ВЕСТНИК УДМУРТСКОГО УНИВЕРСИТЕТА

МАТЕМАТИКА 2009. Вып. 3

УДК 517.977.8

© *Д. Р. Кувшинов*

# АЛГОРИТМ ЧИСЛЕННОГО ПОСТРОЕНИЯ РЕШЕНИЙ ПО НЭШУ В ПОЗИЦИОННОЙ ДИФФЕРЕНЦИАЛЬНОЙ ИГРЕ ДВУХ ЛИЦ[1]

Предлагается численный алгоритм построения аппроксимации множества решений Нэша в линейной неантагонистической позиционной дифференциальной игре двух лиц с терминальными цилиндрическими показателями качества и геометрическими ограничениями на управления игроков.

*Ключевые слова*: позиционные дифференциальные игры, равновесное по Нэшу решение, численный алгоритм.

## Введение

Работа посвящена численному алгоритму построения равновесных по Нэшу решений в позиционной дифференциальной игре двух лиц с терминальными показателями качества игроков и геометрическими ограничениями на управления.

Формализация рассматриваемой неантагонистической игры опирается на формализацию и результаты теории антагонистических позиционных дифференциальных игр, построенной Н. Н. Красовским, А. И. Субботиным и их сотрудниками [1; 2], и излагается в работе А. Ф. Клейменова [3]. В [3] задача построения равновесных по Нэшу решений сведена к решению некоторой задачи (оптимального) управления. Этот подход позволяет использовать известные процедуры численного построения стабильных мостов в антагонистических позиционных дифференциальных играх (см., например, работы В. Н. Ушакова, В. С. Пацко и их учеников [4; 5]).

Предлагаемый алгоритм был создан как обобщение алгоритма численного построения решений Штакельберга в дифференциальной позиционной игре, разработанного С. И. Осиповым [6; 7]. Вначале ставилась задача поиска только неулучшаемых решений Нэша, в дальнейшем полученный алгоритм был дополнен для получения всех решений Нэша.

Изложение построено следующим образом. В § 1 даны формализация рассматриваемой игры и определения решений в достаточно общей форме. Далее описан общий подход к построению решений Нэша, основанный на сведении исходной задачи к нестандартной задаче (оптимального) управления, названной здесь «вспомогательной задачей» (§ 2). Собственно алгоритм для линейной игры, приводимой к игре на плоскости, дан в § 3. Результаты численного эксперимента для задачи движения материальной точки в плоскости представлены в § 4.

Отметим, что ограничение размерности фазового вектора связано с уровнем развития алгоритмов вычислительной геометрии и их программных реализаций. В частности, автору неизвестны алгоритмы, строящие результаты теоретико-множественных операций над многогранниками в пространствах размерности больше 3.

Материал, изложенный в статье, был доложен на Всероссийской конференции «Динамические системы, управление и наномеханика» [8].

## § 1. Постановка задачи

Пусть динамика управляемой системы описывается уравнением вида

$$\dot{\mathbf{x}} = f(t, \mathbf{x}, \mathbf{u}, \mathbf{v}), \quad \mathbf{x}(t_0) = \mathbf{x}_0; \qquad t \in [t_0, \vartheta], \quad \mathbf{x} \in \mathbb{R}^n, \quad \mathbf{u} \in P \subset \mathbb{R}^k, \quad \mathbf{v} \in Q \subset \mathbb{R}^l, \tag{1.1}$$

[1]Работа выполнена при финансовой поддержке РФФИ (грант 09–01–00313).

где $\vartheta$ — фиксированный момент окончания игры; $\mathbf{u}$ и $\mathbf{v}$ — векторные управления, подчиненные первому и второму игрокам соответственно; $P$ и $Q$ — выпуклые компакты.

Будем считать выполненными следующие предположения.

**1.** Существует компактное множество $G \subset [t_0, \vartheta] \times \mathbb{R}^n$ такое, что $(t_0, \mathbf{x}_0) \in G$, и все траектории системы (1.1), начинающиеся в произвольной позиции $(t_*, \mathbf{x}_*) \in G$, остаются в $G$ при всех $t \in (t_*, \vartheta]$.

**2.** Множество $F(t, \mathbf{x}) = \{\, f(t, \mathbf{x}, \mathbf{u}, \mathbf{v}) \mid \mathbf{u} \in P, \mathbf{v} \in Q \,\}$ выпукло в любой позиции $(t, \mathbf{x}) \in G$.

**3.** Функция $f \colon G \times P \times Q \mapsto \mathbb{R}^n$ непрерывна по совокупности переменных.

**4.** Существует константа $L > 0$ такая, что

$$\|f(t, \mathbf{x}', \mathbf{u}, \mathbf{v}) - f(t, \mathbf{x}'', \mathbf{u}, \mathbf{v})\| \leqslant L \|\mathbf{x}' - \mathbf{x}''\|$$

при всех $(t, \mathbf{x}') \in G$, $(t, \mathbf{x}'') \in G$, $\mathbf{u} \in P$, $\mathbf{v} \in Q$.

**5.** Существует константа $\lambda > 0$ такая, что

$$\|f(t, \mathbf{x}, \mathbf{u}, \mathbf{v})\| \leqslant \lambda (1 + \|\mathbf{x}\|)$$

при всех $(t, \mathbf{x}) \in G$, $\mathbf{u} \in P$, $\mathbf{v} \in Q$.

Цель игрока $i$ состоит в максимизации терминального показателя качества

$$I_i = \sigma_i(\mathbf{x}(\vartheta)), \quad i = 1, 2, \tag{1.2}$$

где $\sigma_i \colon \mathbb{R}^n \mapsto \mathbb{R}^1$ — непрерывные функции.

Далее будем считать, что непустые множества уровня показателей качества игроков

$$M_i^c = \{\, \mathbf{x} \in \mathbb{R}^n \mid \sigma_i(\mathbf{x}) \geqslant c \,\} \tag{1.3}$$

выпуклы, а их границы

$$\partial M_i^c = \{\, \mathbf{x} \in \mathbb{R}^n \mid \sigma_i(\mathbf{x}) = c \,\} \tag{1.4}$$

являются гладкими.

