ТРУДЫ ИНСТИТУТА МАТЕМАТИКИ И МЕХАНИКИ УрО РАН

Том 26 № 1 2020

УДК 517.977

ПРИНЯТИЕ РЕШЕНИЙ В ОДНОЙ ГИБРИДНОЙ ЗАДАЧЕ ДИНАМИЧЕСКОГО УПРАВЛЕНИЯ С ТРЕМЯ УЧАСТНИКАМИ

А. Ф. Клейменов

Уравнения движения управляемой системы в рассматриваемой двухшаговой задаче на фиксированном промежутке времени содержат управления либо первого игрока, либо первого и второго игроков, либо первого и третьего игроков, либо всех игроков одновременно. На первом шаге (этапе) управляемого процесса (от начального момента до некоторого заданного момента) на систему действует управление только первого игрока, который решает задачу оптимального управления с заданным терминальным функционалом. В начале второго шага (этапа) процесса первый игрок решает, будут ли участвовать в процессе управления на оставшемся промежутке времени другие игроки или нет. Если да, то участвующие игроки разыгрывают неантагонистическую дифференциальную игру с заданными терминальными функционалами игроков, причем это может быть игра двух или трех лиц. В игре в качестве решения принимается равновесие по Нэшу, неулучшаемое по Парето. Если нет, то первый игрок продолжает решать задачу оптимального управления до окончания процесса.

Ключевые слова: задача оптимального управления, неантагонистическая позиционная дифференциальная игра, терминальные показатели качества, нэшевское равновесие.

A. F. Kleimenov. Decision making in a hybrid two-step problem of dynamic control with three participants.

The equations of motion of a control system in a two-step problem on a fixed time interval contain the controls of either the first player, or the first and second players, or the first and third players, or all players simultaneously. At the first step (stage) of the control process (from the initial time up to a certain predefined moment), the system is controlled only by the first player, who solves an optimal control problem with a given terminal functional. At the beginning of the second step (stage) of the process, the first player decides whether the other players will participate in the control process for the remaining time period. If yes, then the participants play a nonantagonistic differential game with given terminal functionals, and it can be a game of two or three persons. A Pareto-optimal Nash equilibrium is taken as a solution in this game. If no, then the first player continues to solve the optimal control problem until the end of the process.

Keywords: optimal control problem, nonantagonistic positional differential game, terminal payoff functionals, Nash equilibrium.

MSC: 20D10, 20D25

DOI: 10.21538/0134-4889-2020-26-1-131-140

Введение

Работа посвящена приложению теории позиционных дифференциальных игр [1; 2] к анализу одного класса гибридных неантагонистических игр [3–5] и является продолжением статьи [6], в которой рассматривалась гибридная задача управления с двумя участниками.

Управляемая система описывается обыкновенными дифференциальными уравнениями на заданном отрезке времени. Максимальное число участников управляемого процесса (игроков) равно трем (первый игрок $P1$, второй игрок $P2$ и третий игрок $P3$). Игроки действуют в классе позиционных стратегий; движения, порожденные этими стратегиями, определяются аналогично [7; 8].

На первом шаге (этапе) процесса (от начального момента до некоторого заданного момента) правая часть уравнений движения содержит управляющее воздействие только первого игрока, который решает задачу оптимального управления на отрезке с заданным терминальным функционалом выигрыша.

В начале второго шага (этапа) первый игрок решает, будет ли кто из остальных игроков (второй и третий игроки с заданными терминальными функционалами выигрыша) также участвовать в управляемом процессе на оставшемся промежутке времени или нет. Если да, то участвующие игроки разыгрывают неантагонистическую дифференциальную игру двух или трех лиц, в которой в качестве решения берется наилучшее по Парето нэшевское решение. Если нет, то первый игрок по-прежнему решает задачу оптимального управления вплоть до окончания процесса.

Задача состоит в нахождении наилучшей стратегии первого игрока, а также стратегий второго и третьего игроков (в случае их участия в управляемом процессе).

1. Постановка задачи

На заданном отрезке времени $[t_0, \vartheta]$ рассматривается следующая двухшаговая задача управления. Управляемая система описывается обыкновенными дифференциальными уравнениями. Максимальное число участников управляемого процесса (игроков) равно трем (игрок $P1$, игрок $P2$ и игрок $P3$). Игроки действуют в классе позиционных стратегий $U = \{u(t,x,\varepsilon), \beta_1(\varepsilon)\}$, $V = \{v(t,x,\varepsilon), \beta_2(\varepsilon)\}$ и $W = \{w(t,x,\varepsilon), \beta_3(\varepsilon)\}$ соответственно (см. подробно [7, с. 16]). Здесь первая компонента является произвольной функцией позиции (t,x) и положительного параметра точности ε, принимающая значения в ограничивающем множестве. Функция

$$\beta_i \colon (0,\infty) \to (0,\infty)$$

непрерывная монотонная и удовлетворяет условию $\beta_i(\varepsilon) \to 0$ при $\varepsilon \to 0$. Для фиксированного ε величина $\beta_i(\varepsilon)$ является верхней границей шага разбиения отрезка $[t_0, \vartheta]$, которое игрок Pi применяет при формировании пошаговых движений.

