Прикладная математика и механика. Том 42, 1978

УДК 62—50

ОБ ОДНОЙ ЛИНЕЙНОЙ ИГРОВОЙ ЗАДАЧЕ НАВЕДЕНИЯ

М. И. Логинов

(Свердловск)

Рассматривается игровая задача наведения для линейной конфликтно-управляемой системы, когда плата игры имеет смысл евклидова расстояния от фазовой точки до начала координат. Предполагается некоторая модификация правила экстремального прицеливания [1], которая при определенных условиях обеспечивает одному из игроков результат игры, не худший, чем в соответствующей программной задаче на максимин для начальной позиции. Работа опирается на концепцию позиционной дифференциальной игры, разработанную в [1, 2].

1. Рассмотрим конфликтно-управляемую систему, описываемую векторным дифференциальным уравнением

$$y^{\cdot} = A(t)y + B(t)u - C(t)v, \quad u \in P, \quad v \in Q$$

где y — n-мерный фазовый вектор, u, v — r-мерные управляющие воздействия, соответственно первого и второго игроков, $A(t)$, $B(t)$, $C(t)$ — непрерывные по t матрицы соответствующих размерностей P, Q — выпуклые, замкнутые и ограниченные множества.

Игра рассматривается на заданном отрезке времени $t_0 \leqslant t \leqslant \vartheta$, и плата $\gamma[\vartheta]$ изображается равенством

$$\gamma[\vartheta] = \| \{y[\vartheta]\}_m \|$$

Здесь и в дальнейшем $\| x \|$ — евклидова норма вектора x, $\{x\}_m$ — вектор, составленный из первых m компонент вектора x. Известно (см. [2], стр. 160), что рассматриваемая система неособым линейным преобразованием может быть приведена к виду

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

где x — m-мерный вектор, $B(t)$, $C(t)$ — непрерывные по t матрицы, плата игры при этом будет иметь вид

$$\gamma[\vartheta] = \| x[\vartheta] \| \tag{1.2}$$

В дальнейшем удобно работать с системой, преобразованной к виду (1.1).

Первый игрок распоряжается выбором управления $u[t] \in P$ и стремится минимизировать величину $\gamma[\vartheta]$ на траекториях $x[t]$ ($t_0 \leqslant t \leqslant \vartheta$, $x[t_0] = x_0$) системы (1.1), реализующихся под действием его управлений $u[t]$ ($t_0 \leqslant t \leqslant \vartheta$) в паре с любой интегрируемой реализацией $v[t] \in Q$

управления второго игрока. Цель второго игрока противоположна и состоит в максимизации величины $\gamma[\vartheta]$ (1.2).

Допустимые стратегии U и V соответственно первого и второго игроков будем задавать выпуклыми, замкнутыми, полунепрерывными сверху по включению при изменении позиции множествами $U(t, x) \subset P$, $V(t, x) \subset Q$; под движениями будем понимать решения соответствующих уравнений в контингенциях.

Пусть $(\gamma[\vartheta] \mid t_0, x_0, u, v)$ — реализация величины $\gamma[\vartheta]$ (1.2), отвечающая начальной позиции $\{t_0, x_0\}$ при некоторых управлениях u и v.

Задача 1. Среди допустимых стратегий U первого игрока найти стратегию U*, которая при любом допустимом способе управления второго игрока для любой исходной позиции гарантирует результат игры

$$(\gamma[\vartheta] \mid t_0, x_0, \mathrm{U}^*, v) \leqslant \varepsilon_0(t_0, x_0)$$

Задача 2. Среди допустимых стратегий V второго игрока требуется найти стратегию V*, которая при любом допустимом способе управления первого игрока для любой исходной позиции $\{t_0, x_0\}$ гарантирует результат игры

$$(\gamma[\vartheta] \mid t_0, x_0, u, \mathrm{V}^*) \geqslant \varepsilon_0(t_0, x_0)$$

Здесь величина $\varepsilon_0(t_0, x_0)$ — программный максимин для начальной позиции $\{t_0, x_0\}$, который определяется равенством [1]

$$\varepsilon_0(t_0, x_0) = \max_{\|l\|=1} \Big[\int_{t_0}^{\vartheta} \max_{v \in Q} l'C(t)v(t)\,dt - \int_{t_0}^{\vartheta} \max_{u \in P} l'B(t)u(t)\,dt - l'x_0\Big] \tag{1.3}$$

если правая часть этого равенства положительна, иначе $\varepsilon_0(t_0, x_0) = 0$. Штрих означает транспонирование. Будем предполагать, что начальная позиция $\{t_0, x_0\}$ такова, что $\varepsilon_0(t_0, x_0) > 0$.

