Глава 15

Дифференциальные игры

В *дифференциальных играх* исследуются ситуации конфликта или кооперации, в которых игроки осуществляют выбор стратегии во времени [1]. В отличие от материала, изложенного в предыдущих четырех главах, в дифференциальной игре участвуют не менее двух игроков, а выигрыши каждого участника зависят от траекторий управления, принятых всеми участниками игры. С другой стороны, еще одно отличие от игр, рассмотренных в гл. 6, состоит в том, что игроки делают свои ходы в течение некоторого интервала времени, так что число ходов и вместе с ними стратегий бесконечно.

Дифференциальные игры можно классифицировать на основе тех же принципов, по которым классифицировались игры в гл. 6. Один способ — классификация по числу игроков: *дифференциальная игра с двумя, тремя и более участниками*, при этом задачу управления из главы 11 можно рассматривать как особую дифференциальную игру с одним участником. Другая классификация — по характеру платежных функций: *игры с нулевой суммой* и *игры с ненулевой суммой* в зависимости от того, равна или не равна нулю (в общем случае — константе) общая сумма выигрышей всех игроков. Еще один способ — разделение дифференциальных игр на *стохастические*, если в них содержатся случайные переменные, и на *детерминированные* — в противном случае [2]. Если время измеряется в дискретных единицах, то такая игра называется *дискретной дифференциальной игрой*, а если время измеряется как непрерывная величина, то соответствующая игра называется *непрерывной дифференциальной игрой*.

[1] Основная литература по теории игр: [1, 2, 3, 4, 5]. В теории дифференциальных игр используются многие термины теории игр, такие, как «игрок», «стратегия» и «выигрыш». Эти понятия рассматриваются в гл. 6.

[2] Для более глубокого ознакомления со стохастическими дифференциальными играми см. работу Хо [6].

15.1. НЕПРЕРЫВНЫЕ ДЕТЕРМИНИРОВАННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ ИГРЫ ДВУХ УЧАСТНИКОВ

Темой этой главы будут непрерывные детерминированные дифференциальные игры двух игроков. Игра происходит в интервале времени

$$t_0 \leqslant t \leqslant t_1, \tag{15.1.1}$$

где t_0 — фиксированное время начала, а t_1 — время окончания — либо задается, либо определяется самой игрой.

Игра происходит в некоторой системе, описываемой с помощью набора из n фазовых координат, представленных в виде *фазового вектора*. Фазовый вектор — это n-мерный вектор-столбец

$$\mathbf{x} = \mathbf{x}(t) = (x_1(t),\ x_2(t),\ \ldots,\ x_n(t))', \tag{15.1.2}$$

составляющие которого могут изменяться во времени начиная с заданного начального состояния

$$\mathbf{x}(t_0) = \mathbf{x}_0 \tag{15.1.3}$$

и кончая заданным конечным состоянием

$$\mathbf{x}(t_1) = \mathbf{x}_1. \tag{15.1.4}$$

Конечный момент времени определяется *конечной поверхностью*, т. е. поверхностью в E^{n+1}, описываемой уравнением

$$\mathbf{T}(\mathbf{x}_1,\ t_1) = 0. \tag{15.1.5}$$

Предполагается, что рассматриваемая игра — *игра с полной информацией*, т. е. игрокам известны значения всех текущих фазовых координат. Каждый игрок выбирает в течение игры значения своего вектора управляющих параметров, которые образуют траекторию управления.

Игрок 1 выбирает траекторию управления $\{\mathbf{u}^1(t)\}$

$$\{\mathbf{u}^1(t)\} = \{(u_1^1(t),\ u_2^1(t),\ \ldots \\ \ldots,\ u_{r_1}^1(t))' \mid t_0 \leqslant t \leqslant t_1\}, \tag{15.1.6}$$

а игрок 2 выбирает траекторию управления $\{\mathbf{u}^2(t)\}$

$$\{\mathbf{u}^2(t)\} = \{(u_1^2(t), u_2^2(t), \ldots \\ \ldots, u_{r_2}^2(t)' \mid t_0 \leqslant t \leqslant t_1\}. \qquad (15.1.7)$$

Эти траектории управления принадлежат заданным *множествам управлений*

$$\{\mathbf{u}^1(t)\} \in U^1 \\ \{\mathbf{u}^2(t)\} \in U^2. \qquad (15.1.8)$$

Это означает, что управления должны быть кусочно-непрерывными функциями времени, значения которых во всех точках указанного промежутка времени принадлежат непустым компактным множествам

$$\mathbf{u}^1(t) \in \Omega^1 \text{ при всех } t, \\ t_0 \leqslant t \leqslant t_1, \ \Omega^1 \subset E^{r_1} \\ \mathbf{u}^2(t) \in \Omega^2 \text{ при всех } t, \qquad (15.1.9) \\ t_0 \leqslant t \leqslant t_1, \ \Omega^2 \subset E^{r_2}.$$

Уравнения движения — это система дифференциальных уравнений

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t), \qquad (15.1.10)$$

причем предполагается, что $\mathbf{f}(\cdot\cdot\cdot\cdot)$ — это известная, непрерывно дифференцируемая функция. Уравнения движения (15.1.10) с начальным состоянием (15.1.3) и траекториями управления (15.1.6) и (15.1.7), выбранными игроками, определяют *фазовую траекторию* $\{\mathbf{x}(t)\}$

$$\{\mathbf{x}(t)\} = \{(x_1(t), x_2(t), \ldots \\ \ldots, x_n(t))' \mid t_0 \leqslant t \leqslant t_1\}. \qquad (15.1.11)$$

Выигрыши каждого игрока зависят от траекторий управления, выбранных обоими игроками. Выигрыш игрока 1 определяется как

$$J^1 = J^1[\{\mathbf{u}^1(t)\}, \{\mathbf{u}^2(t)\}] = \\ = \int_{t_0}^{t_1} I^1(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t)\, dt + \mathbf{F}^1(\mathbf{x}_1, t_1), \qquad (15.1.12)$$

а выигрыш игрока 2 — как

$$J^2 = J^2[\{\mathbf{u}^1(t)\}, \{\mathbf{u}^2(t)\}] = \\ = \int_{t_0}^{t_1} I^2(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t)\, dt + F^2(\mathbf{x}_1, t_1). \qquad (15.1.13)$$

Каждый игрок пытается максимизировать свой собственный выигрыш выбором траектории управления.

Стратегией каждого игрока называется правило для определения своего управляющего вектора в любой момент времени как функции фазовых координат в этот момент времени

$$\left.\begin{aligned} \mathbf{u}^1(t) &= \mathbf{S}^1(\mathbf{x}(t)) \\ \mathbf{u}^2(t) &= \mathbf{S}^2(\mathbf{x}(t)) \end{aligned}\right\} \text{ для всех } t, \ t_0 \leqslant t \leqslant t_1, \qquad (15.1.14)$$

причем смешанные стратегии не исключаются. Поскольку стратегия определяет выборы, производимые игроком в любой возможной ситуации, описываемой фазовым вектором, то используемое здесь понятие стратегии согласуется с определением стратегии в гл. 6. В терминах задачи управления стратегия представляет собой управление с обратной связью, обсуждавшееся в гл. 11. Так как каждый игрок знает только свою стратегию и получает информацию о противнике только в процессе игры, то он должен выбирать свой управляющий вектор по текущим фазовым координатам. Таким образом, по самому своему существу дифференциальным играм соответствует управление с обратной связью, а не управление по разомкнутому контуру.