Движения (пошаговые и предельные), порожденные этими стратегиями, определяются аналогично [1; 2; 7].

На первом шаге (этапе) процесса от начального момента t_0 до заданного момента T, $t_0 < T < \vartheta$, правая часть уравнений движения содержит управляющее воздействие u только игрока $P1$, который решает следующую задачу оптимального управления Γ^1 с заданным терминальным функционалом выигрыша I_1:

$$\dot{x} = f^{[0]}(t,x,u), \quad x \in \mathbb{R}^n, \quad u \in P^0 \subset \mathbb{R}^p, \quad t_0 \le t \le T, \quad x(t_0) = x_0, \quad I_1 = \sigma_1(x(T)). \tag{1.1}$$

Решение задачи оптимального управления Γ^1 (1.1) обозначим через $u = u^\circ(t)$; оно порождает траекторию $x = x^\circ(t),\ t_0 \le t \le T$.

В начале второго шага (этапа), т. е. в заданный момент времени T, игрок $P1$ должен решить, кто из остальных игроков также будет участвовать в управляемом процессе на оставшемся промежутке времени $[T, \vartheta]$. Всего у игрока $P1$ имеется четыре варианта выбора: 1°) участвует игрок $P2$; 2°) участвует игрок $P3$; 3°) участвуют оба игрока $P2$ и $P3$; 4°) никто из игроков $P2$, $P3$ не участвует.

Примем, что игрок $P2$ располагает ресурсом управления v, $v(t) \in Q^0 \subset \mathbb{R}^q$, и имеет заданный терминальный функционал выигрыша I_2, а игрок $P3$ располагает ресурсом управления w, $w(t) \in S^0 \subset \mathbb{R}^s$, и имеет заданный терминальный функционал выигрыша I_3. Предполагаем также, что выигрыши игроков являются трансферабельными [3] и что за участие в управляемом процессе на отрезке $[T, \vartheta]$ вновь вошедшие игроки выплачивают игроку $P1$ платеж в размере $L > 0$ единиц.

Если игрок $P1$ выбрал вариант 4°), т. е. если никто из игроков $P2$, $P3$ не будет участвовать в процессе, то на отрезке $[T, \vartheta]$ игрок $P1$ продолжает решать задачу оптимального управления Γ^1:

$$\dot{x} = f^{[0]}(t,x,u), \quad u \in P^0, \quad T \le t \le \vartheta, \quad x(T) = x^\circ(T), \quad I_1 = \sigma_1(x(\vartheta)). \tag{1.2}$$

Выигрыш игрока $P1$, получаемый в конечной точке $x^{(0)}(\vartheta)$ оптимальной траектории $x^{(0)}(\cdot)$ в задаче Γ^1 (1.2), обозначим через $I_1^{(0)}$.

Если игрок $P1$ выбрал вариант 1°), то игроки $P1$ и $P2$ на отрезке $[T,\vartheta]$ разыгрывают следующую неантагонистическую позиционную дифференциальную игру двух лиц Γ^{12}:

$$\dot{x} = f^{[1]}(t,x,u,v),\ u\in P,\ v\in Q^0,\ T\le t\le\vartheta,\ x(T)=x^\circ(T),\ I_1=\sigma_1(x(\vartheta)),\ I_2=\sigma_2(x(\vartheta)), \quad (1.3)$$

причем в этой игре игрок $P1$ распоряжается выбором управления u уже из другого множества $u\in P$, а игрок $P2$ распоряжается выбором управления $v\in Q^0$. При этом множество $P\subset P^0$ выбирается таким, что два множества — вектограмма уравнений динамики дифференциальной игры Γ^{12} (где $u\in P,\ v\in Q^0$) и вектограмма уравнений динамики задачи оптимального управления Γ^1 (где $u\in P^0$) — совпадают. Полагаем, что игроки выбирают одно из $P(NE)$-решений [7] игры Γ^{12} (1.3), порождающее траекторию $x^{(1)}(\cdot)$. Обозначим значения выигрышей на выбранном $P(NE)$-решении через $I_1^{(1)}$ — для игрока $P1$ и через $I_2^{(1)}$ — для игрока $P2$.

Аналогично если игрок $P1$ выбрал вариант 2°), то $P1$ вместе с $P3$ на отрезке $[T,\vartheta]$ разыгрывает неантагонистическую позиционную дифференциальную игру двух лиц Γ^{13}:

$$\dot{x} = f^{[2]}(t,x,u,w),\ u\in P,\ w\in S^0,\ T\le t\le\vartheta,\ x(T)=x^\circ(T),\ I_1=\sigma_1(x(\vartheta)),\ I_3=\sigma_3(x(\vartheta)). \quad (1.4)$$

При этом множество $P\subset P^0$ выбирается таким, что два множества — вектограмма уравнений динамики дифференциальной игры Γ^{13} (где $u\in P,\ w\in S^0$) и вектограмма уравнений динамики задачи оптимального управления Γ^1 (где $u\in P^0$) — совпадают. Значения выигрышей на траектории $x^{(2)}(\cdot)$, порожденной выбранным $P(NE)$-решением игры Γ^{13} (1.4), обозначим через $I_1^{(2)}$ — для игрока $P1$ и через $I_3^{(2)}$ — для игрока $P3$.

