u^*(\cdot)$, $v^*(\cdot)$ — решение задачи 3.

б) $\tau_* < \vartheta$. Теперь уже нельзя утверждать, что стратегии U^*, V^* (2.1) — (2.3) будут оптимальными в иерархической дифференциальной игре, поскольку второй игрок заинтересован в отслеживании траектории $x^*(\cdot)$ лишь до момента τ_*. Начиная же с момента τ_* он может «переключиться» на стратегию $V^{(2)^\circ}$ (1.4), гарантируя себе результат $\sigma_2(x^*(\vartheta))$. Такое переключение может оказаться нежелательным для первого игрока. Если же рассматривать иерархическую дифференциальную игру с доброжелательным вторым игроком, то верна следующая

Теорема 2. Пусть выполнены предположения А, Б. Пусть пара измеримых функций $u^*(\cdot)$, $v^*(\cdot)$ решает задачу 3, причем для соответствующей траектории выполнено неравенство $\tau_* < \vartheta$. Тогда стратегии U^*, V^* (2.1) — (2.3) являются оптимальными в иерархической дифференциальной игре с доброжелательным вторым игроком.

Теорема доказывается по той же схеме, что и теорема 1, с вполне понятными изменениями.

Таким образом, можно сделать следующий вывод. Если среди решений задачи 3 существует такое $u^*(\cdot)$, $v^*(\cdot)$, для которого $\tau_* = \vartheta$, то задача 1 имеет решение — стратегию U^* (2.1), (2.3). Стратегия U^* тем более будет решением задачи 2. Если же для любого решения задачи 3 имеет место неравенство $\tau_* < \vartheta$, то стратегия U^* (2.1), (2.3), построенная для произвольного решения $u^*(\cdot)$, $v^*(\cdot)$, будет решением задачи 2. При этом задача 1 не имеет решения в классе рассматриваемых стратегий. Однако на основе U^* может быть построена минимизирующая последовательность стратегий.

Поясним теперь на примере стратегий U^* и V^* (2.1) — (2.3) смысл введенных в начале п. 1 при определении стратегий функций $\beta_1(\cdot)$ и $\beta_2(\cdot)$. Как уже отмечалось, стратегия U^* сконструирована так, что, объявив ее, первый игрок стремится побудить второго игрока отслеживать траекторию $x^*(\cdot)$. При этом ломаные Эйлера, в формировании которых участвуют оба игрока, не должны выходить за пределы ε-трубок, построенных вдоль соответствующих траекторий. Объявление первым игроком подходяще выбранной функции $\beta_1(\cdot)$ позволяет второму игроку выбрать, в свою очередь, функцию $\beta_2(\cdot)$ так, что при разбиениях Δ_1 и Δ_2 с шагами, не превосходящими соответственно $\beta_1(\varepsilon_1)$ и $\beta_2(\varepsilon_2)$, где ε_1 и ε_2 — выбранные значения параметров точности игроков ($\varepsilon_2 \leqslant \varepsilon_1$), ломаные Эйлера не будут выходить за пределы упомянутых трубок.

Замечания. 1°. Идея использования стратегии наказания при конструировании оптимальной стратегии игрока верхнего уровня была предложена в [5] для статических иерархических игр и получила развитие в [6] для динамических иерархических игр.

2°. При построении стратегии U^* (2.1) существенно, что стратегия наказания, задаваемая функцией $u^{(2)^\circ}(t, x, \varepsilon)$, является универсальной [2], т. е. пригодной для всех позиций, которые могут встретиться во время игры.

3°. От предположения выпуклости вектограммы $Q^*(t, x)$ (1.9) можно отказаться. Тогда решение задачи 3 существует в классе функций со значениями во множестве вероятностных мер на $P \times Q$.

3. Стратегии U^*, V^* (2.1) — (2.3) образуют равновесную коалиционную стратегию двух игроков в терминологии работ [9, 10]. Это может быть и равновесная коалиционная смешанная стратегия [9] и равновесная коалиционная контрстратегия [10] в зависимости от того, какие принимаются предположения относительно информированности игроков о реализующихся управлениях партнера, а также от того, в классе измеримых функций или в классе функций-мер достигается решение задачи 3.