Полагаем, что оба игрока имеют полную информацию о текущей позиции игры $(t, \mathbf{x}(t))$. При формализации стратегий игроков и порождаемых ими движений в рассматриваемой неантагонистической игре будем опираться на формализацию, введенную для антагонистических позиционных дифференциальных игр [1; 2]. Используемая формализация неантагонистической позиционной дифференциальной игры взята из [3]. В качестве стратегий игроков рассматриваются чистые позиционные стратегии.

**Определение 1.** Пара $U \div \{\, \mathbf{u}(t, \mathbf{x}, \varepsilon), \beta_1(\varepsilon) \,\}$, где $\mathbf{u} \colon G \times (0, \infty) \mapsto P$ — произвольная функция позиции $(t, \mathbf{x})$ и положительного параметра точности $\varepsilon$, а функция $\beta_1 \colon (0, \infty) \mapsto (0, \infty)$ является непрерывной, монотонной и удовлетворяет условию $\beta_1(\varepsilon) \to 0$ при $\varepsilon \to 0$, называется *стратегией* игрока 1.

Функция $\beta_1(\cdot)$ имеет следующий смысл: при фиксированном $\varepsilon$ величина $\beta_1(\varepsilon)$ служит ограничением сверху на шаг разбиения отрезка $[t_0, \vartheta]$, используемого игроком 1 при построении ломаных Эйлера.

Стратегия $V \div \{\, \mathbf{v}(t, \mathbf{x}, \varepsilon), \beta_2(\varepsilon) \,\}$ игрока 2 вводится аналогично.

Рассматриваются движения двух типов, порождаемые парой стратегий игроков: *аппроксимационные* (ломаные Эйлера) и *идеальные* (предельные). Аппроксимационное движение $\mathbf{x}[\cdot; t_0, \mathbf{x}_0; U, \varepsilon_1, \Delta_1, V, \varepsilon_2, \Delta_2]$ вводится для фиксированных значений параметров точности игроков $\varepsilon_1$ и $\varepsilon_2$ и для фиксированных конечных разбиений $\Delta_1 = \{\, t_j^{(1)} \,\}$ и $\Delta_2 = \{\, t_k^{(2)} \,\}$ отрезка $[t_0, \vartheta]$, выбираемых игроком 1 и игроком 2 соответственно, с соблюдением следующих условий, наложенных на шаги разбиений:

$$\delta(\Delta_i) = \max_j (t_{j+1}^{(i)} - t_j^{(i)}) \leqslant \beta_i(\varepsilon_i), \quad i = 1, 2.$$

Таким образом, *закон управления* игрока 1 определяется как тройка $(U, \varepsilon_1, \Delta_1)$, отвечающая стратегии $U$. Аналогично закон управления игрока 2 определяется как тройка $(V, \varepsilon_2, \Delta_2)$.

Предельное движение, порожденное парой стратегий $(U, V)$ из начальной позиции $(t_0, \mathbf{x}_0)$, определяется как непрерывная функция $\mathbf{x}[t] = \mathbf{x}[t; t_0, \mathbf{x}_0; U, V]$, для которой существует последовательность аппроксимационных движений $\{\,\mathbf{x}[\cdot; t_0^k, \mathbf{x}_0^k; U, \varepsilon_1^k, \Delta_1^k, V, \varepsilon_2^k, \Delta_2^k\,\}$, равномерно сходящаяся к $\mathbf{x}[t]$ на $[t_0, \vartheta]$ при $k \to \infty$, $\varepsilon_1^k \to 0$, $\varepsilon_2^k \to 0$, $t_0^k \to t_0$, $\mathbf{x}_0^k \to \mathbf{x}_0$, $\delta(\Delta_i^k) \leqslant \beta_i(\varepsilon_i^k)$.

Пара стратегий $(U, V)$ порождает непустое компактное (в метрике пространства $\mathrm{C}[t_0, \vartheta]$) множество $X(t_0, \mathbf{x}_0, U, V)$ предельных движений.

Далее будем считать, что законы управления игроков согласованы по параметру точности: $\varepsilon_1 = \varepsilon_2 = \varepsilon$.

**Определение 2.** Пара стратегий $(U^N, V^N)$ называется *равновесным по Нэшу решением* (*$N$-решением*) *игры*, если для любого движения $\mathbf{x}^*[\cdot] \in X(t_0, \mathbf{x}_0, U^N, V^N)$, любого $\tau \in [t_0, \vartheta]$ и любых стратегий $U$ и $V$ выполнены следующие неравенства:

$$\max_{\mathbf{x}[\cdot]} \sigma_1(\mathbf{x}[\vartheta, \tau, \mathbf{x}^*[\tau], U, V^N]) \leqslant \min_{\mathbf{x}[\cdot]} \sigma_1(\mathbf{x}[\vartheta, \tau, \mathbf{x}^*[\tau], U^N, V^N]),$$
$$\max_{\mathbf{x}[\cdot]} \sigma_2(\mathbf{x}[\vartheta, \tau, \mathbf{x}^*[\tau], U^N, V]) \leqslant \min_{\mathbf{x}[\cdot]} \sigma_2(\mathbf{x}[\vartheta, \tau, \mathbf{x}^*[\tau], U^N, V^N]),$$

где максимумы и минимумы вычисляются по соответствующим множествам предельных движений. Траекторию, порожденную $N$-решением, будем называть *$N$-траекторией.*

**Определение 3.** $N$-решение $(U^P, V^P)$, являющееся неулучшаемым по Парето относительно величин $I_1$, $I_2$ (1.2), будем называть *$P$-решением.* Иначе говоря, при переходе от $P$-решения к любому другому $N$-решению строгое увеличение выигрыша одного из игроков возможно лишь при строгом уменьшении выигрыша другого. Соответственно траекторию, порожденную $P$-решением, будем называть *$P$-траекторией.*

**Основная задача.** Построить аппроксимации множества $\mathcal{N}$ всех $N$-решений и множества $\mathcal{P}$ всех $P$-решений в рассматриваемой игре.