Предположение 1. Игрок $P1$ решает, что игрок Pi, $i\in\{2,3\}$, участвует в управляемом процессе на промежутке времени $[T,\vartheta]$, если имеют место неравенства

$$I_1^{(i-1)} + L > I_1^{(0)}, \quad (1.5)$$

$$I_i^{(i-1)} - L > I_i^{(0)}, \quad (1.6)$$

где $I_i^{(0)}$ — значение функционала I_i в точке $x^0(\vartheta)$.

Неравенство (1.5) означает, что при участии игрока Pi в управляемом процессе игрок $P1$ получает выигрыш (с учетом полученного платежа в размере L) больший, чем в случае, если это участие не состоится. Неравенство (1.6) означает, что игроку Pi тоже выгодно участвовать в управляемом процессе, даже заплатив за участие игроку $P1$ платеж в размере L.

Итак, определены задачи оптимального управления Γ^1 (1.1), (1.2) и неантагонистические позиционные дифференциальные игры двух лиц Γ^{12} (1.3) и Γ^{13} (1.4). Предполагаем, что функции $f^{[0]}(\cdot,\cdot,\cdot)$ (1.1), (1.2), $f^{[1]}(\cdot,\cdot,\cdot,\cdot)$ (1.3), $f^{[2]}(\cdot,\cdot,\cdot,\cdot)$(1.4) непрерывны по совокупности переменных, липшицевы по x, удовлетворяют условию подлинейного роста по x, а также условию седловой точки в маленькой игре [1], а функции $\sigma_i(\cdot)$, задающие терминальные выигрыши игроков, непрерывны.

З а д а ч а 1.i, $i\in\{2,3\}$. Найти $P(NE)$-решение игры Γ^{1i} (1.3),(1.4) и число $L>0$ такие, что для порожденной этим решением траектории выполняются неравенства (1.5) и (1.6).

В общем случае задачи 1.i решений не имеют.

Наконец, если игрок $P1$ выбрал вариант 3°), то игроки $P1$, $P2$ и $P3$ на отрезке $[T,\vartheta]$ разыгрывают следующую неантагонистическую позиционную дифференциальную игру трех лиц Γ^{123}:

$$\dot{x} = f^{[3]}(t,x,u,v,w),\quad u\in P,\quad v\in Q\subset Q^0,\quad w\in S\subset S^0,\quad T\le t\le\vartheta,\quad x(T)=x^\circ(T),$$

$$I_1=\sigma_1(x(T)),\quad I_2=\sigma_2(x(T)),\quad I_3=\sigma_3(x(T)); \quad (1.7)$$

причем в игре Γ^{123} игрок $P1$ распоряжается выбором управления из множества $u \in P$, игрок $P2$ — выбором управления $v \in Q$, а игрок $P3$ — выбором управления $w \in S$. При этом множество $P \subset P^0$ выбирается так, что два множества — вектограмма уравнений динамики дифференциальной игры Γ^{123} (где $u \in P,\ v \in Q,\ w \in S$) и вектограмма уравнений динамики задачи оптимального управления Γ^1 (где $u \in P^0$) — совпадают. В игре Γ^{123} (1.7) игроки выбирают одно из $P(NE)$-решений игры; выигрыши, получаемые в конечной точке $x^{(3)}(\vartheta)$ траектории $x^{(3)}(\cdot)$, порожденной выбранным $P(NE)$-решением, обозначим через $I_1^{(3)}$ — для игрока $P1$, через $I_2^{(3)}$ — для игрока $P2$ и через $I_3^{(3)}$ — для игрока $P3$.

Предположение 2. Игрок $P1$ решает, что игроки $P2$ и $P3$ участвуют в управляемом процессе на промежутке времени $[T, \vartheta]$, если имеют место следующие неравенства:

$$I_1^{(3)} + L > I_1^{(0)}, \tag{1.8}$$

$$I_2^{(3)} + I_3^{(3)} - L > I_2^{(0)} + I_3^{(0)}. \tag{1.9}$$

Примем, что функция $f^{[3]}(\cdot,\cdot,\cdot,\cdot,\cdot)$ (1.7) непрерывна по совокупности переменных, липшицева по x, удовлетворяют условию подлинейного роста по x, а также условию седловой точки в соответствующих маленьких играх [1].

З а д а ч а 2. Найти $P(NE)$-решение игры трех лиц Γ^{123} (1.7) и число $L > 0$ такие, что для порожденной этим решением траектории выполняются неравенства (1.8) и (1.9).

В общем случае задача 2 также решений не имеет.

З а м е ч а н и е. Можно дать следующую интерпретацию действий игроков на втором шаге (см. также работу [4]). Игрок $P1$ продает часть своих активов за цену L либо игроку $P2$ (вариант 1°)), либо игроку $P3$ (вариант 2°)), либо игрокам $P2$ и $P3$ (вариант 3°)), если такая продажа является взаимовыгодной (оба неравенства хотя бы в одной из систем — в системе (1.5), (1.6) или в системе (1.8), (1.9) — выполнены). Если продажа невыгодна (по крайней мере, одно из неравенств как в системе (1.5), (1.6), так и в системе (1.8), (1.9) не выполнено), то она не состоится (вариант 4°)).