Стремясь выбрать такую стратегию, чтобы максимизировать свой выигрыш, каждый игрок приходит к оптимальной стратегии $\mathbf{S}^{1*}(\mathbf{x})$ или $\mathbf{S}^{2*}(\mathbf{x})$ соответственно. Считая их известными, уравнения движения записываем следующим образом:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{S}^{1*}(\mathbf{x}), \mathbf{S}^{2*}(\mathbf{x}), t). \qquad (15.1.15)$$

Чтобы получить фазовую траекторию $\{\mathbf{x}(t)\}$ и величины выигрышей каждого игрока, можно, начиная с заданного начального состояния, проинтегрировать эти уравнения. В результате получим

$$J^{1*} = J^1[\{\mathbf{S}^{1*}(\mathbf{x})\}, \{\mathbf{S}^{2*}(\mathbf{x})\}] \\ J^{2*} = J^2[\{\mathbf{S}^{1*}(\mathbf{x})\}, \{\mathbf{S}^{2*}(\mathbf{x})\}]. \qquad (15.1.16)$$

15.2. ДИФФЕРЕНЦИАЛЬНЫЕ ИГРЫ ДВУХ УЧАСТНИКОВ С НУЛЕВОЙ СУММОЙ

В дифференциальной игре двух участников с нулевой суммой выигрыш игрока 2 равен выигрышу игрока 1, взятому со знаком минус. Пусть J — выигрыш игрока 1 — равен

$$J[\{\mathbf{u}^1(t)\}, \{\mathbf{u}^2(t)\}] = \int_{t_0}^{t_1} I(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t)\,dt + F(\mathbf{x}_1, t_1). \qquad (15.2.1)$$

Игрок 1 пытается максимизировать J выбором $\{\mathbf{u}^1(t)\}$, а игрок 2 — минимизировать J выбором $\{\mathbf{u}^2(t)\}$. Задача, таким образом, заключается в отыскании таких стратегий

$$\mathbf{u}^{1*}(t) = \mathbf{S}^{1*}(\mathbf{x}(t))$$
$$\mathbf{u}^{2*}(t) = \mathbf{S}^{2*}(\mathbf{x}(t)), \qquad (15.2.2)$$

при которых $\{\mathbf{u}^{1*}(t)\} \in U^1$ и $\{\mathbf{u}^{2*}(t)\} \in U^2$ образуют седловую точку платежного функционала

$$J[\{\mathbf{u}^1(t)\}, \{\mathbf{u}^{2*}(t)\}] \leqslant J[\{\mathbf{u}^{1*}(t)\}, \{\mathbf{u}^{2*}(t)\}] \leqslant J[\{\mathbf{u}^{1*}(t)\}, \{\mathbf{u}^2(t)\}]$$

для всех $\{\mathbf{u}^1(t)\} \in U^1$, $\{\mathbf{u}^2(t)\} \in U^2$. (15.2.3)

Величина $J[\{\mathbf{u}^{1*}(t)\}, \{\mathbf{u}^{2*}(t)\}]$ называется *ценой* дифференциальной игры. Необходимые условия для того, чтобы управления могли образовывать седловую точку данного функционала, можно вывести с помощью принципа максимума по аналогии с прежним выводом условий оптимальности управления [1].

Введем n-мерный вектор-строку сопряженных переменных

$$\mathbf{y}(t) = (y_1(t), y_2(t), \ldots, y_n(t)) \qquad (15.2.4)$$

[1] Доказательства аналогичны приведенным в предыдущих трех главах. Доказательство, использующее вариационное исчисление, см. в работе Берковича [7], доказательства, использующие динамическое программирование, рассматриваются в книгах Айзекса [1] и Берковича [8]; доказательство, использующее принцип максимума, см. в [9]. В этих доказательствах обычно предполагается, что для обоих игроков существуют оптимальные стратегии и что игра имеет конечную цену. О существовании решения см. в [10].

и определим гамильтониан

$$H(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, \mathbf{y}, t) = I(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t) + \mathbf{y}\mathbf{f}(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t). \qquad (15.2.5)$$

Необходимые условия для оптимальных стратегий двух игроков состоят в том, что в любой момент из заданного интервала времени игрок 1 максимизирует гамильтониан, выбирая свой управляющий вектор, а игрок 2 минимизирует гамильтониан, выбирая свой управляющий вектор. Если дифференциальная игра удовлетворяет некоторым условиям регулярности и *вполне определена*, т. е. решением будут чистые, а не смешанные стратегии, то необходимое условие существования решения состоит в том, что в любой момент из заданного интервала времени функция Гамильтона имеет седловую точку [1]

$$H(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^{2*}, \mathbf{y}, t) \leqslant H(\mathbf{x}, \mathbf{u}^{1*}, \mathbf{u}^{2*}, \mathbf{y}, t) \leqslant H(\mathbf{x}, \mathbf{u}^{1*}, \mathbf{u}^2, \mathbf{y}, t) \qquad (15.2.6)$$

для всех $\mathbf{u}^1 \in \Omega^1$, $\mathbf{u}^2 \in \Omega^2$, t, $t_0 \leqslant t \leqslant t_1$.

Иначе говоря,

$$\max_{\mathbf{u}^1 \in \Omega^1} \min_{\mathbf{u}^2 \in \Omega^2} H(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, \mathbf{y}, t) = \min_{\mathbf{u}^2 \in \Omega^2} \max_{\mathbf{u}^1 \in \Omega^1} H(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, \mathbf{y}, t) = H(\mathbf{x}, \mathbf{u}^{1*}, \mathbf{u}^{2*}, \mathbf{y}, t). \qquad (15.2.7)$$

Таким образом, в соответствии с этим результатом, который по аналогии с принципом максимума может быть назван *принципом минимаксимума*, дифференциальная игра двух игроков с нулевой суммой в каждый момент заданного

[1] Из гл. 6 известно, что игры с полной информацией всегда являются вполне определенными, если только это конечные игры. В отличие от игр с полной информацией дифференциальные игры бесконечны, поэтому могут оказаться необходимыми смешанные стратегии, т. е. будет использоваться набор возможных альтернативных стратегий из множества траекторий управлений, взятых с некоторыми вероятностями. Примеры не полностью определенных дифференциальных игр, решениями которых являются смешанные стратегии, см. в работах Берковича [4] и Оуэна [5]. Однако если как подынтегральная функция $I(\cdots)$, так и функция $\mathbf{f}(\cdots)$, входящая в уравнения движения, сепарабельны в том смысле, что гамильтониан можно представить в виде суммы двух функций, одна из которых зависит только от $\mathbf{u}^1$, а другая только от $\mathbf{u}^2$, то дифференциальная игра является вполне определенной и имеет решением чистую стратегию. Примером служит случай, когда и $I(\cdots)$ и $\mathbf{f}(\cdots)$ линейны. Он рассмотрен в книге Понтрягина и др. [9].