2. Пусть выполнено условие [2]: функция

$$\varkappa(l, t) = \max_{u \in P} l'B(t)u - \max_{v \in Q} l'C(t)v \tag{2.1}$$

выпукла по l при всех $t \in [t_0, \vartheta]$ (условие А).

Известно, что это условие необходимое и достаточное для того, чтобы максимум в правой части равенства (1.3) достигался на единственном векторе $l_0 = l_0(t_0, x_0)$. Кроме того, при выполнении этого условия функция $\varkappa(l, t)$ будет [2,3] опорной функцией выпуклого замкнутого множества

$$H(t) = \bigcap_{v \in Q} \{B(t)P - C(t)v\} \tag{2.2}$$

Будем рассматривать программные управления $u^0(t, l_0)$ и $v^0(t, l_0)$, $t_0 \leqslant t \leqslant \vartheta$, удовлетворяющие при почти всех t условиям максимума

$$l_0'B(t)u^0(t, l_0) = \max_{u \in P} l_0'B(t)u \tag{2.3}$$

$$l_0'C(t)v^0(t, l_0) = \max_{v \in Q} l_0'C(t)v \tag{2.4}$$

где l_0 — тот вектор $l_0 = l_0(t_0, x_0)$, на котором достигается максимум в правой части равенства (1.3).

Лемма 1. Если множества P и Q выпуклы и справедливо условие А, то существуют такие измеримые по t программные управления $u^\circ(t, l_0)$ и $v^\circ(t, l_0)$, удовлетворяющие при почти всех $t \in [t_0, \vartheta]$ условиям максимума (2.3) и (2.4), что почти всюду на отрезке $[t_0, \vartheta]$ имеет место включение

$$h^\circ(t, l_0) = \{B(t)\, u^\circ(t, l_0) - C(t)\, v^\circ(t, l_0)\} \in H(t) \tag{2.5}$$

Доказательство. Функции $\max_{u\in P} l'B(t)\,u$, $\max_{v\in Q} l'\,C(t)\,v$ — опорные для выпуклых замкнутых и ограниченных множеств $\{B(t)P\}$, $\{C(t)Q\}$. Следовательно, множества $\{B(t)\,U_1\}$, $\{C(t)\,V_1\}$ векторов u° и v°, на которых достигается максимум в правой части равенств (2.2), (2.3) при $l = l_0$,— субдифференциалы соответствующих опорных функций в точке l_0 [4]. Так как функция $\varkappa(l, t)$ выпукла по l и является [2, 3] опорной функцией множества $H(t)$ (2.5), то ее субдифференциал $H_1(t)$ в точке $l = l_0$ в сумме с $\{C(t)\,V_1\}$ дает множество $\{B(t)\,U_1\}$. Отсюда следует справедливость включения (2.5). Осталось показать, что функции $u^\circ(t, l_0)$, $v^\circ(t, l_0)$ можно выбрать измеримыми. Действительно, множества $\{B(t)\,U_1\}$, $\{C(t)\,V_1\}$, $H_1(t)$ полунепрерывны сверху по включению при изменении t, поэтому можно выбрать [1, 5] измеримые функции $C(t)\,v^\circ(t, l_0) \in \{C(t)\,V_1\}$ и $h^\circ(t, l_0) \in H_1(t)$, тогда $B(t)\,u^\circ(t, l_0)$ будет также измеримой как сумма двух измеримых функций.