2. Вспомогательные антагонистические позиционные дифференциальные игры

Выше было отмечено, что в возникающих в вариантах 1°)–3°) неантагонистических дифференциальных играх двух или трех лиц используются $P(NE)$-решения этих игр. $P(NE)$-решения в играх двух лиц введены в [7, с. 28–29], где они были обозначены как P^*-решения. Там же выявлена структура NE-решений и $P(NE)$-решений в играх двух лиц. Обобщение понятия $P(NE)$-решения на игру трех лиц можно найти в недавней работе автора (Альтруистический и агрессивный типы поведения в неантагонистической позиционной дифференциальной игре трех лиц // Тр. Ин-та математики и механики УрО РАН. 2019. Т. 25, № 3. С. 108–117.)

Введем вспомогательные антагонистические позиционные дифференциальные игры Γ_1^{12}, Γ_2^{12}, Γ_1^{13}, Γ_3^{13}, Γ_1^{123}, Γ_2^{123}, Γ_3^{123} . Динамика игр Γ_1^{12} и Γ_2^{12} описывается уравнением (1.3). В игре Γ_i^{12}, $i = 1, 2$, игрок i максимизирует свой функционал I_i , а игрок $(3-i)$ противодействует ему.

Аналогично динамика игр Γ_1^{13} и Γ_3^{13} описывается уравнением (1.4). В игре Γ_i^{13}, $i = 1, 3$, игрок i максимизирует свой функционал I_i , а игрок $(4-i)$ противодействует ему.

Наконец, динамика игр Γ_1^{123}, Γ_2^{123} и Γ_3^{123} описывается уравнением (1.7). В игре Γ_i^{123}, $i = 1, 2, 3$, игрок i максимизирует свой функционал I_i, а два других игрока совместно противодействуют ему. Из [1; 2] следует, что при сделанных предположениях относительно правых частей уравнений движений и функционалов выигрыша игроков каждая из упомянутых семи антагонистических игр имеет универсальную седловую точку и непрерывную функцию цены $\gamma_i^{12}(t,x)$, $\gamma_j^{13}(t,x)$, $\gamma_k^{123}(t,x)$, $i=1,2$, $j=1,3$, $k=1,2,3$. Свойство универсальности стратегий означает, что они являются оптимальными не только для фиксированной начальной позиции (t_0,x_0), но и для любой позиции (t_*,x_*), рассматриваемой в качестве начальной.

Заметим, что величина $\gamma_i(t,x)$ представляет собою гарантированный выигрыш игрока i в позиции (t,x) игры. Для каждой NE- и $P(NE)$-траектории $x^*(t)$ имеет место следующее свойство [7, с. 26]:

С в о й с т в о A. Точка $t=\vartheta$ является точкой максимума функции гарантированного выигрыша игрока i, вычисленной вдоль этой траектории, т. е.

$$\max_{t\in[t_0,\vartheta]}\gamma_i(t,x^*(t))=\gamma_i(\vartheta,x^*(\vartheta)),\quad i=1,2,3.$$

В следующем разделе рассмотрим один класс задач, в котором задачи 1 и 2 имеют решения.

3. Задача управления на плоскости с динамикой простых движений

Рассмотрим следующую двухшаговую задачу управления на плоскости с динамикой простых движений и с тремя участниками.

На первом шаге решается задача оптимального управления Γ^1 для игрока $P1$:

$$\dot x=u,\quad x,u\in\mathbb{R}^2,\quad \| u \|\le 2\mu,\quad t_0\le t\le T,\quad x(t_0)=x_0,\quad I_1=\sigma_1(x(T)). \tag{3.1}$$

На втором шаге, т. е. на отрезке $[T,\vartheta]$, в зависимости от варианта выбора игроком решается одна из следующих четырех задач:

4°) задача оптимального управления Γ^1 для игрока $P1$:

$$\dot x=u,\quad \| u \|\le 2\mu,\quad T\le t\le\vartheta,\quad x(T)=x^\circ(T),\quad I_1=\sigma_1(x(\vartheta)); \tag{3.2}$$

1°) неантагонистическая дифференциальная игра двух лиц Γ^{12}:

$$\dot x=u+v,\ \| u \|\le\mu,\ \| v \|\le\mu,\ T\le t\le\vartheta,\ x(T)=x^\circ(T),\ I_1=\sigma_1(x(T)),\ I_2=\sigma_2(x(T)); \tag{3.3}$$

2°) неантагонистическая дифференциальная игра двух лиц Γ^{13}:

$$\dot x=u+w,\ \| u \|\le\mu,\ \| w \|\le\mu,\ T\le t\le\vartheta,\ x(T)=x^\circ(T),\ I_1=\sigma_1(x(T)),\ I_3=\sigma_3(x(T)); \tag{3.4}$$

3°) неантагонистическая дифференциальная игра трех лиц Γ^{123}:

$$\dot x=u+v+w,\quad \| u \|\le\mu,\quad \| v \|\le 0.5\mu,\quad \| w \|\le 0.5\mu,\quad T\le t\le\vartheta,\quad x(T)=x^\circ(T), \tag{3.5}$$

$$I_1=\sigma_1(x(T)),\quad I_2=\sigma_2(x(T)),\quad I_3=\sigma_3(x(T)).$$

Зададим следующие функционалы выигрыша игроков $P1$, $P2$ и $P3$:

$$I_i=\sigma_i(x(\vartheta))=M-\|x(\vartheta)-a^{(i)}\|,\quad i=1,2,3,$$

т. е. игрок Pi стремится привести точку $x(\vartheta)$ как можно ближе к целевой точке $a^{(i)}$.