## § 2. Вспомогательные антагонистические игры и структура решений

Введем вспомогательные антагонистические позиционные дифференциальные игры $\Gamma_1$ и $\Gamma_2$. Динамика обеих игр описывается (1.1). В игре $\Gamma_i$ игрок $i$ максимизирует свой показатель качества $\sigma_i(\mathbf{x}[\vartheta])$ (1.2), в то время как игрок $3-i$ противодействует ему. Игроки действуют в классах чистых позиционных стратегий $U \div \mathbf{u}(t, \mathbf{x}, \varepsilon)$ и $V \div \mathbf{v}(t, \mathbf{x}, \varepsilon)$.

Далее будем считать выполненным условие седловой точки в маленькой игре [1, с. 56]:

$$\max_{\mathbf{u}\in P} \min_{\mathbf{v}\in Q} \mathbf{s}^\top f(t, \mathbf{x}, \mathbf{u}, \mathbf{v}) = \min_{\mathbf{v}\in Q} \max_{\mathbf{u}\in P} \mathbf{s}^\top f(t, \mathbf{x}, \mathbf{u}, \mathbf{v}).$$

Тогда в играх $\Gamma_1$ и $\Gamma_2$ существуют универсальные седловые точки [1, 2]

$$\{\,\mathbf{u}^{(i)}(t, \mathbf{x}, \varepsilon),\ \mathbf{v}^{(i)}(t, \mathbf{x}, \varepsilon)\,\}, \quad i = 1, 2, \tag{2.1}$$

и непрерывные функции цены $\gamma_1(t, \mathbf{x})$ и $\gamma_2(t, \mathbf{x})$.

**Вспомогательная задача** (см. [3, с. 26]). Найти измеримые управления $\mathbf{u}(\cdot) = \{\,\mathbf{u}(t),\ t_0 \leqslant t \leqslant \vartheta,\ \mathbf{u}(t) \in P$ п.в.$\,\}$ и $\mathbf{v}(\cdot) = \{\,\mathbf{v}(t),\ t_0 \leqslant t \leqslant \vartheta,\ \mathbf{v}(t) \in Q$ п. в.$\,\}$ такие, что порожденная ими траектория $\mathbf{x}(t)$, $t_0 \leqslant t \leqslant \vartheta$ удовлетворяет неравенствам

$$\gamma_i(t, \mathbf{x}(t)) \leqslant \gamma_i(\vartheta, \mathbf{x}(\vartheta)) = \sigma_i(\mathbf{x}(\vartheta)), \quad t_0 \leqslant t \leqslant \vartheta, \quad i = 1, 2. \tag{2.2}$$

Пусть измеримые управления $\mathbf{u}^*(t)$ и $\mathbf{v}^*(t)$ порождают траекторию $\mathbf{x}^*(t)$ системы (1.1). Рассмотрим стратегии игрока 1 и игрока 2

$$U^0 \div \{\,\mathbf{u}^0(t, \mathbf{x}, \varepsilon), \beta_1^0(\varepsilon)\,\}, \quad V^0 \div \{\,\mathbf{v}^0(t, \mathbf{x}, \varepsilon), \beta_2^0(\varepsilon)\,\}, \tag{2.3}$$

где

$$\mathbf{u}^0(t,\mathbf{x},\varepsilon) = \begin{cases} \mathbf{u}^*(t), & \text{при } \|\mathbf{x}-\mathbf{x}^*(t)\| < \varepsilon\varphi(t), \\ \mathbf{u}^{(2)}(t,\mathbf{x},\varepsilon), & \text{при } \|\mathbf{x}-\mathbf{x}^*(t)\| \geqslant \varepsilon\varphi(t), \end{cases}$$
$$\mathbf{v}^0(t,\mathbf{x},\varepsilon) = \begin{cases} \mathbf{v}^*(t), & \text{при } \|\mathbf{x}-\mathbf{x}^*(t)\| < \varepsilon\varphi(t), \\ \mathbf{v}^{(1)}(t,\mathbf{x},\varepsilon), & \text{при } \|\mathbf{x}-\mathbf{x}^*(t)\| \geqslant \varepsilon\varphi(t) \end{cases} \tag{2.4}$$

для всех $t \in [t_0, \vartheta]$. Функции $\beta_i(\cdot)$ и положительная непрерывная неубывающая функция $\varphi(\cdot)$ выбираются так, что неравенство

$$\left\|\mathbf{x}(t,t_0,\mathbf{x}_0,U^0,\varepsilon,\Delta_1,V^0,\varepsilon,\Delta_2) - \mathbf{x}^*(t))\right\| < \varepsilon\varphi(t) \tag{2.5}$$

выполняется при $t \in [t_0, \vartheta]$, $\varepsilon > 0$, $\delta(\Delta_i) \leqslant \beta_i(\varepsilon)$. Стратегии $\mathbf{u}^{(2)}(\cdot,\cdot,\cdot)$ и $\mathbf{v}^{(1)}(\cdot,\cdot,\cdot)$ определены в (2.1). Они могут быть интерпретированы как универсальные «стратегии наказания», используемые одним игроком в случае, когда другой игрок отказывается в некоторый момент времени $t \in [t_0, \vartheta]$ отслеживать траекторию $\mathbf{x}^*(\cdot)$.

**Теорема 1** (см. [3, с. 28]). *Пусть управления $\mathbf{u}^*(\cdot)$ и $\mathbf{v}^*(\cdot)$ являются решением вспомогательной задачи. Тогда пара стратегий $(U^0, V^0)$ (2.3) – (2.5) доставляет N-решение. Обратно, для любого N-решения найдется эквивалентное N-решение в форме $(U^0, V^0)$, где $\mathbf{u}^*(\cdot)$ и $\mathbf{v}^*(\cdot)$ есть решение вспомогательной задачи.*

Приведенная теорема устанавливает связь между множеством решений вспомогательной задачи и множеством $\mathcal{N}$, а теорема существования $N$-решений является ее следствием. Одновременно определяется структура решений игры. Согласно изложенному $P$-решения определяются как $N$-решения, неулучшаемые по Парето относительно показателей качества игроков.