Дадим теперь определение стратегии $\mathbf{U}^*$ первого игрока. Пусть реализовалась некоторая позиция $\{t, x[t]\}$. На отрезке времени $t \leqslant \tau \leqslant \vartheta$ выберем управляющие воздействия $u^\circ(\cdot\, l_0) = u^\circ(\tau, l_0)$ и $v^\circ(\cdot\, l_0) = v^\circ(\tau, l_0)$, которые при почти всех $\tau \in [t, \vartheta]$ удовлетворяют условиям максимума (2.3), (2.4) и таковы, что имеет место включение (2.5). Рассмотрим движение $x(\tau; t, x[t], u^\circ(\cdot\, l_0), v^\circ(\cdot\, l_0))$, $\tau \in [t, \vartheta]$ системы (1.1), порожденное воздействиями $u = u^\circ(\cdot\, l_0)$ и $v = v^\circ(\cdot\, l_0)$ при начальном условии $x(t; t, x[t], u^\circ(\cdot\, l_0), v^\circ(\cdot\, l_0)) = x[t]$.

Определение 1. Пусть m-мерный вектор $s(t)$ определяется равенством

$$s(t) = -x(\vartheta; t, x[t], u^\circ(\cdot\, l_0), v^\circ(\cdot\, l_0)) \tag{2.6}$$

Тогда стратегия $\mathbf{U}^*$ первого игрока определяется следующим образом:

1) если позиция $\{t, x[t]\}$ такова, что $s(t)$ — ненулевой вектор, то с этой позицией будем сопоставлять множество $U^*(t, x[t])$ всех векторов u^*, которые удовлетворяют условию максимума

$$s'(t)B(t)u^* = \max_{u\in P} s'(t)\, B(t)\, u \tag{2.7}$$

2) если же позиция $\{t, x[t]\}$ такова, что $s(t)$ — нулевой вектор, то будем полагать, что $U^*(t, x[t]) = P$.

Из формулы Коши, определяющей величину $x(\vartheta; t, x[t], u^\circ(\cdot\, l_0), v^\circ(\cdot\, l_0))$, и из результатов работы [1] следует, что стратегия $\mathbf{U}^*$, определенная условиями 1) и 2), допустима.

Теорема 1. Если множества P и Q выпуклы и выполнено условие А, то стратегия $\mathbf{U}^*$ первого игрока, построенная в соответствии с определением 1, гарантирует ему результаты игры $(\gamma[\vartheta] \mid t_0, x_0, \mathbf{U}^*, v) \leqslant \varepsilon_0(t_0, x_0)$ при любом допустимом способе управления второго игрока.

Доказательство. Рассмотрим функцию

$$\varepsilon[t] = \varepsilon(t, x[t]) = \| x(\vartheta; t, x[t], u^\circ(\cdot\, l_0), v^\circ(\cdot\, l_0)) \|^2$$

Стратегия $\mathbf{U}^*$ допустима, поэтому при почти всех t существует производная $d\varepsilon\,[t]\,/\,dt$, которая определяется равенством

$$d\varepsilon\,[t]/dt = 2s'\,(t)\,[h^\circ\,(t,\ l_0) - \{B\,(t)\ u\,[t] - C\,(t)\ v\,[t]\}]$$

По построению множества $H\,(t)$ для любой допустимой реализации $v\,[t]$ найдется допустимое управление $u^{(1)}\,(t)$, такое, что

$$h^\circ\,(t,\ l_0) = \{B\,(t)\ u^{(1)}\,(t) - C\,(t)\,v\,[t]\}$$

Поэтому

$$d\varepsilon\,[t]/dt = 2s'\,(t)\,\{B\,(t)\,u^{(1)}\,(t) - B\,(t)\,u\,[t]\}$$

Из этого равенства и из условия максимума (2.7) следует, что при $u\,[t] = u^*$ для любой позиции $\{t,\,x\}$, где $\varepsilon\,[t] > 0$, при почти всех t справедливо неравенство $d\varepsilon\,[t]\,/\,dt \leqslant 0$. Учитывая теперь, что по определению вспомогательной функции $\varepsilon\,[t]$ имеют место равенства $\varepsilon\,[t_0] = \varepsilon_0{}^2\,(t_0,\ x_0)$, $\gamma^2\,[\vartheta] = \varepsilon\,[\vartheta]$, приходим к выводу о справедливости утверждения теоремы.