Очевидно, что вектограммы уравнений динамики дифференциальных игр Γ^{12} (3.3), Γ^{13} (3.4) и Γ^{123} (3.5) совпадают друг с другом, а также с вектограммой уравнений динамики задачи оптимального управления Γ^1 (3.2).

Зададим следующие начальные условия и значения параметров:

$$t_0=0,\quad x_0=(-2,0),\quad \vartheta=5.0,\quad T=1.0,\quad \mu=0.5,\quad a^{(1)}=(12,0),$$

$$a^{(2)}=(-8.5,8.5),\quad a^{(3)}=(8.5,8.5),\quad M=20.$$

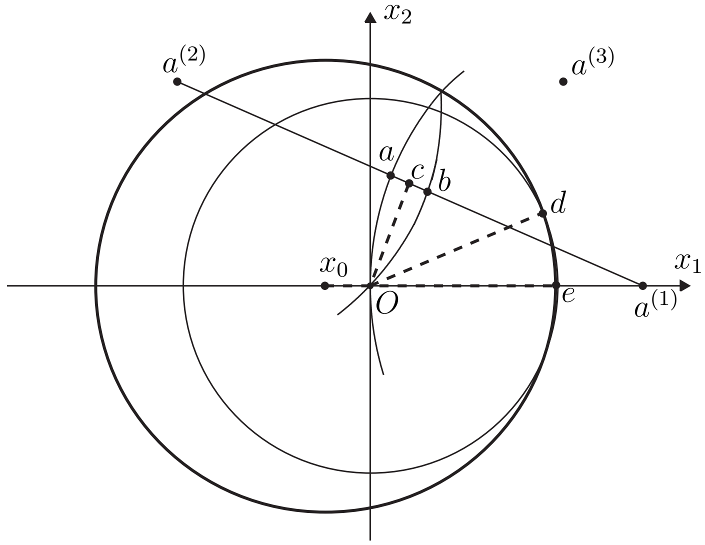

Рис. 1. Опции 4°), 1°), 2°).

На рис. 1 и рис. 2 большой круг с центром в точке x_0 есть множество достижимости $G(\vartheta, t_0, x_0)$, построенное для момента ϑ. Очевидно, что рассматриваемое на первом шаге решение задачи оптимального управления Γ^1 (3.1) может быть представлено как $u = u^\circ(t) = (2, 0),\ 0 \le t \le 1$. Оптимальной траекторией будет $x = x^\circ(t) = (-2 + 2t, 0),\ 0 \le t \le 1$; на рис. 1 и рис. 2 она изображена отрезком x_0O.

Переходим ко второму шагу. Прежде всего отметим, что на рис. 1 и рис. 2 малый круг с центром в точке O изображает множество достижимости $G(\vartheta, T, x^\circ(T))$ из состояния O в момент $T = 1$, построенное для момента $\vartheta = 5$. На этом шаге решение задачи оптимального управления Γ^1 (3.2) (опция 4°)) следующее: $u = u^0(t) = (2, 0),\ 1 \le t \le 5$. Оптимальной траекторией будет $x = x^0(t) = (2t, 0),\ 1 \le t \le 5$; на рис. 1 она изображена пунктирной линией Oe. В точке e выигрыши игроков составят величины $I_1^{(0)} = 16.0,\ I_2^{(0)} = 1.4,\ I_3^{(0)} = 11.6$.

Переходем к нахождению $P(NE)$-решений игр (3.3)–(3.5); сначала выпишем формулы для функций цены $\gamma_i^{12}(t,x),\ \gamma_j^{13}(t,x), \gamma_k^{123}(t,x)$, $i = 1, 2$, $j = 1, 3$, $k = 1, 2, 3$:

$$\gamma_i^{12}(t,x) = 20 - \|x - a^{(i)}\|, \quad i = 1, 2, \tag{3.6}$$

$$\gamma_j^{13}(t,x) = 20 - \|x - a^{(j)}\|, \quad j = 1, 3,$$

$$\gamma_1^{123}(t,x) = 20 - \|x - a^{(1)}\|,$$

$$\gamma_k^{123}(t,x) = 20 - \|x - a^{(k)}\| - (\vartheta - t), \quad k = 2, 3.$$

Нетрудно проверить, что в игре двух лиц Γ^{12} (3.3) (опция 1°)) множество концов траекторий, порожденных $P(NE)$-решений игры, составляет отрезок ab, где a и b – точки пересечения отрезка $a^{(1)}a^{(2)}$ с дугами окружностей $\|x - a^{(1)}\| = \|Oa^{(1)}\|$ и $\|x - a^{(2)}\| = \|Oa^{(2)}\|$ соответственно. Эти окружности являются сечениями поверхностей уровня функций $\gamma_1^{12}(t,x)$ и $\gamma_2^{12}(t,x)$ (3.6) плоскостями $t = \text{const}$.