## § 3. Алгоритм

Ввиду того что функции $\gamma_i(t,\mathbf{x})$, как правило, в аналитической форме не находятся, предлагаемый алгоритм использует *максимальный стабильный мост* — известную конструкцию из теории антагонистических позиционных дифференциальных игр. Для этого вводятся вспомогательные позиционные дифференциальные игры сближения-уклонения [1; 2] $\Gamma_i^c$, динамика которых описывается (1.1), в качестве целевых множеств в этих играх берутся множества уровня $M_i^c$ (1.3). Цель игрока $i$ в игре $\Gamma_i^c$ состоит в приведении фазового вектора $\mathbf{x}[\vartheta]$ на множество $M_i^c$.

Обозначим через $W_i^c$ максимальный стабильный мост в задаче сближения с множеством $M_i^c$. Из теоремы об альтернативе [1] и теоремы из § 2 следует

**Утверждение 1.** *Траектория $\mathbf{x}[t]$ системы (1.1) будет N-траекторией тогда и только тогда, когда найдутся числа $c_1$ и $c_2$ такие, что $(t,\mathbf{x}[t]) \notin \mathrm{int}W_1^{c_1} \cup \mathrm{int}W_2^{c_2}$ для $t \in [t_0, \vartheta)$, и $\mathbf{x}[\vartheta] \in \partial M_1^{c_1} \cap \partial M_2^{c_2}$.*

Далее приведем алгоритм, строящий аппроксимацию множества концов всех $P$-траекторий или всех $N$-траекторий для линейной системы с динамикой

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)\mathbf{u} + \mathbf{C}(t)\mathbf{v}, \quad \mathbf{x}(t_0) = \mathbf{x}_0, \quad \mathbf{x} \in \mathbb{R}^n, \tag{3.1}$$

где матричные функции $\mathbf{A}(t)$, $\mathbf{B}(t)$, $\mathbf{C}(t)$ покомпонентно непрерывны на $[t_0, \vartheta]$ и имеют размерности $n \times n$, $n \times k$ и $n \times l$ соответственно.

Кроме того, будем считать, что показатели качества игроков $\sigma_i(\cdot)$ зависят только от двух координат фазового вектора $x_\kappa$ и $x_\lambda$, $1 \leqslant \kappa < \lambda \leqslant n$, тогда система (3.1) известной заменой переменных [1] приводится к виду (для удобства изложения новую переменную обозначим снова буквой $\mathbf{x}$) :

$$\dot{\mathbf{x}} = \mathbf{B}^*(t)\mathbf{u} + \mathbf{C}^*(t)\mathbf{v}, \quad \mathbf{x}(t_0) = \mathbf{x}_0, \quad \mathbf{x} \in \mathbb{R}^2. \tag{3.2}$$

При этом существенно, что показатели качества игроков сохранят прежний вид. Ограничения на выбор управлений игроками $P$ и $Q$ будем считать выпуклыми многоугольниками.

Сформулированное утверждение 1 является базовым в предлагаемом алгоритме. Будем считать, что игроки используют одинаковое разбиение временного отрезка при построении ломаных Эйлера: $\Delta_1 = \Delta_2 = \Delta$. Система (3.2) заменяется конечно-разностным уравнением с разбиением временного отрезка $\Delta = \{\, t_j \,\}$, $j = 1,\ldots,N$, $t_1 = t_0$, $t_N = \vartheta$. Обозначим $\delta_j = t_{j+1} - t_j$.

Заметим, что числа $c_i$ в утверждении 1 равны значениям выигрышей игроков на рассматриваемой $N$-траектории $\sigma_i(\mathbf{x}[\vartheta])$. При заданных $c_1, c_2$ будем называть *незапрещенными* те позиции $(t,\mathbf{x})$, для которых выполняются неравенства $\gamma_i(t,\mathbf{x}) \leqslant c_i$, $i = 1,2$. *Множество незапрещенных позиций*, соответствующее фиксированным выигрышам игроков $c_1$, $c_2$, будем обозначать через $G^{c_1,c_2}$. Аппроксимации сечений $W_i^{c_i}$ и $G^{c_1,c_2}$ плоскостью $t = \mathrm{const}$ будем обозначать через $\widetilde{W}_{i,t}^{c_i}$ и $\widetilde{G}_t^{c_1,c_2}$.

Для заданных значений выигрышей мы можем построить аппроксимацию множества незапрещенных позиций $\widetilde{G}^{c_1,c_2}$ в виде набора полигональных сечений в моменты времени, заданные разбиением $\Delta$, используя следующую дискретную схему:

$$\widetilde{G}_{t_{j+1}}^{c_1,c_2} = \mathrm{cl}\left\{ \left[ \widetilde{G}_{t_j}^{c_1,c_2} \oplus \delta_j \left( \mathbf{B}^*(t_j) P \oplus \mathbf{C}^*(t_j) Q \right) \right] \setminus \left( \widetilde{W}_{1,t_{j+1}}^{c_1} \cup \widetilde{W}_{2,t_{j+1}}^{c_2} \right) \right\}, \tag{3.3}$$

где $\widetilde{G}_{t_0}^{c_1,c_2} = \{\, \mathbf{x}_0 \,\}$. Операция $A \oplus B = \{\, a+b \mid a \in A, b \in B \,\}$ обозначает алгебраическую сумму двух множеств $A$ и $B$. Знак $\mathrm{cl}\, A$ обозначает замыкание множества $A$.

Аппроксимации максимальных стабильных мостов $\widetilde{W}_i^{c_i}$ в линейной игре можно строить используя, например, процедуру из [4].

Нетрудно заключить, что если найдется точка $p \in \widetilde{G}_{i,\vartheta}^{c_1,c_2} \cap \widetilde{W}_{1,\vartheta}^{c_1} \cap \widetilde{W}_{2,\vartheta}^{c_2}$, то можно построить аппроксимацию $N$-траектории, находящуюся внутри $\widetilde{G}_{i,\vartheta}^{c_1,c_2}$ и приводящую в точку $p$ при $t = \vartheta$. Ей будет соответствовать $N$-решение вида (2.3), (2.4), доставляющее игроку 1 и игроку 2 выигрыши $c_1$ и $c_2$ соответственно.