интервала времени должна быть вполне определенной (статической) игрой. Другими словами, эта статическая игра двух участников с нулевой суммой имеет седловую точку. Остальные необходимые условия — это канонические уравнения, и граничные условия такие же, как и в принципе максимума

$$\dot{\mathbf{x}} = \frac{\partial H}{\partial \mathbf{y}}, \quad \mathbf{x}(t_0) = \mathbf{x}_0$$
$$\dot{\mathbf{y}} = -\frac{\partial H}{\partial \mathbf{x}}, \quad \mathbf{y}(t_1) = \frac{\partial F}{\partial \mathbf{x}_1} + \mathbf{v}\frac{\partial T}{\partial \mathbf{x}_1}. \quad (15.2.8)$$

Здесь $\mathbf{v}$ — вектор-строка множителей Лагранжа. Преобразуя полученную систему, приходим к условию трансверсальности

$$\left(H + \frac{\partial F}{\partial t_1}\right) + \left(\frac{\partial F}{\partial \mathbf{x}_1} - \mathbf{y}\right)\left(\frac{d\mathbf{x}}{dt}\right)_{\mathbf{T}(\cdot\cdot)=0} = 0. \quad (15.2.9)$$

Значения переменных и производных здесь берутся в конечный момент времени t_1.

Если задача автономна, т. е. функции $I(\cdot\cdot\cdot\cdot)$ и $\mathbf{f}(\cdot\cdot\cdot\cdot)$ не зависят явно от времени, то минимаксное значение гамильтониана является константой, которую можно положить равной нулю. Итак, в данном случае [1]

$$\max_{\mathbf{u}^1\in\Omega^1}\min_{\mathbf{u}^2\in\Omega^2} I(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t) + \mathbf{y}\mathbf{f}(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t) = 0. \quad (15.2.10)$$

Примером дифференциальной игры двух участников с нулевой суммой, которую можно решить с помощью принципа минимаксимума, является игра с квадратичным функционалом и линейными автономными уравнениями движения.

[1] Айзекс [1] заменяет вектор y представлением его в виде $\partial J^*/\partial \mathbf{x}$ (см. об этом в разделе 14.4) и называет уравнение

$$\max_{\mathbf{u}^1\in\Omega^1}\min_{\mathbf{u}^2\in\Omega^2}\left[I(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t) + \frac{\partial J^*}{\partial \mathbf{x}} f(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t)\right] = 0$$

главным уравнением. Это уравнение представляет собой не что иное, как уравнение Беллмана для данной задачи. Айзекс [1] записывает канонические уравнения как уравнения траекторий, взятых в обратном направлении:

$$\mathring{\mathbf{x}} = -\frac{\partial H}{\partial \mathbf{y}}, \quad \mathring{\mathbf{y}} = \frac{\partial H}{\partial \mathbf{x}}.$$

Здесь кружок сверху обозначает, что дифференцирование производится по времени, остающемуся до окончания, т. е.

$$\mathring{z} = \frac{dz}{d\tau}, \text{ где } \tau = t_1 - t \text{ — остающееся время.}$$

Аналогом этой задачи в теории управления является задача управления по минимуму энергии, разобранная в разделе 14.5. В этой дифференциальной игре фазовый вектор можно разложить на две части

$$\mathbf{x} = \begin{pmatrix}\mathbf{x}^1\\ \mathbf{x}^2\end{pmatrix}, \quad (15.2.11)$$

где $\mathbf{x}^1$ включает фазовые координаты игрока 1, а $\mathbf{x}^2$ — фазовые координаты игрока 2. Уравнения движения линейны и автономны

$$\dot{\mathbf{x}}^1 = \mathbf{A}^1\mathbf{x}^1 + \mathbf{B}^1\mathbf{u}^1$$
$$\dot{\mathbf{x}}^2 = \mathbf{A}^2\mathbf{x}^2 + \mathbf{B}^2\mathbf{u}^2. \quad (15.2.12)$$

Здесь $\mathbf{u}^1$ и $\mathbf{u}^2$ — это управляющие векторы игроков 1 и 2 соответственно, причем никаких ограничений на $\mathbf{u}^1$ и $\mathbf{u}^2$ не накладывается. Время окончания t_1 задано, а выигрыш игрока 1 определяется как

$$J = \int_{t_0}^{t_1} [\mathbf{x}^{1\prime}\mathbf{C}^1\mathbf{x}^1 + \mathbf{x}^{2\prime}\mathbf{C}^2\mathbf{x}^2 + \mathbf{u}^{1\prime}\mathbf{D}^1\mathbf{x}^1 +$$
$$+ \mathbf{u}^{2\prime}\mathbf{D}^2\mathbf{u}^2 + \mathbf{u}^{1\prime}\mathbf{D}^3\mathbf{u}^2]\, dt +$$
$$+ [\mathbf{x}_1^{1\prime}E^1\mathbf{x}_1^1 + \mathbf{x}^{2\prime}_1\mathbf{E}^2\mathbf{x}_1^2 + \mathbf{x}_1^{1\prime}\mathbf{E}^3\mathbf{x}_1^2], \quad (15.2.13)$$

где $\mathbf{D}^1$ — отрицательно определенная матрица, а $\mathbf{D}^2$ — положительно определенная матрица. Гамильтониан в этом случае имеет вид

$$H = [\mathbf{x}^{1\prime}\mathbf{C}^1\mathbf{x}^1 + \mathbf{x}^{2\prime}\mathbf{C}^2\mathbf{x}^2 + \mathbf{u}^{1\prime}\mathbf{D}^1\mathbf{u}^1 +$$
$$+ \mathbf{u}^{2\prime}\mathbf{D}^2\mathbf{u}^2 + \mathbf{u}^{1\prime}\mathbf{D}^3\mathbf{u}^2] + \mathbf{y}^1[\mathbf{A}^1\mathbf{x}^1 + \mathbf{B}^1\mathbf{u}^1] +$$
$$+ \mathbf{y}^2[\mathbf{A}^2\mathbf{x}^2 + \mathbf{B}^2\mathbf{u}^2]. \quad (15.2.14)$$

Здесь вектор сопряженных переменных представлен в виде

$$\mathbf{y} = (\mathbf{y}^1\mathbf{y}^2). \quad (15.2.15)$$

По принципу минимаксимума получаем необходимые условия оптимальности

$$\frac{\partial H}{\partial \mathbf{u}^1} = 2\mathbf{u}^{1\prime}\mathbf{D}^1 + \mathbf{u}^{2\prime}\mathbf{D}^3 + \mathbf{y}^1\mathbf{B}^1 = \mathbf{0}$$
$$\frac{\partial H}{\partial \mathbf{u}^2} = 2\mathbf{u}^{2\prime}\mathbf{D}^2 + \mathbf{u}^{1\prime}\mathbf{D}^3 + \mathbf{y}^2\mathbf{B}^2 = \mathbf{0}. \quad (15.2.16)$$