Правда, в [9, 10] равновесные коалиционные стратегии имели более сложную структуру, обусловленную тем, что игроков больше, чем два, а наказание игрока-уклониста, формируемое остальными игроками, зависело от индекса нарушителя, сообщаемого извне. В рассматриваемой же игре двух лиц наказание второго игрока-уклониста включено в конструкцию стратегии первого игрока U^*, а первый игрок, естественно, не выигрывает при уклонении от стратегии U^*.

Найдем среди решений задачи 3 такое, которое доставляет минимум показателю $\sigma_2(x(\vartheta))$, и обозначим его через $u^p(\cdot)$, $v^p(\cdot)$. При предположениях работы такое решение существует. Обозначим через U^p, V^p стратегии первого и второго игроков, которые получаются соответственно из U^* и V^* (2.1) — (2.3) заменой $u^*(\cdot)$ на $u^p(\cdot)$, а $v^*(\cdot)$ на $v^p(\cdot)$. Можно проверить, что стратегии U^p, V^p образуют паретовскую точку множества равновесных коалиционных стратегий, неулучшаемую для первого игрока.

4. Пусть движение управляемой системы описывается уравнением вида

(4.1) $$\frac{d^2\xi}{dt^2} = L(\varphi)F_1 + F_2$$

$$\xi = \begin{Vmatrix}\xi_1\\ \xi_2\end{Vmatrix}, \quad L(\varphi) = \begin{Vmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{Vmatrix}, \quad F_i = \begin{Vmatrix}F_{i1}\\ F_{i2}\end{Vmatrix}$$

Выбор вектора F_1 производится первым игроком, вектора F_2 и скалярной величины φ — вторым игроком, т. е.

$$u = F_1, \quad v = \begin{Vmatrix}F_2\\ \varphi\end{Vmatrix}$$

Управления стеснены ограничениями

(4.2) $$\|F_i\|^2 = F_{i1}^2 + F_{i2}^2 \leqslant 1, \quad |\varphi| \leqslant \varphi_0, \quad 0 < \varphi_0 < \pi/2$$

Заданы начальные условия $\xi[t_*] = \xi_*$, $\xi^\cdot[t_*] = \xi_*^\cdot$ и момент окончания игры ϑ. Первый (второй) игрок стремится минимизировать показатель $\sigma_1(\xi[\vartheta])$ $(\sigma_2(\xi[\vartheta]))$ вида

(4.3) $$\sigma_i(\xi[\vartheta]) = \|\xi[\vartheta] - a^{(i)}\|, \quad i = 1, 2$$

где $a^{(i)}$ — заданные точки в плоскости (ξ_1, ξ_2).

Уравнение (4.1) можно трактовать как уравнение движения материальной точки единичной массы в плоскости (ξ_1, ξ_2) под действием силы, формируемой двумя игроками. Можно проверить, что условие седловой точки в маленькой игре для системы (4.1) не выполняется, поэтому классы стратегий игроков по существу определяются теми предположениями, которые принимаются относительно информированности игроков о реализующихся в текущий момент времени управлениях партнера. Будем предполагать, что второй игрок в текущий момент времени t знает кроме реализовавшейся позиции также реализацию управления первого игрока и, следовательно, имеет возможность формировать свое управление в классе контрстратегий ([1], с. 355, [2]). Первый игрок знает только реализовавшуюся позицию и формирует свое управление в классе чистых стратегий. Как отмечалось в начале п. 1, результаты данной работы остаются верными и в этом случае; необходимо лишь в соответствующих выкладках заменить чистую стратегию второго игрока на контрстратегию. Таким образом, в текущий момент времени t первый игрок выбирает управление-силу $F_1(t, \xi[t], \xi^\cdot[t], \varepsilon)$, а второй игрок — управление-силу $F_2(t, \xi[t], \xi^\cdot[t], F_1[t], \varepsilon)$ и управление-угол $\varphi(t, \xi[t], \xi^\cdot[t], F_1[t], \varepsilon)$, на который осуществляется поворот силы F_1 (условимся, что положительному углу соответствует поворот против часовой стрелки). Цель i-го игрока — привести материальную точку как можно ближе к точке $a^{(i)}$.