Перед тем как переходить к описанию собственно алгоритма, необходимо отметить, что выигрыш игрока $i$ на $N$-решении не может быть меньше $\gamma_i(t_0,\mathbf{x}_0)$. Далее, пусть выигрыш $c_i$ игрока $i$ зафиксирован, тогда на множестве $\partial M_i^{c_i} \neq \varnothing$ найдутся точки минимума и максимума показателя $\sigma_{3-i}(\cdot)$ игрока $3-i$.

Во *внешнем цикле* алгоритма происходит поочередный выбор выигрышей $g_s^1$ игрока 1 и $g_s^2$ игрока 2 так, что $g_s^i = \gamma_i(t_0,\mathbf{x}_0) + s\delta g^i + \epsilon$, где $s \in \mathbb{N}$, $\delta g^i$ — параметры точности, а $\epsilon$ — некоторая малая величина. Тело этого цикла может состоять из одной (если производится поиск концов только $P$-траекторий) или двух частей (если производится поиск концов всех $N$-траекторий).

Положим $c_i = g_s^i$. Идея *первой части* внешнего цикла состоит в том, что в качестве начальной итерации берется $\max \sigma_{3-i}(\mathbf{x})$ при $\mathbf{x} \in \partial M_i^{c_i}$, а при поиске очередного приближения используется (3.3). Одновременно находится конец соответствующей $N$-траектории, которая может быть аппроксимацией $P$-траектории.

*Вторая часть* внешнего цикла представляет собой перебор значений (уменьшение) $c_{3-i}$ с шагом $\delta g^{3-i}$, начиная с найденного в первой части максимального значения и до значения $g_s^{3-i}$. Для каждой пары выигрышей $c_1$, $c_2$ с помощью (3.3) находится конец соответствующей $N$-траектории.

В дальнейшем интерес будет представлять множество $\widetilde{D}_i^{c_1,c_2} = \widetilde{G}_\vartheta^{c_1,c_2} \cap \widetilde{W}_{i,\vartheta}^{c_i}$. Можно считать $\widetilde{D}_i^{c_1,c_2}$ набором вершин из последнего сечения множества незапрещенных позиций, попавших на мост для игрока $i$.

1. Положим, что известны $c_1^*$ и $c_2^*$ такие, что $\mathbf{x}_0 \in \partial W_{1,t_0}^{c_1^*} \cap \partial W_{2,t_0}^{c_2^*}$, заданы положительные шаги $\delta g^1$, $\delta g^2$ и параметр точности $\epsilon_{\mathbf{p}} > 0$.

2. Обозначим через $\widetilde{K}_1$, $\widetilde{K}_2$ части строимого множества $\widetilde{K}$, где $\widetilde{K}_i$ наполняется при переборе значений выигрыша игрока $i$, и пусть $g^1 = c_1^* + \epsilon_1$, $g^2 = c_2^* + \epsilon_2$, где $\epsilon_i$ — некие малые положительные величины, меньшие $\delta g^i$.

3. Вначале $\widetilde{K}_1 = \widetilde{K}_2 = \varnothing$, «признаки останова внешнего цикла» $stop_1 = stop_2 = \text{false}$.

4. Для $i = 1, 2$, если не $stop_i$, то найдем конец $N$-траектории $\mathbf{p}_i$ для фиксированного выигрыша $g^i$ игрока $i$ («первая часть»):

   (a) Обозначим $c_i = g^i$ (здесь $i$ фиксировано: либо 1, либо 2).

   (b) Пусть $c_{3-i} = \max\limits_{\mathbf{x}\in\partial M_i^{c_i}} \sigma_{3-i}(\mathbf{x})$, а $\mathbf{p}$ — точка на $\partial M_i^{c_i}$, доставляющая этот выигрыш.

   (c) Строим $\widetilde{G}^{c_1,c_2}$ используя (3.3) (при этом $\mathbf{p}$ должна принадлежать аппроксимации границы целевого множества $M_{3-i}^{c_{3-i}}$ перед построением соответствующего моста).

   (d) Если существует $\mathbf{p}^* = \arg\max \sigma_i(\mathbf{x})$ при $\mathbf{x} \in \widetilde{D}_i^{c_1,c_2}$, то

   - если $\|\mathbf{p} - \mathbf{p}^*\| \leqslant \epsilon_{\mathbf{p}}$, тогда $\mathbf{p}_i = \mathbf{p}^*$;
   - иначе $\mathbf{p} := \mathbf{p}^*$, $c_{3-i} := \sigma_{3-i}(\mathbf{p}^*)$, повтор с 4c.

   (e) Иначе $stop_i = \text{true}$.

5. Если $stop_1 \wedge stop_2$ или $\sigma_1(\mathbf{p}_2) \leqslant g^1 \wedge \sigma_2(\mathbf{p}_1) \leqslant g^2$, то переходим к 8.

6. Пополним $\widetilde{K}_1 \Leftarrow \mathbf{p}_1$, $\widetilde{K}_2 \Leftarrow \mathbf{p}_2$.

7. Изменим $g^1 := g^1 + \delta g^1$ и $g^2 := g^2 + \delta g^2$, повтор с 4.

8. Возвращаем множество концов траекторий $\widetilde{K} = \widetilde{K}_1 \cup \widetilde{K}_2$, выход.

После исключения доминируемых элементов из $\widetilde{K}$ оно будет представлять собой аппроксимацию множества концов $P$-траекторий. Под доминируемыми элементами понимаются точки $\mathbf{p}$, для которых найдется $\mathbf{p}^* \in \widetilde{K}$ такая, что либо $\sigma_1(\mathbf{p}) < \sigma_1(\mathbf{p}^*) \wedge \sigma_2(\mathbf{p}) \leqslant \sigma_2(\mathbf{p}^*)$, либо $\sigma_1(\mathbf{p}) \leqslant \sigma_1(\mathbf{p}^*) \wedge \sigma_2(\mathbf{p}) < \sigma_2(\mathbf{p}^*)$.

В случае построения аппроксимации множества концов всех $N$-траекторий пункт 6 дополняется двумя циклами.