Условия второго порядка выполняются, так как, по предположению, $\mathbf{D}^1$ — отрицательно определенная матрица, а $\mathbf{D}^2$ — положительно определенная матрица. Оптимальный вектор управления, выраженный через вектор сопряженных переменных, имеет вид

$$
\begin{aligned}
\mathbf{u}^1 &= -\frac{1}{2}\,\mathbf{D}^{1^{-1}}\{\mathbf{D}^3\mathbf{u}^2+\mathbf{B}^{1\prime}\mathbf{y}^{1\prime}\}\\
\mathbf{u}^2 &= -\frac{1}{2}\,\mathbf{D}^{2^{-1}}\{\mathbf{D}^3\mathbf{u}^1+\mathbf{B}^{2\prime}\mathbf{y}^{2\prime}\},
\end{aligned}
\qquad (15.2.17)
$$

а дифференциальные уравнения для сопряженных переменных имеют вид

$$
\begin{aligned}
\dot{\mathbf{y}}^1 &= -\frac{\partial H}{\partial \mathbf{x}^1} = -2\mathbf{x}^{1\prime}\mathbf{C}^1-\mathbf{y}^1\mathbf{A}^1\\
\dot{\mathbf{y}}^2 &= -\frac{\partial H}{\partial \mathbf{x}^2} = -2\mathbf{x}^{2\prime}\mathbf{C}^2-\mathbf{y}^2\mathbf{A}^2.
\end{aligned}
\qquad (15.2.18)
$$

Тогда, предположив, что решение линейно и что его можно представить в виде

$$
\begin{aligned}
\mathbf{y}^1 &= \mathbf{x}^{1\prime}\mathbf{Q}^1(t)\\
\mathbf{y} &= \mathbf{x}^{2\prime}\mathbf{Q}^2(t),
\end{aligned}
\qquad (15.2.19)
$$

получаем, как и в разделе 14.5, матричные уравнения Рикатти для $\mathbf{Q}^1(t)$ и $\mathbf{Q}^2(t)$. Итак, найдены оптимальные управления с обратной связью

$$
\begin{aligned}
\mathbf{u}^1 &= -\frac{1}{2}\,\mathbf{D}^{1^{-1}}\{\mathbf{D}^3\mathbf{u}^2+\mathbf{B}^{1\prime}\mathbf{Q}^{1\prime}\mathbf{x}^1\}\\
\mathbf{u}^2 &= -\frac{1}{2}\,\mathbf{D}^{2^{-1}}\{\mathbf{D}^{3\prime}\mathbf{u}^1+\mathbf{B}^{2\prime}\mathbf{Q}^{2\prime}\mathbf{x}^2\}.
\end{aligned}
\qquad (15.2.20)
$$

Отсюда видно, что оптимальное управление для каждого участника — это линейные функции своих собственных фазовых координат и управляющих параметров другого участника. *Точка равновесия* достигается тогда, когда при выборе игроком 1 управления $\mathbf{u}^1$ на основе управления $\mathbf{u}^2$ игрока 2 оптимальным выбором для последнего является именно $\mathbf{u}^2$, что в свою очередь принуждает игрока 1 повторить свое решение. Точка равновесия отыскивается в результате решения системы уравнений (15.2.9) относительно $\mathbf{u}^1$ и $\mathbf{u}^2$

$$
\begin{aligned}
\mathbf{u}^1 = {} & \left[I-\frac{1}{4}\,\mathbf{D}^{1^{-1}}\mathbf{D}^3\mathbf{D}^{2^{-1}}\mathbf{D}^{3\prime}\right]^{-1}\cdot\left[-\frac{1}{2}\,\mathbf{D}^{1^{-1}}\mathbf{B}^{1\prime}\mathbf{Q}^{1\prime}(t)\,\mathbf{x}^1+\right.\\
& \left.+\frac{1}{4}\,\mathbf{D}^{1^{-1}}\mathbf{D}^3\mathbf{D}^{2^{-1}}\mathbf{B}^{2\prime}\mathbf{Q}^{2\prime}(t)\,\mathbf{x}^2\right]\\
\mathbf{u}^2 = {} & \left[I-\frac{1}{4}\,\mathbf{D}^{2^{-1}}\mathbf{D}^{3\prime}\mathbf{D}^{1^{-1}}\mathbf{D}^3\right]^{-1}\times\\
& \times\left[\frac{1}{4}\,\mathbf{D}^{2^{-1}}\mathbf{D}^{3\prime}\mathbf{D}^{1^{-1}}\mathbf{B}^{1\prime}\mathbf{Q}^{1\prime}(t)\,\mathbf{x}^1-\frac{1}{2}\,\mathbf{D}^{2^{-1}}\mathbf{B}^{2\prime}\mathbf{Q}^{2\prime}(t)\,\mathbf{x}^2\right].
\end{aligned}
\qquad (15.2.21)
$$

Причем предполагается, что обратные матрицы существуют. При этом оптимальные управляющие векторы каждого участника представляют собой линейные функции фазовых векторов обоих участников, т. е. каждый участник при оптимальном поведении пользуется *правилом линейного решения*, выбирая линейную форму связи линейных управляющих параметров и фазовых координат [1].

15.3. ИГРЫ ПРЕСЛЕДОВАНИЯ

В теоретическом плане и с точки зрения практических приложений наиболее важным классом дифференциальных игр двух участников с нулевой суммой являются *игры преследования*. В них игрок 1 является *преследователем*, а игрок 2 — *преследуемым* [9, 11, 12, 1, 3]. Игра заканчивается, когда преследователь «захватил» преследуемого. Продолжительность игры равна времени, нужному для «захвата». Цель преследователя — минимизировать «время захвата», в то время как цель преследуемого — максимизировать его. Если преследователь не может подойти достаточно близко к преследуемому для захвата последнего, то преследуемый «спасен» и время захвата бесконечно. Такое определение игры преследования достаточно широкое, и оно позволяет отнести к играм такого типа множество примеров преследования, включая такие различные ситуации, как преследование игрока в футболе или преследование ракеты антиракетой.

Простейшей игрой преследования является *преследование на плоскости*, когда участники управляют двумя

---

[1] Ср. с решением соответствующей задачи управления (раздел 14.5). — *Прим. перев.*

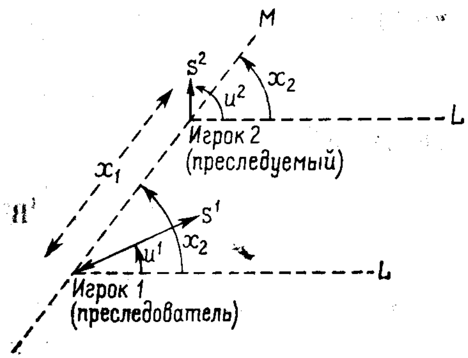

Рис. 15.1. Задача преследования на плоскости.

точками на плоскости, двигающимися с фиксированными скоростями, причем скорость преследователя превосходит скорость преследуемого. Здесь управляющими параметрами являются направления движения точек. Фазовые координаты и управляющие параметры показаны на рис. 15.1. Прямая L — ось отсчета, а прямая M в любое время проходит через эти две точки. В качестве фазовых координат в движущейся системе отсчета выбраны

x_1 — расстояние от игрока 1 до игрока 2
x_2 — угол между L и M. (15.3.1)

Характеристики направления движения

u^1 — угол между вектором скорости участника 1 и L
u^2 — угол между вектором скорости участника 2 и L (15.3.2)

являются управляющими параметрами.