Полагая в системе (4.1) $y_1 = \xi_1$, $y_2 = \xi_2$, $y_3 = \xi_1^\cdot$, $y_4 = \xi_2^\cdot$ и производя замену переменных $x_1 = y_1 + (\vartheta - t)\,y_3$, $x_2 = y_2 + (\vartheta - t)\,y_4$, $x_3 = y_3$, $x_4 = y_4$, получим систему, два первых уравнения которой имеют вид

(4.4) $$z_1^\cdot = (\vartheta - t)\,(F_{11}\cos\varphi - F_{12}\sin\varphi + F_{21})$$
$$z_2^\cdot = (\vartheta - t)\,(F_{11}\sin\varphi + F_{12}\cos\varphi + F_{22})$$

Показатели (4.3) в переменных z_1, z_2 примут вид

(4.5) $$\sigma_i(z[\vartheta]) = \|z[\vartheta] - a^{(i)}\|, \quad z = \begin{Vmatrix}z_1\\ z_2\end{Vmatrix}, \quad i = 1, 2$$

Поскольку показатели (4.5) определяются значениями лишь координат z_1 и z_2, а правые части системы (4.4) не зависят от других координат, то можно заключить, что рассматриваемую дифференциальную игру достаточно исследовать лишь для укороченной системы (4.4) с показателями (4.5). При этом начальные условия для системы (4.4) будут $z_1[t_*] = z_{*1} = \xi_{*1} - (\vartheta - t_*)\ \xi_{*1}^{\cdot}$, $z_2[t_*] = z_{*2} = \xi_{*2} - (\vartheta - t_*)\ \xi_{*2}^{\cdot}$.

Считаем выполненными условия А, Б.

Функция цены $\gamma_2(t, z)$ антагонистической игры Γ_2, в которой действия игроков формализуются в классах {чистые стратегии первого игрока — контрстратегии второго игрока}, а динамика описывается системой (4.4), а также функции $u^{(2)\circ}(t, z, \varepsilon) = F_1^{(2)\circ}(t, z, \varepsilon)$ и $v^{(2)\circ}(t, z, u, \varepsilon) = \{F_2^{(2)\circ}(t, z, F_1, \varepsilon), \varphi^{(2)\circ}(t, z, F_1, \varepsilon)$, являющиеся аналогами функций (1.4), будут следующими:

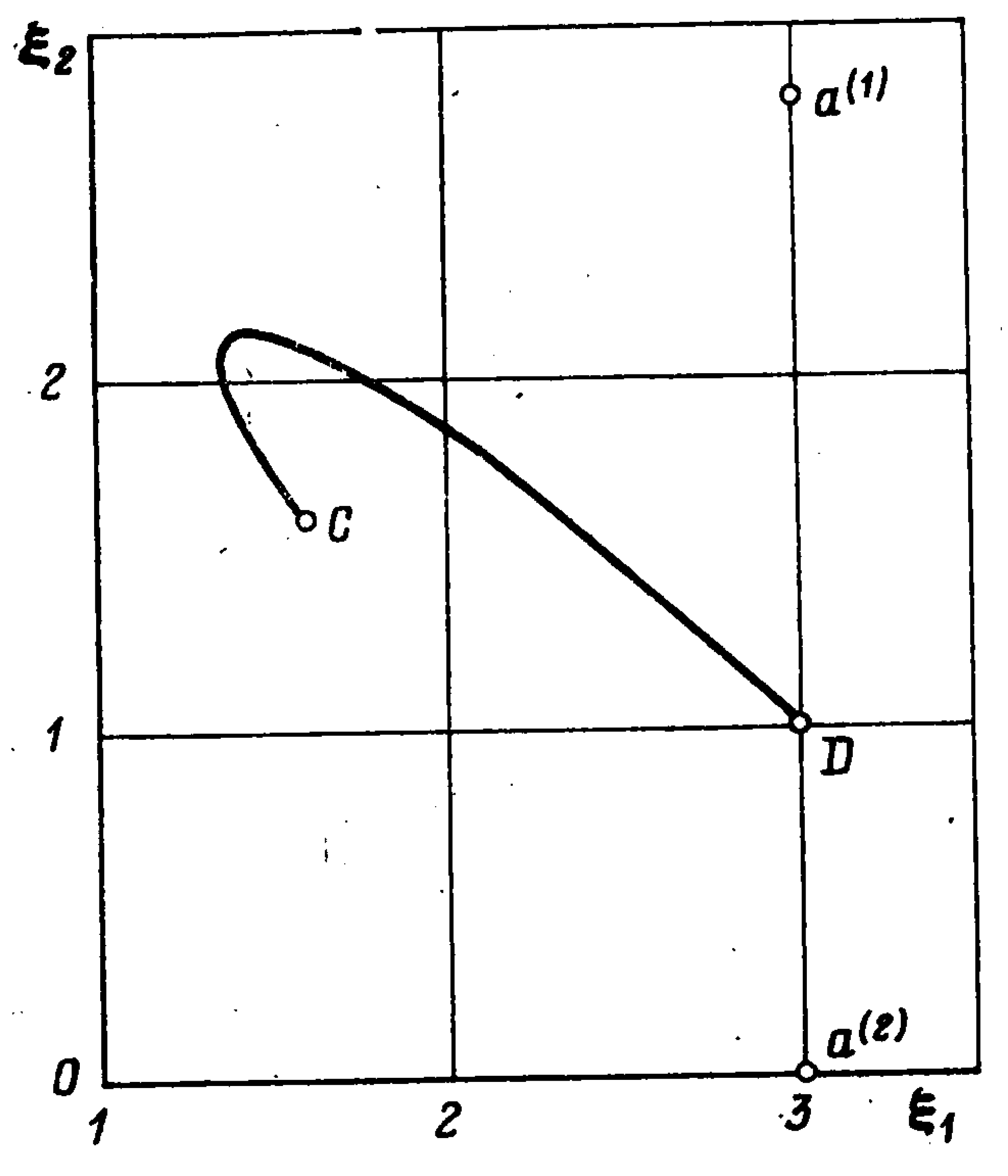
$$\gamma_2(t, z) = \max\left\{\| z - a^{(2)} \| - \frac{1}{2}(1 - \cos\varphi_0)(\vartheta - t)^2, 0\right\} \tag{4.6}$$