Для $i = 1, 2$, для $1 \leqslant s < \left\lceil \frac{\sigma_{3-i}(\mathbf{p}_i) - g^{3-i}}{\delta g^{3-i}} \right\rceil$ найти $\mathbf{p}_s^i$ — конец $N$-траектории, доставляющей выигрыш $c_i$ игроку $i$ и выигрыш $c_{3-i} = \sigma_{3-i}(\mathbf{p}_i) - s\delta g^{3-i}$ игроку $3 - i$, используя (3.3), пополнить $\widetilde{K}_i$ найденной точкой $\mathbf{p}_s^i$.

Теперь, взяв из полученного множества концов $P$- (или $N$-) траекторий любую точку, можно восстановить попятной процедурой некоторую соответствующую ей аппроксимационную траекторию, двигаясь внутри соответствующего множества незапрещенных позиций $\widetilde{G}^{c_1,c_2}$, одновременно вычисляя порождающие эту траекторию кусочно-постоянные управления $\mathbf{u}(t) \equiv \mathbf{u}_{t_k}$, $\mathbf{v}(t) \equiv \mathbf{v}_{t_k}$, $t \in [t_k, t_{k+1})$.

Переход от $t_k$ к $t_{k-1}$ можно совершать аналогично дискретной схеме (3.3), обратив ее во времени и вместо высекания объединения сечений мостов $\widetilde{W}_{i,t_{k-1}}^{c_i}$ брать на каждом шаге пересечение с сечением множества незапрещенных позиций $\widetilde{G}_{t_{k-1}}^{c_1,c_2}$ (что и означает «двигаться внутри») с запоминанием элементов $P$ и $Q$, давших результирующую вершину в процедуре построения алгебраической суммы.

## § 4. Пример

Пусть уравнение

$$\ddot{\xi} = \mathbf{u} + \mathbf{v}, \quad \xi(t_0) = \xi_0, \quad \dot{\xi}(t_0) = \dot{\xi}_0; \quad \xi, \mathbf{u}, \mathbf{v} \in \mathbb{R}^2, \quad \|\mathbf{u}\| \leqslant \mu, \quad \|\mathbf{v}\| \leqslant \nu \tag{4.1}$$

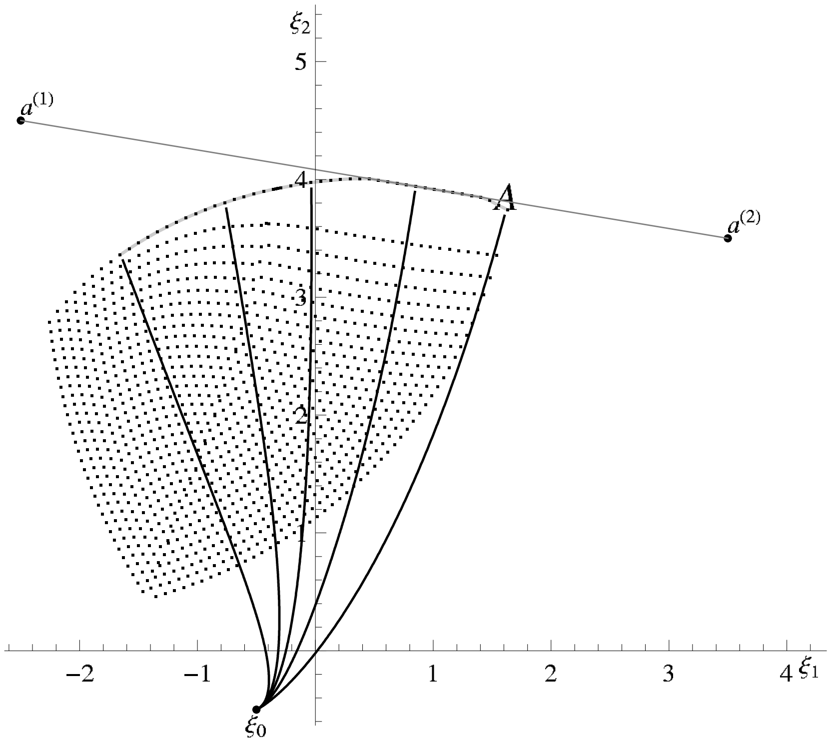


**Рис. 1.** Множество концов $N$-траекторий, $\mu = 1.3$, $\nu = 0.7$

описывает движение материальной точки единичной массы на плоскости $(\xi_1, \xi_2)$ под действием силы $\mathbf{F} = \mathbf{u} + \mathbf{v}$. Игрок 1 (игрок 2), выбирающий управление $\mathbf{u}$ $(\mathbf{v})$, стремится привести материальную точку как можно ближе к заданной целевой точке $a^{(1)}$ $(a^{(2)})$ в момент времени $\vartheta$. Итак, имеем следующие показатели качества игроков:

$$\sigma_i(\xi(\vartheta)) = -\|\xi(\vartheta) - a^{(i)}\|, \tag{4.2}$$
$$\xi = (\xi_1, \xi_2), \quad a^{(i)} = (a_1^{(i)}, a_2^{(i)}), \quad i = 1, 2,$$

где $\vartheta$ — момент окончания игры.

Обозначив $y_1 = \xi_1$, $y_2 = \dot{\xi}_1$, $y_3 = \xi_2$, $y_4 = \dot{\xi}_2$ и выполнив замену переменных $x_1 = y_1 + (\vartheta - t)y_3$, $x_2 = y_2 + (\vartheta - t)y_4$, $x_3 = y_3$, $x_4 = y_4$, мы получаем систему со следующими двумя первыми уравнениями:

$$\dot{x}_1 = (\vartheta - t)(u_1 + v_1), \tag{4.3}$$
$$\dot{x}_2 = (\vartheta - t)(u_2 + v_2).$$

Далее, (4.2) может быть записано как

$$\sigma_i(\mathbf{x}(\vartheta)) = -\|\mathbf{x}(\vartheta) - a^{(i)}\|, \quad \mathbf{x} = (x_1, x_2), \quad i = 1, 2. \tag{4.4}$$

Так как показатели (4.4) зависят только от переменных $x_1$ и $x_2$, а правая часть (4.3) не зависит от других переменных, то можно заключить, что достаточно рассматривать только приведенную систему (4.3) с показателями качества (4.4).