Выбираем $P(NE)$-решение, приводящее в точку $c = (1.7, 4.25)$ — середину отрезка ab. Убеждаемся, что это решение порождено следующими стратегиями игроков $P1$ и $P2$:

$$U^{*1} = \{u^{*1}(t,x,\varepsilon), \beta_1^{*1}(\varepsilon)\},\ \ V^{*1} = \{v^{*1}(t,x,\varepsilon), \beta_2^{*1}(\varepsilon)\}, \tag{3.7}$$

$$u^{*1}(t,x,\varepsilon) = \begin{cases} (0.372, 0.929), & \left\|x - x^{*1}(t)\right\| < \varepsilon, \\ (0, -1), & \left\|x - x^{*1}(t)\right\| \ge \varepsilon, \end{cases}$$

$$v^{*1}(t,x,\varepsilon) = (0.372, 0.929),$$

где

$$x^{*1}(t) = \big((0.744(t-1), 1.858(t-1)),\ 1 \le t \le 3.288;\ (1.7, 4.25),\ 3.288 < t \le 5\big),$$

$$\beta_i^{*1}(\varepsilon) = 0.1\varepsilon, \quad i = 1, 2.$$

Порожденная выбранным $P(NE)$-решением траектория представляет собой отрезок Oc (на рис. 1 выделен пунктирной линией). Выигрыши игроков $P1$ и $P2$ в точке c составляют величины $I_1^{(1)} = 8.9$, $I_2^{(1)} = 8.9$. Подставляя эти значения в неравенства (1.5),(1.6), имеем $8.9 + L > 16$, $8.9 - L > 1.4$ или

$$7.1 < L < 7.5. \tag{3.8}$$

Таким образом, получаем следующее утверждение.

Утверждение 1. *Пара стратегий игроков* $P1$ *и* $P2$ (U^{*1}, V^{*1}) (3.7) *и число* L (3.8) *доставляют решение задачи* 1.2.

Далее, в игре двух лиц Γ^{13} (3.4) (опция 2°)) можно проверить, что точка $d = (7.4, 3.04)$, являющаяся граничной точкой множества достижимости $G(\vartheta, T, x^{\circ}(T))$, будет концом траектории, порожденной $P(NE)$-решением игры. Убеждаемся, что это решение порождено следующими стратегиями игроков $P1$ и $P3$:

$$U^{*2} = \{u^{*2}(t, x, \varepsilon), \beta_1^{*2}(\varepsilon)\}, \ \ W^{*2} = \{w^{*2}(t, x, \varepsilon), \beta_3^{*2}(\varepsilon)\}, \tag{3.9}$$

$$u^{*2}(t, x, \varepsilon) = \begin{cases} (0.925, 0.380), & \|x - x^{*2}(t)\| < \varepsilon, \\ (0, -1), & \|x - x^{*2}(t)\| \ge \varepsilon, \end{cases}$$

$$w^{*1}(t, x, \varepsilon) = (0.372, 0.929),$$

где $x^{*2}(t) = \big((1.850(t-1),\ 0.760(t-1)),\ 1 \le t \le 5\big)$, $\beta_i^{*2}(\varepsilon) = 0.1\varepsilon$, $i = 1, 2$.

Эта траектория представляет собой отрезок Od (на рис. 1 выделен пунктирной линией). Выигрыши игроков $P1$ и $P3$ в точке d составляют величины $I_1^{(2)} = 14.5$, $I_3^{(2)} = 14.5$. Подставляя эти значения в неравенства (1.5),(1.6), получаем $14.5 + L > 16$, $14.5 - L > 11.6$ или

$$1.5 < L < 2.9. \tag{3.10}$$

Утверждение 2. *Пара стратегий игроков* $P1$ *и* $P3$ (U^{*2}, W^{*2}) (3.9) *и число* L (3.10) *доставляют решение задачи* 1.3.

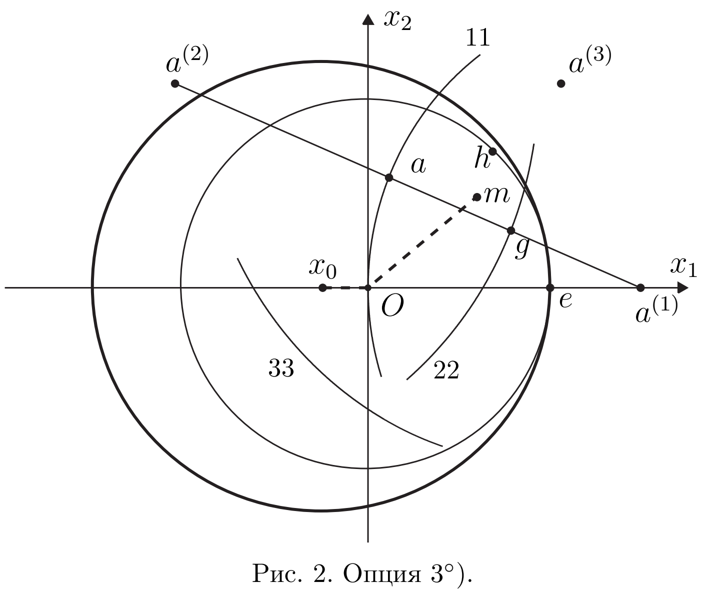

Рис. 2. Опция 3°).