$$F_1^{(2)\circ} = \frac{z - a^{(2)}}{\| z - a^{(2)} \|}, \quad F_2^{(2)\circ} = -\frac{z - a^{(2)}}{\| z - a^{(2)} \|} \tag{4.7}$$

$$\varphi^{(2)\circ} = \begin{cases} \psi, & 0 \leqslant \psi \leqslant \varphi_0 \\ -\varphi_0, & \varphi_0 \leqslant \psi \leqslant \pi \\ \varphi_0, & \pi \leqslant \psi \leqslant 2\pi - \varphi_0 \\ 2\pi - \psi, & 2\pi - \varphi_0 \leqslant \psi < 2\pi \end{cases}$$

где $\psi \in [0, 2\pi)$ — угол между векторами $a^{(2)} - z$ и F_1, отсчитываемый от вектора $a^{(2)} - z$ в положительном направлении.

Зададим начальные условия $t_* = 0$, $\xi_{*1} = 1{,}6$, $\xi_{*1}^{\cdot} = -0{,}4$, $\xi_{*2} = 1{,}6$, $\xi_{*2}^{\cdot} = 0{,}6$ и следующие значения параметров $\vartheta = 4$, $a_1^{(1)} = 3$, $a_2^{(1)} = 2{,}8$, $a_1^{(2)} = 3$, $a_2^{(2)} = 0$. Тогда имеем $z_{*1} = 0$, $z_{*2} = 4$. Пусть $\varphi_0 = \pi/3$.



Рассмотрим следующую вспомогательную задачу: найти измеримую вектор функцию $p(t)$, $t_* \leqslant t \leqslant \vartheta$, доставляющую минимум величине $\| z(\vartheta) - a^{(1)} \|$ при условии

$$\| z(\vartheta) - a^{(2)} \| \leqslant \gamma_2(t, z(t)),\ t \in [t_*, \vartheta]$$

где $z(\cdot)$ — траектория системы

$$z^{\cdot}(t) = p(t), \| p(t) \| \leqslant 2(\vartheta - t) \tag{4.8}$$

удовлетворяющая начальному условию $z[t_*] = z_*$.