Тогда начальные условия для (4.3) задаются формулами

$$x_i(t_0) = x_{0i} = \xi_{0i} + (\vartheta - t_0)\dot{\xi}_{0i}, \quad i = 1, 2.$$

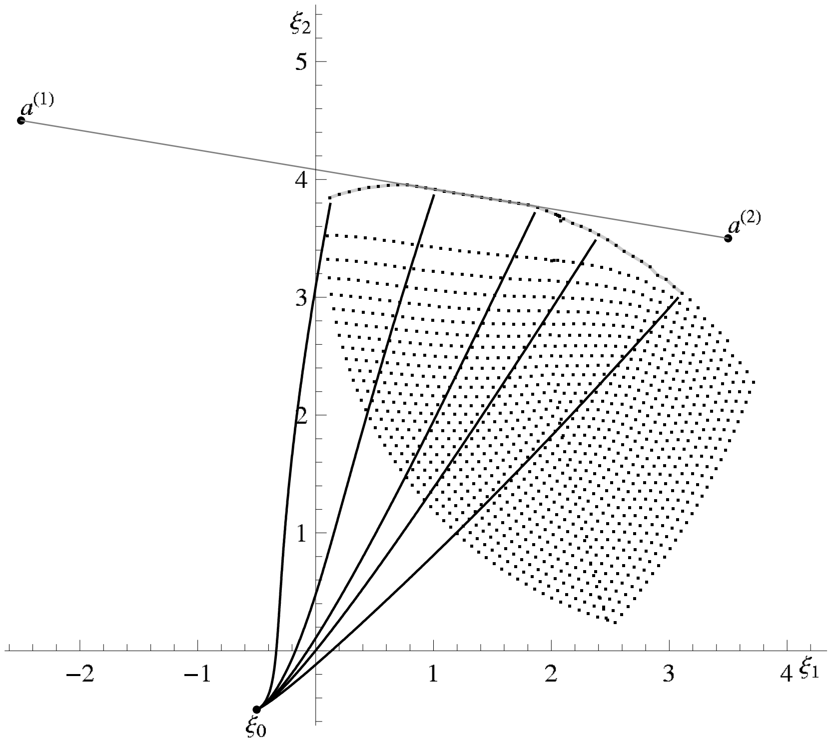


**Рис. 2.** Множество концов $N$-траекторий, $\mu = 0.7$, $\nu = 1.3$

Легко показать, что функции цены в антагонистических дифференциальных играх $\Gamma_1$ и $\Gamma_2$ представимы в виде

$$\gamma_1(t, \mathbf{x}) = \min\{\, -\|\mathbf{x} - a^{(1)}\| + \frac{(\vartheta - t)^2}{2}(\mu - \nu), 0\,\},$$
$$\gamma_2(t, \mathbf{x}) = \min\{\, -\|\mathbf{x} - a^{(2)}\| - \frac{(\vartheta - t)^2}{2}(\mu - \nu), 0\,\},$$

а универсальные оптимальные стратегии (2.1) задаются как

$$\mathbf{u}^{(i)}(t, \mathbf{x}, \varepsilon) = (-1)^i \mu \frac{\mathbf{x} - a^{(i)}}{\|\mathbf{x} - a^{(i)}\|}, \qquad \mathbf{v}^{(i)}(t, \mathbf{x}, \varepsilon) = -(-1)^i \nu \frac{\mathbf{x} - a^{(i)}}{\|\mathbf{x} - a^{(i)}\|}.$$

Были выбраны следующие значения параметров игры: $\vartheta = 2$, целевые точки игроков $a^{(1)} = (-2.5, 4.5)$, $a^{(2)} = (3.5, 3.5)$, $\xi_0 = (-0.5, -0.5)$, $\dot{\xi}_0 = (0.5, 0.25)$. Отсюда получаем $\mathbf{x}_0 = (0.5, 0)$. Использовалось разбиение временного отрезка с постоянным шагом 0.005. Были выбраны шаги $\delta g^1 = \delta g^2 = 0.08$.

Рассматривались два варианта значений $\mu$ и $\nu$. На рис. 1, 3 и 4 приведены результаты численного расчета для значений $\mu = 1.3$ и $\nu = 0.7$. На рис. 1 черными точками показана аппроксимация множества концов всех $N$-траекторий, на рис. 4 представлено отображение этого множества в плоскость $(I_1, I_2)$ значений выигрышей игроков. Серой ломаной и на рис. 1, и на рис. 4 соединены (и таким образом выделены) концы $P$-траекторий. Кроме того, на рис. 1 приведены пять $P$-траекторий (в координатах $(\xi_1, \xi_2)$, полученные интегрированием управлений, вычисленных при расчете траекторий в координатах $(x_1, x_2)$). Параметр точности $\epsilon_{\mathbf{p}}$ был выбран равным $5 \cdot 10^{-4}$. Кроме того, на каждом шаге дискретной схемы (3.3) производилось прореживание контуров с удалением ребер длиной менее $10^{-4}$. Расчет этого варианта на ЭВМ на основе процессора AMD K7 2ГГц занял 17 часов, число полученных концов $N$-траекторий — 1153. Буквой $A$ помечена траектория, для которой на рис. 3 приведены управления $\mathbf{u}(t_k), \mathbf{v}(t_k)$, ее порождающие (каждая четвертая пара).