Игрок 1 (преследователь) движется со скоростью s^1, игрок 2 (преследуемый) движется со скоростью s^2 ($s^1 > s^2$), и

$$0 \leqslant u^1 < 2\pi, \qquad 0 \leqslant u^2 < 2\pi. \tag{15.3.3}$$

Уравнения движения имеют вид

$$\dot{x}_1 = -s^1 \cos(u^1 - x_2) + s^2 \cos(u^2 - x_2), \qquad \dot{x}_2 = \frac{-s^1 \sin(u^1 - x_2) + s^2 \sin(u^2 - x_2)}{x_1}. \tag{15.3.4}$$

Отметим, что если преследователь движется прямо за преследуемым, а преследуемый удаляется от преследователя по прямой линии, то

$$u^1 = x_2, \qquad u^2 = x_2 \tag{15.3.5}$$

и уравнения движения принимают вид

$$\dot{x}_1 = s^2 - s^1, \qquad \dot{x}_2 = 0. \tag{15.3.6}$$

Первое уравнение показывает, что расстояние между преследователем и преследуемым уменьшается со скоростью, равной разнице их скоростей.

Конечный момент времени определяется временем, когда расстояние между участниками сократится до заданной величины l, т. е.

$$x_1(t_1) = l. \tag{15.3.7}$$

В момент t_1 преследователь «захватывает» преследуемого. Выигрыш преследователя (игрок 1) равен

$$J = -\int_{t_0}^{t_1} dt = -(t_1 - t_0).$$

В этой задаче гамильтониан имеет следующий вид:

$$H = -1 + y_1(-s^1 \cos(u^1 - x_2) + s^2 \cos(u^2 - x_2)) + \frac{y_2}{x_1}(-s^1 \sin(u^1 - x_2) + s^2 \sin(u^2 - x_2)). \tag{15.3.8}$$

Согласно принципу минимаксимума, нужно найти максимум функции Гамильтона по u^1 и минимум по u^2. Из условий первого порядка

$$\frac{\partial H}{\partial u^1} = y_1 s^1 \sin(u^1 - x_2) - \frac{y_2}{x_1} s^1 \cos(u^1 - x_2) = 0, \qquad \frac{\partial H}{\partial u^2} = -y_1 s^2 \sin(u^2 - x_2) + \frac{y_2}{x_1} s^2 \cos(u^2 - x_2) = 0 \tag{15.3.9}$$

следует, что

$$\operatorname{tg}(u^1 - x_2) = \operatorname{tg}(u^2 - x_2) = \frac{y_2}{y_1 x_1}. \tag{15.3.10}$$

Запишем дифференциальные уравнения для сопряженных переменных

$$\dot{y}_1 = -\frac{\partial H}{\partial x_1} = -\frac{y_2}{x_1^2}\left(-s^1 \sin(u^1 - x_2) + s^2 \sin(u^2 - x_2)\right)$$

$$\dot{y}_2 = -\frac{\partial H}{\partial x_2} = y_1\left(-s^1 \sin(u^1 - x_2) + s^2 \sin(u^2 - x_2)\right) + \frac{y_2}{x_1}\left(s^1 \cos(u^1 - x_2) - s^2 \cos(u^2 - x_2)\right). \qquad (15.3.11)$$

Но, согласно (15.3.10),

$$\sin(u^1 - x_2) = \frac{y_2}{y_1 x_1}\cos(u^1 - x_2)$$
$$\sin(u^2 - x_2) = \frac{y_2}{y_1 x_1}\cos(u^2 - x_2). \qquad (15.3.12)$$

Отсюда следует, что

$$\dot{y}_2 = 0, \qquad (15.3.13)$$

т. е. y_2 есть величина, постоянная во времени. Поскольку никаких ограничений на конечное значение x_2 нет, то

$$y_2(t_1) = 0 \qquad (15.3.14)$$

и

$$y_2(t) = 0, \quad t_0 \leqslant t \leqslant t_1. \qquad (15.3.15)$$

Итак, цена игры не зависит от величины начального угла $x_2(t)$, так как сопряженная переменная, являющаяся характеристикой чувствительности, равна нулю

$$y_2(t_0) = \frac{\partial J^*}{\partial x_2(t_0)} = 0.$$

Из (15.3.10) видно, что u^1 и u^2 являются решениями, если

$$\sin(u^1 - x_2) = \sin(u^2 - x_2) = 0$$
$$\operatorname{tg}(u^1 - x_2) = \operatorname{tg}(u^2 - x_2) = 0. \qquad (15.3.16)$$

Следовательно, оптимальная траектория управления должна удовлетворять условиям

$$u^1 = x_2$$
$$u^2 = x_2. \qquad (15.3.17)$$

Это, как отмечалось раньше, соответствует случаю, когда преследователь настигает преследуемого по прямой, а преследуемый удаляется от преследователя также по прямой.

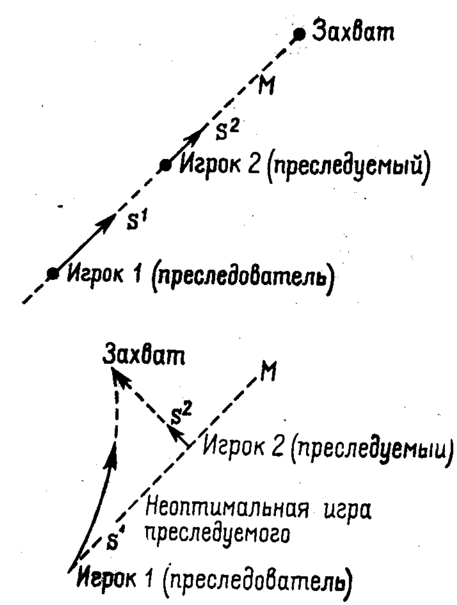

Рис. 15.2. Оптимальная игра преследования и неоптимальная игра преследования на плоскости.

В этом случае скорость уменьшения расстояния между ними равна

$$\dot{x}_1 = s^2 - s^1. \qquad (15.3.18)$$

Следовательно,

$$x_1(t) = (s^1 - s^2)(t_0 - t) + x_1(t_0), \qquad (15.3.19)$$

где $x_1(t_0)$ — заданное первоначальное расстояние между участниками. Так как по определению t_1

$$x_1(t_1) = (s^1 - s^2)(t_0 - t) = x_1(t_0) = l, \qquad (15.3.20)$$

то цена игры для игрока 1 (преследователя) равна

$$J^* = -(t_1 - t_0) = -\left(\frac{x_1(t_0) - l}{s^1 - s^2}\right). \qquad (15.3.21)$$

На рис. 15.2 показаны оптимальная и неоптимальная игры преследования на плоскости. Верхний чертеж демонстрирует оптимальное поведение: преследователь движется прямо в направлении преследуемого, а преследуе-

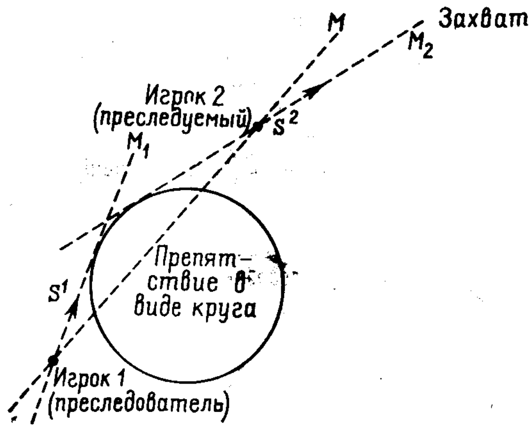

Рис. 15.3. Игра преследования на плоскости. Преследователь и преследуемый разделены препятствием в виде круга.