При заданных начальных условиях и значениях параметров одно из решений $p^*(t)$, $0 \leqslant t \leqslant 4$ поставленной вспомогательной задачи, которое порождает траекторию

$$z_1^*(t) = 3 - \frac{3}{16}(4 - t)^2 \tag{4.9}$$

$$z_2^*(t) = \left(\frac{7}{256}(4 - t)^4 + \frac{1}{2}(4 - t)^2 + 1\right)^{1/2}$$

определяется непрерывными функциями

$$p_1^*(t) = \frac{dz_1^*(t)}{dt}, \quad p_2^*(t) = \frac{dz_2^*(t)}{dt} \tag{4.10}$$

Заметим, что вдоль траектории $z^*(\cdot)$ (4.9) функция (4.6) остается постоянной.

Можно убедиться в том, что любая пара непрерывных функций $(u^*(\cdot), v^*(\cdot))$, удовлетворяющая условиям

$$u^*(t) + v^*(t) = p^*(t), \ \| u^*(t) \| \leqslant (4 - t) \tag{4.11}$$

$$\| v^*(t) \| \leqslant (4 - t), \ \forall t \in [0, 4]$$

доставляет решение задачи 3 для системы (4.4) с показателями (4.5) при заданных начальных условиях и значениях параметров. Условиям (4.11) удовлетворяют, в частности, функции $u^*(t) = v^*(t) = {}^1/_2\, p^*(t)$.

Зафиксируем некоторую пару $(u^*(\cdot), v^*(\cdot))$, являющуюся решением задачи 3. Поскольку имеет место случай б) п. 2, то, согласно теореме 2, стратегии U^*, V^* (2.1) — (2.3), построенные на основе пары $(u^*(\cdot), v^*(\cdot))$ и траектории $z^*(\cdot)$ (4, 9), являются оптимальными в иерархической дифференциальной игре с доброжелательным вторым игроком.

На фигуре изображено в плоскости (ξ_1, ξ_2) движение $z^*[\cdot]$ (4.9), которое является единственным движением пучка, порожденного оптимальными стратегиями U^* и V^*. Движение начинается при $t_* = 0$ в точке C (1,6; 1, 6) и заканчивается в момент $\vartheta = 4$ в точке D (3,0; 1,0). Получаются следующие значения показателей игроков: $\sigma_1(z^*[4]) = 1{,}8$, $\sigma_2(z^*[4]) = 1{,}0$.

Без подробного обсуждения заметим, что для любого сколь угодно малого $\zeta > 0$ можно указать стратегию первого игрока U_ζ, которая гарантирует ему в рассматриваемой иерархической игре без предположения о доброжелательности второго игрока результат, отличающийся от приведенного выше на величину, не превосходящую ζ.

ЛИТЕРАТУРА

1. *Красовский Н. Н., Субботин А. И.* Позиционные дифференциальные игры. М.: Наука, 1974. 456 с.
2. *Красовский Н. Н.* Дифференциальные игры. Аппроксимационные и формальные модели.— Матем. сб., 1978, т. 107, № 4, с. 541—571.
3. *Экланд И.* Элементы математической экономики. М.: Мир, 1983. 248 с.
4. *Von Stackelberg H.* The theory of the market economy. L.: Hodge, 1952. 328 p.
5. *Гермейер Ю. Б.* Об играх двух лиц с фиксированной последовательностью ходов.— Докл. АН СССР, 1971, т. 198, № 5, с. 1001—1004.
6. *Кононенко А. Ф.* О многошаговых конфликтах с обменом информацией.— Ж. вычисл. матем. и матем. физики, 1977, т. 17, № 4, с. 922—931.
7. *Cruz J. B., Jr.* Leader-follower strategies for multilevel systems.— IEEE Trans. Automatic Control, 1978, v. 23, No. 2, p. 244—255.
8. *Basar T., Selbuz H.* Closed-loop Stackelberg strategies with applications in the optimal control of multilevel systems.— IEEE Trans. Automatic Control, 1979, v. 24, No. 2, p. 166—179.
9. *Kleimenov A. F.* Equilibrium coalitional mixed strategies in differential games of m players.— Probl. Control and Inform. Theory, 1982, v. 11, No. 2, p. 85—95.
10. *Клейменов А. Ф.* Равновесные коалиционные контрстратегии в дифференциальных играх многих лиц.— ПММ, 1982, т. 46, вып. 5, с. 714—721.
11. *Kleimenov A. F.* Optimal strategies in a hierarchical differential game.— Probl. Control and Inform. Theory, 1983, v. 12, No. 6, p. 369—377.

Свердловск

Поступила в редакцию
1.VIII.1983