Аналогичным образом строится стратегия $\mathbf{V}^*$ второго игрока, разрешающая задачу 2. Пусть функция $\varkappa\,(l,t)$ (2.1), фигурирующая в условии A, вогнута по l при каждом $t \in [t_0,\ \vartheta]$, тогда, рассматривая вместо множества $H\,(t)$ множество

$$G(t) = \bigcap_{u \in P} \{B(t)\,u - C(t)\,Q\}$$

можно доказать лемму, аналогичную лемме 1. Стратегию $\mathbf{V}^*$ второго игрока зададим множеством $V^*\,(t,\ x\,[t])$ векторов v^*, удовлетворяющих условию максимума

$$s'\,(t)\,C\,(t)\ v^* = \max\nolimits_{v \in Q}\ s'(t)\,C\,(t)\,v$$

в позициях $\{t,\,x\,[t]\}$, для которых $\|\ s\,(t)\ \| \neq 0$, а в позициях, для которых $s\,(t) = 0$, полагаем $V^*\,(t,\ x\,[t]) = Q$. По тому же плану, как и при доказательстве теоремы 1, можно проверить справедливость следующего утверждения.

Теорема 2. Если множества P и Q выпуклы и функция $\varkappa\,(l,\ t)$ (2.1) вогнута по l при каждом $t \in [t_0,\vartheta]$, то стратегия $\mathbf{V}^*$ второго игрока гарантирует ему результат игры $(\gamma\,[\vartheta]\,|\,t_0,\,x_0,\,u,\,\mathbf{V}^*) \geqslant \varepsilon_0\,(t_0,\ x_0)$ при любом допустимом способе управления первого игрока.

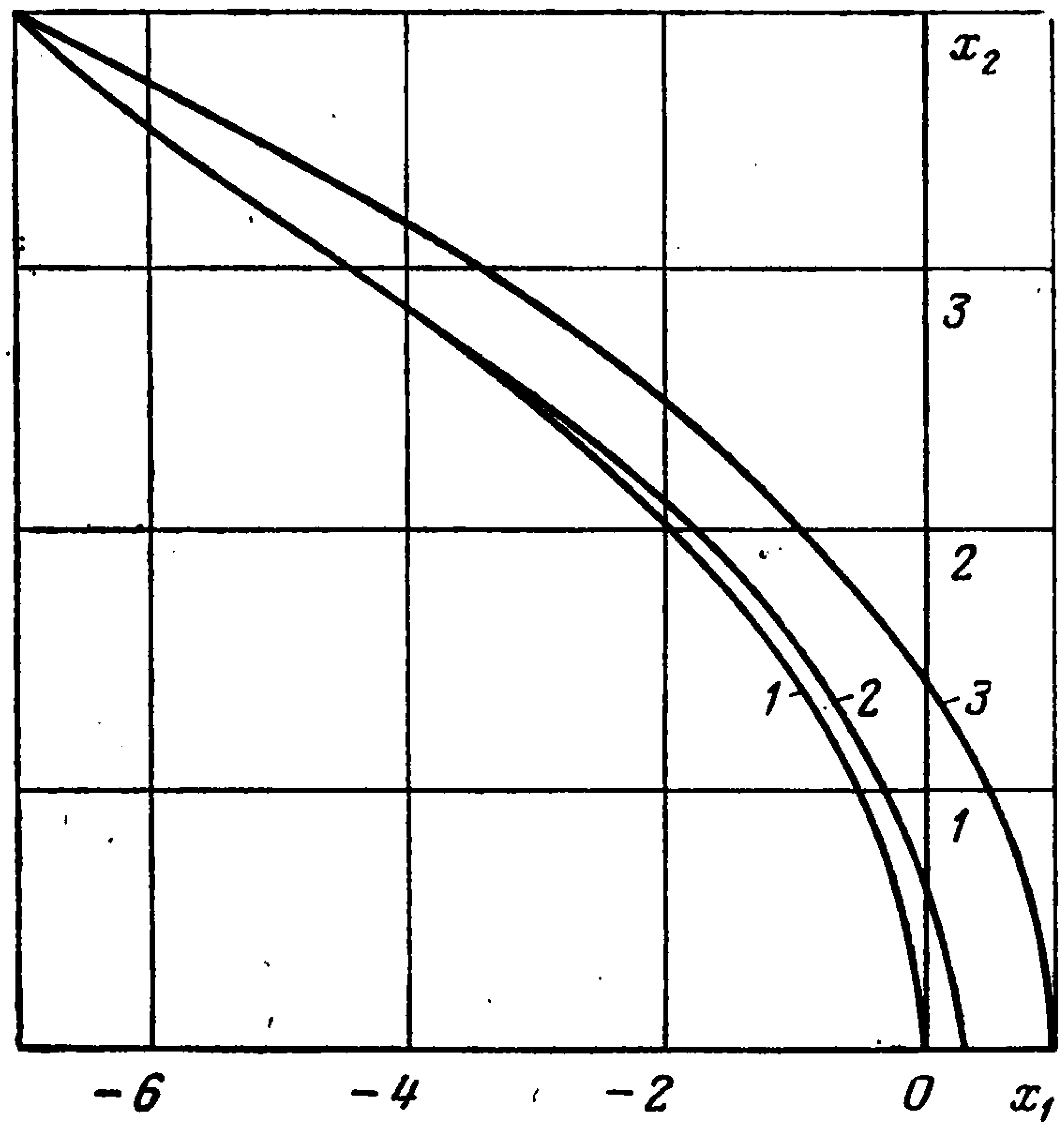
Замечания. 1°. Условие A можно ослабить. Как показывает доказательство теоремы 1, для построения стратегии $\mathbf{U}^*$, разрешающей задачу 1, достаточно, чтобы для начальной позиции $\{t_0,\ x_0\}$ существовали такие оптимальные программные управления $u^\circ\,(t,\ l_0)$ и $v^\circ\,(t,\ l_0)$, $t_0 \leqslant t \leqslant \vartheta$, удовлетворяющие условиям максимума (2.3) и (2.4), что при почти всех $t \in [t_0,\ \vartheta]$ выполняется вложение