мый движется от преследователя по прямой M, соединяющей обоих игроков. На нижнем чертеже показано неоптимальное поведение, когда преследуемый движется под прямым углом к линии M. Преследователь выбирает оптимальное поведение: движется все время в направлении преследуемого ($u^1 = x^2$) и «захватывает» его за меньшее время.

Возможны различные усложенения игры преследования на плоскости. Как одно из таких усложнений рассмотрим случай, когда между преследователем и преследуемым лежит препятствие, например, в виде круга, как на рис. 15.3. Оптимальной стратегией для игрока 2 будет движение по прямой M_2, касательной к окружности и проходящей через начальное положение этого игрока. Оптимальной стратегией игрока 1 будет движение по прямой M_1, касательной к окружности и проходящей через его начальное положение, потом по окружности и, наконец, по прямой M_2, на которой и происходит захват. Оптимальное поведение для каждого игрока показано на рис. 15.3. Никакие другие стратегии игрока 1 не могут сократить время захвата, и никакие другие стратегии участника 2 не приводят к продлению времени захвата.

Если прямая M, связывающая начальные положения участников, проходит через центр окружности, то в распоряжении каждого игрока имеются две касательные, одинаково выгодные для движения. В этом случае игроки могут использовать смешанные стратегии, выбирая путь случайным образом так, что каждый путь может быть выбран с вероятностью $^1/_2$. Множество таких симметричных положений называется *рассеивающей поверхностью.* Эта поверхность исчезает, как только выбор сделан, при этом один или оба участника могут изменить свое решение. Если оба изменяют свои решения, тогда они могут оказаться на другой рассеивающей поверхности [1].

15.4. КООРДИНИРОВАННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ ИГРЫ

В игре с нулевой суммой участники находятся в состоянии прямого конфликта, так как выигрыш одного участника равен проигрышу другого. В *координированных играх*, напротив, игроки действуют в полном согласии друг с другом, стремясь максимизировать одинаковые выигрыши, равные

$$J[\{\mathbf{u}^1(t)\},\ \{\mathbf{u}^2(t)\}] = \int_{t_0}^{t_1} I(\mathbf{x},\ \mathbf{u}^1,\ \mathbf{u}^2,\ t)\,dt + F(\mathbf{x}_1,\ t_1) \quad (15.4.1)$$

за счет выбора своих управляющих траекторий $\{\mathbf{u}(t)\}$ и $\{\mathbf{u}^2(t)\}$ соответственно. Иллюстрацией подобной игры служит задача о совместном стремлении избежать столкновения двух движущихся машин (например, автомобилей, кораблей, самолетов), причем выигрыш равен либо 0, либо 1 в зависимости от того, будет ли расстояние между ними в момент наибольшего сближения меньше или больше некоторого критического.

Решение кооперативной дифференциальной игры двух участников можно получить аналогично решению задачи управления с помощью принципа максимума. В нашем случае при предположении, что дифференциальная игра удовлетворяет некоторым условиям регулярности, оптимальное управление обязательно удовлетворяет следующему условию, налагаемому на функцию Гамильтона:

$$\max_{\mathbf{u}^1 \in \Omega^1} \max_{\mathbf{u}^2 \in \Omega^2} H(\mathbf{x},\ \mathbf{u}^1,\ \mathbf{u}^2,\ \mathbf{y},\ t) = H(\mathbf{x},\ \mathbf{u}^{1*},\ \mathbf{u}^{2*},\ \mathbf{y},\ t) \quad (15.4.2)$$

во всех точках указанного интервала времени. Это условие можно назвать *принципом максимаксимума*. Канонические уравнения те же самые, что и в предыдущем параграфе.

В качестве примера кооперативной дифференциальной игры двух участников рассмотрим игру, в которой двое участников управляют ускорением единичной массы с координатами $(x_1\ x_2)'$ в одном и том же направлении. Дифференциальные уравнения имеют следующий вид:

$$\begin{aligned} \ddot{x}_1 &= u^1 \\ \ddot{x}_2 &= u^2. \end{aligned} \tag{15.4.3}$$

Ограничения на управляющие параметры

$$\begin{aligned} |u^1| &\leqslant 1 \\ |u^2| &\leqslant 1 \end{aligned} \tag{15.4.4}$$

показывают, что максимальное ускорение для каждого игрока не превосходит единицы. Цель игры — достичь начала координат за минимальное время, т. е.

$$J = -\int_{t_0}^{t_1} dt = -(t_1 - t_0). \tag{15.4.5}$$

В заданном начальном положении масса находится в покое

$$\begin{aligned} x_1(t_0) &= x_{10} \\ x_2(t_0) &= x_{20} \\ \dot{x}_1(t_0) &= 0 \\ \dot{x}_2(t_0) &= 0. \end{aligned} \tag{15.4.6}$$

Эта координированная игра представляет собой дифференциальную игру, аналогом которой в теории управления является задача о минимальном времени перехода, где в качестве управляющего параметра берется вторая производная фазовой координаты (см. раздел 14.5). Воспользовавшись приемом этого раздела, можно наши дифференциальные уравнения свести к уравнениям первого порядка. Введем новые фазовые переменные x_3 и x_4

$$\begin{aligned} \dot{x}_1 &= x_3, & x_1(t_0) &= x_{10} \\ \dot{x}_2 &= x_4, & x_2(t_0) &= x_{20} \\ \dot{x}_3 &= u^1, & x_3(t_0) &= 0 \\ \dot{x}_4 &= u^2, & x_4(t_0) &= 0. \end{aligned} \tag{15.4.7}$$

Тогда гамильтониан принимает вид

$$H = -1 + y_1x_3 + y_2x_4 + y_3u^1 + y_4u^2, \tag{15.4.8}$$

и по принципу максимаксимума

$$\begin{aligned} u^1 &= \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}, \quad \text{если} \quad y_3 \begin{Bmatrix} > \\ < \end{Bmatrix} 0 \\ u^2 &= \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}, \quad \text{если} \quad y_4 \begin{Bmatrix} > \\ < \end{Bmatrix} 0. \end{aligned} \tag{15.4.9}$$

Канонические уравнения для сопряженных переменных выглядят следующим образом:

$$\begin{aligned} \dot{y}_1 &= -\frac{\partial H}{\partial x_1} = 0 \\ \dot{y}_2 &= -\frac{\partial H}{\partial x_2} = 0 \\ \dot{y}_3 &= -\frac{\partial H}{\partial x_3} = -y_1 \\ \dot{y}_4 &= -\frac{\partial H}{\partial x_4} = -y_2. \end{aligned} \tag{15.4.10}$$