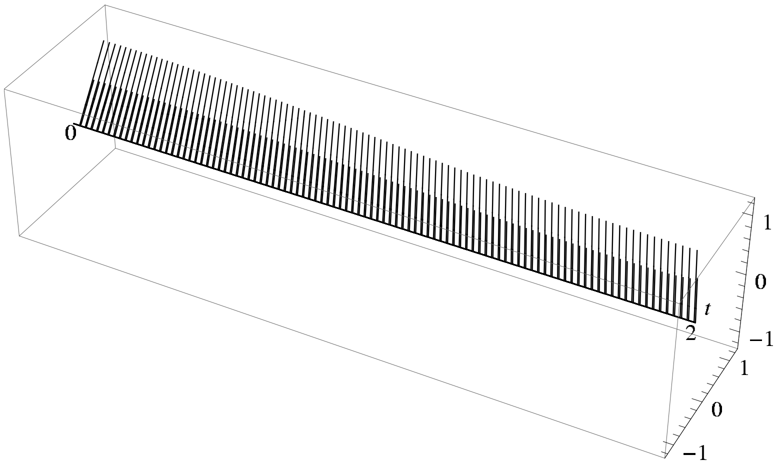


**Рис. 3.** Управления игроков $\mathbf{u}(t_k), \mathbf{v}(t_k)$, $\mu = 1.3$, $\nu = 0.7$

На рис. 2 и 5 приведены результаты расчета с $\mu = 0.7$ и $\nu = 1.3$. В этом варианте была повышена погрешность: $\epsilon_{\mathbf{p}} = 2 \cdot 10^{-3}$, а порог длины ребра при прореживании был выбран равным $5 \cdot 10^{-4}$. Расчет на той же ЭВМ занял менее 2 часов, число полученных концов $N$-траекторий — 1144.

## Заключение

Предполагается разработка обобщения описанных алгоритмов для игры с динамикой с правой частью (1.1) вида $f(t, \mathbf{x}, \mathbf{u}, \mathbf{v}) = f_1(t, \mathbf{x}, \mathbf{u}) + f_2(t, \mathbf{x}, \mathbf{v})$ с использованием процедуры из [5] для построения максимальных стабильных мостов.

Кроме того, планируется переписать текущую реализацию с использованием подхода *обобщенного программирования* (на языке C++) для обеспечения высокого уровня гибкости и оформить ее в виде свободно распространяемой библиотеки. В качестве основной реализации алгоритмов вычислительной геометрии будет использована библиотека (являющаяся открытым ПО) Computational Geometry Algorithms Library (CGAL).

## СПИСОК ЛИТЕРАТУРЫ

1. Красовский Н. Н., Субботин А. И. Позиционные дифференциальные игры. — М.: Наука, 1974. — 456 с.
2. Красовский Н. Н. Управление динамической системой. — М.: Наука, 1985. — 520 с.
3. Клейменов А. Ф. Неантагонистические позиционные дифференциальные игры. — Екатеринбург: Наука, 1993. — 184 с.
4. Исакова Е. А., Логунова Г. В., Пацко В. С. Построение стабильных мостов в линейной дифференциальной игре с фиксированным моментом окончания // Алгоритмы и программы решения линейных дифференциальных игр (материалы по математическому обеспечению ЭВМ). Свердловск: ИММ УНЦ АН СССР, 1984. — С. 127–158.
5. Тарасьев А. М., Ушаков В. Н., Хрипунов А. П. Об одном вычислительном алгоритме решения игровых задач управления // Прикладная математика и механика. — 1987. — Т. 51, № 2. — С. 216–222.
6. Осипов С. И. О реализации алгоритма построения решений для класса иерархических игр Штакельберга // Автоматика и телемеханика. — 2007. — № 11. — С. 195–208.
7. Осипов С. И. Решение одного класса иерархических дифференциальных игр (методы, алгоритмы, программы): дис. . . . канд. физ.-матем. наук / УрГУ. — Екатеринбург, 2007. — 128 с.
8. Кувшинов Д. Р. Алгоритм численного построения решений по Нэшу в позиционной дифференциальной игре двух лиц // Динамические системы, управление и наномеханика: Тез. докл. Всеросс. конф. Ижевск, 2009. — С. 42.

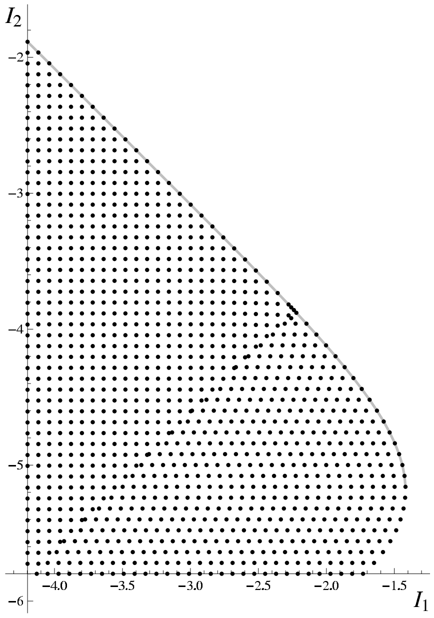

**Рис. 4.** Выигрыши на $N$-решениях,
$\mu = 1.3, \quad \nu = 0.7$

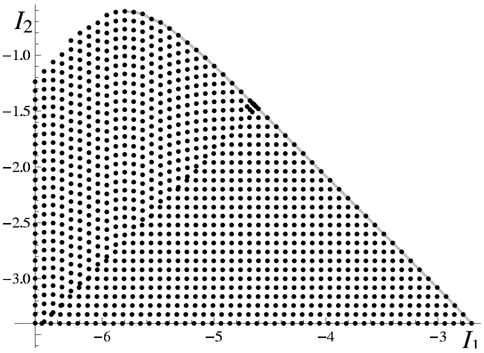

**Рис. 5.** Выигрыши на $N$-решениях,
$\mu = 0.7, \quad \nu = 1.3$

Поступила в редакцию 20.07.09

***D. R. Kuvshinov***

**Numerical construction algorithm for Nash solutions in a two-person positional differential game**

The article presents a numerical algorithm for building an approximation of the Nash solution set in a linear non-zero sum positional differential two-person game with terminal cylindrical cost functionals and geometrical constraints on players' controls.

*Keywords*: positional differential games, Nash equilibrium solution, numerical algorithm.

Mathematical Subject Classifications: 91A05, 91A23, 91A24, 49N70

Кувшинов Дмитрий Рустамович, математик, Институт математики и механики УрО РАН, 620219, Россия, г. Екатеринбург, ул. С. Ковалевской, 16, E-mail: evetro.here@gmail.com