$$\{B\,(t)\ P\} \supset \{C\,(t)\ Q\} + h^\circ\,(t,\ l_0)$$

При этом предположение о выпуклости множеств P и Q несущественно, и его можно отбросить.

2°. Особенность предлагаемого способа управления по сравнению с правилом экстремального прицеливания, разработанного в [1], заключается в том, что вектор $s\,(t)$, используемый при определении стратегии игроков, вообще говоря, вычислять проще, чем соответствующий вектор $l^\circ\,[t] = l^\circ\,(t,\ x\,[t])$ в экстремальной конструкции.

Это связано с тем, что для определения вектора $l^0[t]$ необходимо для каждой текущей позиции $\{t, x[t]\}$ решать экстремальную задачу (1.3). В то же время для вычисления вектора $s(t)$ (2.6) требуется знать решение задачи (1.3) только для начальной позиции $\{t_0, x_0\}$. Разумеется, результат получается хуже, чем при использовании правила экстремального прицеливания [1], из-за того, что используются не все «ошибки» противника. Следует отметить, что по сравнению с прямыми методами в теории игр [6] и априори стабильными дорожками [2] предлагаемый в работе способ управления сложнее, но дает лучший результат с точки зрения одного из игроков.



Таким образом, описанный выше способ решения задачи 1 и 2 занимает промежуточное место между правилом экстремального прицеливания и прямыми методами в теории дифференциальных игр.

3°. Можно проверить, что предлагаемая процедура управления для первого игрока гарантирует приведение системы (1.1) в положение $\{x\} = 0$ не позже, чем в момент программного поглощения $\vartheta_0(t_0, x_0)$ при любой допустимой реализации $v[t]$, $t_0 \leqslant t \leqslant \vartheta_0$ управления второго игрока.