Их решениями являются

$$\begin{aligned} y_1 &= c_1 \\ y_2 &= c_2 \\ y_3 &= c_3 - c_1t \\ y_4 &= c_4 - c_2t, \end{aligned} \tag{15.4.11}$$

где c_1, c_2, c_3 и c_4 — постоянные. Поскольку на значения конечных скоростей не наложено ограничений, то

$$\begin{aligned} y_3(t_1) &= 0 \\ y_4(t_1) &= 0. \end{aligned} \tag{15.4.12}$$

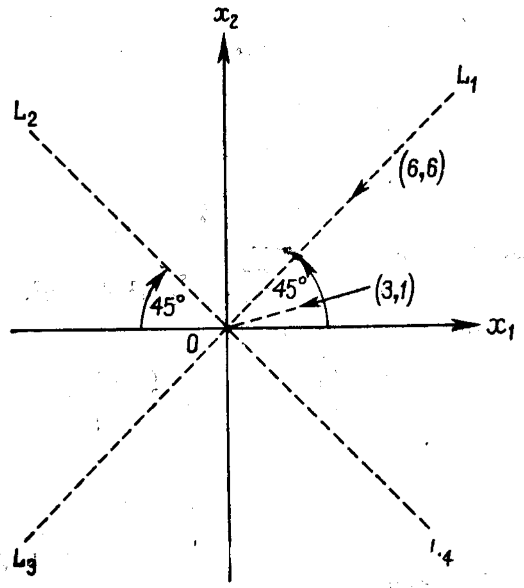

Рис. 15.4. Кооперативная дифференциальная игра.

Из конечных условий и вида решения для сопряженных переменных видно, что y_3 и y_4 не могут менять знак — они или всегда положительны, или всегда отрицательны, или равны нулю. Решение задачи показано на рис. 15.4. Решение, начинающееся в некоторой точке на прямой OL_1, например, если начальная точка — это точка (6.6), является очевидным

$$u^1 = -1, \quad u^2 = -1, \tag{15.4.13}$$

при этом y_3 и y_4 отрицательны. Так как начальные значения сопряженных переменных можно интерпретировать как характеристики чувствительности, то

$$\begin{aligned} y_3(t_0) &= \frac{\partial J^*}{\partial x_{30}} \\ y_4(t_0) &= \frac{\partial J^*}{\partial x_{40}}. \end{aligned} \tag{15.4.14}$$

Отрицательные начальные значения y_3 и y_4 показывают в данном примере, что при увеличении положительных начальных скоростей в любых направлениях, началом

которых служат точки на прямой OL, увеличилось бы время достижения начала координат.

Аналогично получается решение, начинающееся из какой-либо точки на прямой OL_2

$$u^1 = 1, \quad u^2 = -1. \tag{15.4.15}$$

Точно так же решением, начальная точка которого лежит на прямой OL_3, является

$$u^1 = 1, \quad u^2 = 1, \tag{15.4.16}$$

а если начальная точка лежит на прямой OL_4, то

$$u^1 = -1, \quad u^2 = 1. \tag{15.4.17}$$

Посмотрим, что получится для точек, не лежащих ни на одной из этих прямых, например таких, как точка (3.1). Решение по-прежнему проходит по прямой и

$$u^1 = -1, \quad u^2 = -\frac{1}{3}. \tag{15.4.18}$$

Это решение удовлетворяет указанным ранее необходимым условиям, хотя u^2 и не находится на границе. В этом случае

$$\begin{aligned} y_2 &= 0 \\ y_4 &= 0. \end{aligned} \tag{15.4.19}$$

Отсюда, поскольку сопряженные переменные измеряют чувствительность, следует, что целевой функционал (минимум времени) не зависит от начального положения и от скорости в вертикальном направлении. Это очевидно, так как в данном случае единственное влияющее на время направление — горизонтальное.

Начиная с более высокой точки или с большей скоростью в вертикальном направлении, мы просто изменяем значение u^2 без изменения J^*. По приведенным соображениям решение всегда проходит по прямой, а оптимальный выигрыш и минимальное время зависят только от большей из начальных координат.

15.5. НЕКООПЕРАТИВНЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ ИГРЫ

Некооперативная дифференциальная игра — это дифференциальная игра с ненулевой суммой, в которой участники не могут принимать связывающие их обязательства об используемых ими стратегиях до начала игры. В дифференциальной игре двух участников с ненулевой суммой, где выигрыши участника 1 и участника 2 определяются соответственно как

$$J^1[\{\mathbf{u}^1(t)\}, \{\mathbf{u}^2(t)\}] = \int_{t_0}^{t_1} I^1(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t)\,dt + F^1(\mathbf{x}_1 t_1)$$
$$J^2[\{\mathbf{u}^1(t)\}, \{\mathbf{u}^2(t)\}] = \int_{t_0}^{t_1} I^2(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t)\,dt + F^2(\mathbf{x}_1 t_1), \tag{15.5.1}$$

некооперативным равновесием (по Нэшу) называется пара стратегий

$$\mathbf{u}^{1*}(t) = \mathbf{S}^{1*}(\mathbf{x}(t))$$
$$\mathbf{u}^{2*}(t) = \mathbf{S}^{2*}(\mathbf{x}(t)), \tag{15.5.2}$$

таких, что ни один из игроков не имеет намерения изменить свою стратегию при заданной стратегии противника. Следовательно,

$$J^1[\{\mathbf{u}^{1*}(t)\}, \{\mathbf{u}^{2*}(t)\}] \geqslant J[\{\mathbf{u}^1(t)\}, \{\mathbf{u}^{2*}(t)\}]$$
$$\text{для всех } \{\mathbf{u}^1(t)\} \in U^1 \tag{15.5.3}$$
$$J^2[\{\mathbf{u}^{1*}(t)\}, \{\mathbf{u}^{2*}(t)\}] \geqslant J[\{\mathbf{u}^{1*}(t)\}, \{\mathbf{u}^2(t)\}]$$
$$\text{для всех } \{\mathbf{u}^2(t)\} \in U^2.$$

Если выполнены некоторые условия регулярности, то можно получить необходимые условия некооперативного равновесия (по Нэшу) с помощью понятия функции Гамильтона, действуя по аналогии с методом принципа максимума [13].