Переходим к игре трех лиц Γ^{123} (3.5)(опция 3°)).

На рис. 2 линии 11, 22 и 33 суть дуги окружностей $\|x - a^{(1)}\| = \|Oa^{(1)}\|$, $\|x - a^{(2)}\| = \|Oa^{(2)}\| + (\vartheta - T)$, $\|x - a^{(3)}\| = \|Oa^{(3)}\| + (\vartheta - T)$ соответственно. Эти окружности являются сечениями поверхностей уровня функций $\gamma_1^{123}(t,x)$ (3.19), $\gamma_2^{123}(t,x)$ и $\gamma_3^{123}(t,x)$ (3.20) плоскостью $t = T$. Согласно свойству A из разд. 2 множество, являющееся пересечением соответствующих кругов, содержит концы траекторий, порожденных $P(NE)$-решениями игры. При этом траектории, заканчивающиеся в точках $g = (6.3, 2.4)$, $a = (0.9, 6.5)$ и $h = (5.7, 5.7)$, порождены $P(NE)$-решениями, наилучшими для игроков $P1$, $P2$ и $P3$ соответственно. Построим точку $m = (4.8, 3.8)$, являющуюся взвешенным средним точек g, a и h (с весами 0.5, 0.25 и 0.25, соответственно). На рис. 2 $P(NE)$-траектория, приводящая в точку m, выделена пунктирной линией Om. Убеждаемся, что эта траектория порождена следующими стратегиями игроков $P1$, $P2$ и $P3$:

$$U^{*3} = \{u^{*3}(t,x,\varepsilon), \beta_1^{*3}(\varepsilon)\}, \;\; V^{*3} = \{v^{*3}(t,x,\varepsilon), \beta_2^{*3}(\varepsilon)\}, \;\; W^{*3} = \{w^{*3}(t,x,\varepsilon), \beta_3^{*3}(\varepsilon)\}, \tag{3.11}$$

$$u^{*3}(t,x,\varepsilon) = \begin{cases} (0.621, 0.784), & \left\|x - x^{*3}(t)\right\| < \varepsilon, \\ (0,-1), & \left\|x - x^{*3}(t)\right\| \geq \varepsilon, \end{cases}$$

$$v^{*3}(t,x,\varepsilon) = (0.311, 0.392),$$

$$w^{*}(t,x,\varepsilon) = (0.311, 0.392),$$

где

$$x^{*3}(t) = \big((1.241(t-1), 1.568(t-1)),\; 1 \leq t \leq 4.061;\; (4.8, 3.8),\; 4.061 < t \leq 5\big),$$

$$\beta_i^{*3}(\varepsilon) = 0.1\varepsilon, \quad i = 1,2,3.$$

Выигрыши игроков $P1$, $P2$ и $P3$ в точке m составляют величины $I_1^{(3)} = 11.9$, $I_2^{(3)} = 5.9$, $I_3^{(3)} = 14.1$. Подставляя эти значения в неравенства (1.8),(1.9), получим $11.9 + L > 16$, $5.9 + 14.1 - L > 1.4 + 11.6$ или

$$4.1 < L < 7 \tag{3.12}$$

Утверждение 3. *Тройка стратегий игроков $P1$, $P2$ и $P3$ (U^{*3}, U^{*3}, W^{*3}) (3.11) и число L (3.12) доставляют решение задачи 2.*

Заключение

В работе рассматривается двухшаговая задача динамического управления на фиксированном отрезке времени $[t_0, \vartheta]$, динамика которой описывается обыкновенными дифференциальными уравнениями и формируется под воздействием управлений трех участников (игроков). У каждого игрока имеется свой терминальный функционал выигрыша.

На первом шаге $[t_0, T]$ в формировании управляющего воздействия участвует только первый игрок, который решает задачу оптимального управления с терминальным функционалом выигрыша I_1.

На оставшемся промежутке времени $[T, \vartheta]$ в соответствии с выбором первого игрока может реализоваться один из следующих четырех вариантов: 1) разыгрывается неантагонистическая дифференциальная игра первого и второго игроков с терминальными функционалами I_1 и I_2; 2) разыгрывается неантагонистическая дифференциальная игра первого и третьего игроков с терминальными функционалами I_1 и I_3; 3) разыгрывается неантагонистическая дифференциальная игра первого, второго и третьего игроков с терминальными функционалами I_1, I_2 и I_3; 4) продолжается решение задачи оптимального управления с терминальным функционалом выигрыша I_1. Предполагается, что в каждом из четырех вариантов вектограмма правой части уравнений движения одна и та же. В вариантах 1)–3) игроки действуют в классе позиционных стратегий. В качестве решений здесь принимаются $P(NE)$-решения игры двух или трех лиц.