3. В качестве примера рассмотрим задачу наведения для конфликтно-управляемой материальной точки, движущейся по горизонтальной прямой. Уравнения движения точки имеют вид

(3.1) $$x_1^{\cdot} = x_2, \quad x_2^{\cdot} = u - v; \quad |u| \leqslant \mu, \quad |v| \leqslant \nu, \quad \mu > \nu.$$

Пусть плата игры γ оценивает расстояние фазовой точки $x[\vartheta]$ в заданный момент времени ϑ до начала координат $x_1 = x_2 = 0$, т. е.

$$\gamma[\vartheta] = \{x_1^2[\vartheta] + x_2^2[\vartheta]\}^{1/2}$$

Для системы (3.1) выполнены все условия теоремы 1, поэтому стратегию U^* первого игрока можно строить в соответствии с определением 1. Как и в работе [1], выберем следующие исходные данные: $x_{01} = -7$, $x_{02} = 4$, $t_0 = 0$, $\vartheta = 4$, $\mu = 2$, $\nu = 1$.

Проделав необходимые вычисления, получаем, что $\varepsilon_0(t_0, x_0) = 1$, максимум в правой части равенства (1.3) достигается на векторе $l_0 = (-1, 0)$ и вектор $s(t)$ (2.6) определяется равенствами

$$s_1(t) = -x_1[t] - x_2[t](\vartheta - t) + {}^1/_2(\vartheta - t)^2$$
$$s_2(t) = -x_2[t] + \vartheta - t$$

Стратегия U^* первого игрока определяется следующим образом:

1) если позиция $\{t, x_1[t], x_2[t]\}$ такова, что $s_1(t)(\vartheta - t) + s_2(t) \neq 0$, то множество $U^*(t, x_1[t], x_2[t])$ состоит из единственной точки

$$u^*[t] = 2\,\mathrm{sign}\,\{s_1(t)(\vartheta - t) + s_2(t)\}$$

2) если позиция $\{t, x_1[t], x_2[t]\}$ такова, что $s_1(t)(\vartheta - t) + s_2(t) = 0$, то $U^*(t, x_1[t], x_2[t]) = P$, т. е. $u^*[t]$ — произвольная величина, удовлетворяющая неравенству $-2 \leqslant u^*[t] \leqslant 2$, для определенности будем полагать в этом случае, что $u^*[t] = 0$.

Реализации движений, диктуемых различным выбором стратегий первого и второго игроков, были просчитаны на ЭЦВМ и представлены на фигуре.

Кривая *1* изображает фазовую траекторию, порожденную оптимальной экстремальной стратегией U° [1] первого игрока, при условии, что второй игрок выбирает управление $v \equiv 0$. Кривая *2* изображает фазовую траекторию, отвечающую описанной в данной статье стратегии U^* первого игрока при управлении второго игрока $v \equiv 0$. Как и следовало ожидать, видно, что в первом случае реализуется величина $\gamma[\vartheta] = 0$, а во втором случае реализуется большее значение платы $\gamma[\vartheta]$, равное 0.258. Кривая *3* порождается парой $\{U^\circ, V^\circ\}$ оптимальных экстремальных стратегий [1], по этой же кривой проходит движение, отвечающее паре стратегий $\{U^*, V^\circ\}$. Отметим еще, что априори стабильная дорожка [2], построенная для данного примера, также изображается кривой *3*.

Автор благодарит Э. Г. Альбрехта и А. И. Субботина за обсуждение работы и критические замечания.

Поступила 22 XI 1977

ЛИТЕРАТУРА

1. *Красовский Н. Н.* Игровые задачи о встрече движений. М., «Наука», 1970.
2. *Красовский Н. Н., Субботин А. И.* Позиционные дифференциальные игры. М., «Наука», 1974.
3. *Рокафеллар Р.* Выпуклый анализ. М., «Мир», 1973.
4. *Пшеничный Б. Н.* Необходимые условия экстремума. М., «Наука», 1969.
5. *Филиппов А. Ф.* Дифференциальные уравнения с разрывной правой частью. Матем. сб., 1960, т. 51 (93), вып. 1.
6. *Понтрягин Л. С.* О линейных дифференциальных играх. 2. Докл. АН СССР, 1967, т. 175, № 4.