Выигрыши игроков считаются трансферабельными. За участие в управляемом процессе на отрезке $[T, \vartheta]$ вновь вошедшие игроки выплачивают первому игроку платеж в размере $L > 0$ единиц. Ставятся задачи 1.i и 2.

Получено решение задач 1.i и 2 в примере с уравнениями простой динамики на плоскости и фунционалами игроков, имеющих смысл минимума расстояния от целевой точки до конечной точки на траектории (утверждения 1–3).

СПИСОК ЛИТЕРАТУРЫ

1. **Красовский Н.Н., Субботин А.И.** Позиционные дифференциальные игры. Москва: Наука, 1974. 456 с.
2. **Красовский Н.Н.** Управление динамической системой. М.: Наука, 1985. 516 с.
3. **Петросян Л.А., Зенкевич Н.А., Шевкопляс Е.В.** Теория игр. СПб.: БХВ-Петербург, 2012. 432 с.
4. **Peter M. Kort, Stefan Wrzaczek.** Optimal firm growth under the threat of entry // Eur. J. Oper. Res. 2015. Vol. 246, no. 1. P. 281–292. doi: 10.1016/j.ejor.2015.04.030 .
5. **Gromov D., Gromova E.** On a class of hybrid differential games // Dyn. Games Appl. 2017. Vol. 7, no. 2. P. 266–288. doi: 10.1007/s13235-016-0185-3 .
6. **Клейменов А.Ф.** Принятие решений в одной гибридной двухшаговой задаче динамического управления // Вестн. Тамбов. ун-та. Сер.: Естественные и технические науки. 2018. Т. 23, №. 123. С. 415–423.
7. **Клейменов А.Ф.** Неантагонистические позиционные дифференциальные игры. Екатеринбург: Наука, 1993. 185 с.
8. **Клейменов А.Ф.** О решениях в неантагонистической позиционной дифференциальной игре // Прикл. математика и механика. 1997. Т. 61, вып. 5. С. 739–746.

Поступила 25.12.2019
После доработки 30.01.2020
Принята к публикации 3.02.2020

Клейменов Анатолий Федорович
д-р физ.-мат. наук, профессор,
ведущий науч. сотрудник
Институт математики и механики им. Н. Н. Красовского УрО РАН;
Уральский федеральный университет
г. Екатеринбург
e-mail: kleimenov@imm.uran.ru

REFERENCES

1. Krasovskii N.N., Subbotin A.I. *Game-theoretical control problems.* N Y: Springer, 1988, 517 p. ISBN: 978-1-4612-8318-8 . Original Russian text published in Krasovskii N.N., Subbotin A.I. *Pozitsionnye differentsial'nye igry.* Moscow: Nauka Publ., 1974, 456 p.
2. Krasovskii N.N. *Upravlenie dinamicheskoi sistemoi* [Control of a dynamical system]. Moscow: Nauka Publ., 1985, 516 p.
3. Petrosyan L.A., Zenkevich N.A., Shevkoplyas E.V. *Teoriya igr* [Game theory]. Saint Petersburg: BKhV-Peterburg Publ., 2012, 432 p. ISBN: 978-5-9775-0484-3 .
4. Kort P.M., Wrzaczek S. Optimal firm growth under the threat of entry. *Eur. J. Oper. Res.*, 2015, vol. 246, no. 1, pp. 281–292. doi: 10.1016/j.ejor.2015.04.030 .
5. Gromov D., Gromova E. On a class of hybrid differential games. *Dyn. Games Appl.*, 2017, vol. 7, no. 2, pp. 266–288. doi: 10.1007/s13235-016-0185-3 .
6. Kleimenov A.F. Decision-making in a hybrid two-step problem of dynamic control. *Tambov University Reports. Series: Natural and Technical Sciences*, 2018, vol. 23, no. 123, pp. 415–423. DOI: 10.20310/1810-0198-2018-23-123-415-423 .
7. Kleimenov A.F. *Neantagonisticheskie pozitsionnye differentsial'nye igry* [Nonantagonistic positional differential games]. Ekaterinburg: Nauka Publ., 1993, 185 p. ISBN: 5-7691-0353-1 .

8. Kleimenov A.F. On solutions in a nonantagonistic positional differential game. *J. Appl. Math. Mech.*, 1997, vol. 61, no. 5, pp. 717–723. doi: 10.1016/S0021-8928(97)00094-4 .

Received December 25, 2019
Revised January 30, 2020
Accepted February 3, 2020

Anatolii Fedorovich Kleimenov, Dr. Phys.-Math. Sci., Krasovskii Institute of Mathematics and Mechanics of the Ural Branch of the Russian Academy of Sciences, Yekaterinburg, 620108 Russia; Ural Federal University, Yekaterinburg, 620083 Russia, e-mail: kleimenov@imm.uran.ru .

Cite this article as: A. F. Kleimenov. Decision making in a hybrid two-step problem of dynamic control with three participants, *Trudy Instituta Matematiki i Mekhaniki URO RAN*, 2020, vol. 26, no. 1, pp. 131–140 .