Гамильтонианы для участника 1 и участника 2 соответственно имеют вид

$$H^1(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, \mathbf{y}^1, t) = I^1(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t) + $$
$$+ \mathbf{y}^1\mathbf{f}(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t) \tag{15.5.4}$$
$$H^2(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, \mathbf{y}^2, t) = I^2(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t) + $$
$$+ \mathbf{y}^2\mathbf{f}(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t),$$

где $\mathbf{y}^1$ — вектор-строка сопряженных переменных игрока 1, а $\mathbf{y}^2$ — вектор-строка сопряженных переменных игрока 2. Необходимым условием некооперативного равновесия (по Нэшу) является условие, что векторы управления в любой точке заданного интервала времени должны образовывать некооперативные равновесия (по Нэшу) для статической игры с ненулевой суммой, в которой выигрыши равны $H^1(\cdot\cdot\cdot\cdot\cdot)$ и $H^2(\cdot\cdot\cdot\cdot\cdot)$, т. е.

$$H^1(\mathbf{x}, \mathbf{u}^{1*}, \mathbf{u}^{2*}, \mathbf{y}^1, t) \geqslant H^1(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^{2*}, \mathbf{y}^1, t)$$
$$\text{для всех } \mathbf{u}^1 \in \Omega^1 \tag{15.5.5}$$
$$H^2(\mathbf{x}, \mathbf{u}^{1*}, \mathbf{u}^{2*}, \mathbf{y}^2, t) \geqslant H^2(\mathbf{x}, \mathbf{u}^{1*}, \mathbf{u}^2, \mathbf{y}^2, t)$$
$$\text{для всех } \mathbf{u}^2 \in \Omega^2.$$

Иначе говоря,

$$\mathbf{u}^{1*}(t) = \mathbf{S}^{1*}(\mathbf{x}(t)) \text{ максимизирует}$$
$$H^1(\mathbf{x}, \mathbf{u}^1, \mathbf{S}^{2*}(\mathbf{x}), \mathbf{y}^1, t) \tag{15.5.6}$$
$$\mathbf{u}^{2*}(t) = \mathbf{S}^{2*}(\mathbf{x}(t)) \text{ максимизирует}$$
$$H^2(\mathbf{x}, \mathbf{S}^{1*}(\mathbf{x}), \mathbf{u}^2, \mathbf{y}^2, t).$$

Канонические уравнения принимают следующий вид:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t) = \frac{\partial H^1}{\partial \mathbf{y}^1} = \frac{\partial H^2}{\partial \mathbf{y}^2}$$
$$\dot{\mathbf{y}}^1 = -\frac{\partial H^1}{\partial \mathbf{x}} - \frac{\partial H^1}{\partial \mathbf{u}^2}\frac{\partial \mathbf{S}^{2*}}{\partial \mathbf{x}} \tag{15.5.7}$$
$$\dot{\mathbf{y}}^2 = -\frac{\partial H^2}{\partial \mathbf{x}} - \frac{\partial H^1}{\partial \mathbf{u}^1}\frac{\partial \mathbf{S}^{1*}}{\partial \mathbf{x}},$$

где последние члены в двух последних дифференциальных уравнениях определяют влияние стратегии одного из участников на гамильтониан другого.

ЗАДАЧИ

15-А. Найти решение игры двух лиц с нулевой суммой с выигрышем по окончании времени игры. Уравнения движения заданы

$$\dot{x}_1 = au^1 + b\sin u^2$$
$$\dot{x}_2 = -1 + b\cos u^2.$$

Управляющие параметры удовлетворяют следующим условиям [1]:

$$-1 \leqslant u^1 \leqslant 1$$
$$0 \leqslant u^2 < 2\pi.$$

15-Б. В дифференциальной игре уравнения движения имеют вид

$$\dot{x}_1 = u_1\left(1 + 2\sqrt{|x_1|}\right) + u_2$$
$$\dot{x}_2 = -1.$$

Управляющие параметры удовлетворяют ограничениям

$$0 \leqslant u^1 \leqslant 1$$
$$0 \leqslant u^2 \leqslant 1 .$$

Игра начинается в $x_2(t_0) > 0$ и оканчивается при $x_2(t_1) = 0$, а выигрыш игрока 1 определяется как

$$J = \frac{1}{1 + [x_1(t_1)]^2} .$$

Показать, что ось x_2 является «сингулярной поверхностью» в том смысле, что все оптимальные траектории — это кривые, начинающиеся на этой оси [5].

15-В. В дифференциальной игре двух участников с нулевой суммой уравнения движения имеют вид

$$\dot{x}_1 = (u^1 - u^2)^2$$
$$\dot{x}_2 = -1.$$

Скалярные управляющие параметры удовлетворяют условиям

$$|u^1| \leqslant 1$$
$$|u^2| \leqslant 1.$$

Игра начинается в $x_2(t_0) > 0$ и оканчивается в $x_2(t_1) = 0$. Выигрыш участника 1 равен

$$J = x_1(t_1).$$

Показать, что игра не имеет решений в чистых стратегиях. Дайте геометрическую иллюстрацию в плоскости $(x_1, x_2)'$.

15-Г. Предположим, что в игре преследования уравнения движения линейны и сепарабельны, т. е.

$$\dot{\mathbf{x}}^1 = \mathbf{A}^1\mathbf{x}^1 + \mathbf{b}^1 u^1$$
$$\dot{\mathbf{x}}^2 = \mathbf{A}^2\mathbf{x}^2 + \mathbf{b}^2 u^2,$$

а скалярные параметры удовлетворяют ограничениям

$$0 \leqslant |u^1| \leqslant 1$$
$$0 \leqslant |u^2| \leqslant 1.$$

Начальные состояния $x^1(t_0)$ и $x^2(t_0)$ фиксированы. Игра заканчивается, когда $x_1^1(t_1) = x_1^2(t_1)$. Цель игрока 1 (2) — минимизировать (максимизировать) время окончания $t_1 - t_0$. Найти решение. См. [9, 11].

15-Д. В игре преследования на плоскости преследователь (игрок 1) осуществляет управление по координате x_1, преследуемый (игрок 2) — по координате x_2, где

$$\ddot{x}_1 + \alpha\dot{x}_1 = u^1 \quad |u^1| \leqslant 1$$
$$\ddot{x}_2 + \beta\dot{x}_2 = u^2 \quad |u^2| \leqslant 1.$$

Игра оканчивается в t_1, когда

$$x_1(t_1) = x_2(t_2).$$

Показать, что выигрыш конечен (т. е. игра может быть окончена), если $\alpha < \beta$ [14].

15-Е. Вывести «главное уравнение», указанное в примечании на стр. 446, используя методы динамического программирования.

15-Ж. В *дифференциальной игре о вратаре* игрок 1 защищает зону нападения, к которой приближается игрок 2. Игра похожа на хоккей, где участник 1 — вратарь. Игра происходит на плоскости $(x_1, x_2)'$. Зона нападения лежит на оси x_1, ось x_2 делит ее на две равные части длины L.

Игрок 1 начинает движение из зоны нападения, покидая ее с постоянной скоростью v^1, но он может управлять своей горизонтальной скоростью

$$\dot{x}_1^1 = u^1,\ x_1^1(t_0), \quad \text{причем } |u^1| < \bar{u}^1$$
$$\dot{x}_2^1 = v^1,\ x_2^1(t_0) = 0.$$

Игрок 2 начинает движение к зоне нападения из точки, расположенной над зоной нападения с постоянной скоростью v^2, управляя своей горизонтальной скоростью

$$\dot{x}_1^2 = u^2, \quad x_1^2(t_0), \text{ причем } |u^2| \leqslant \bar{u}^2,\ \bar{u}^2 > \bar{u}^1$$
$$\dot{x}_2^2 = -v^2,\ x_2^2(t_0) > 0.$$

Игра заканчивается, когда

$$x_2^1(t_1) = x_2^2(t_1).$$

Здесь выигрыш (проигрыш) участника 1 (2) определяется так:

$$J = \left\{ \begin{matrix} 1 \\ -(x_1^1(t_1) - x_1^2(t_1))^2 \end{matrix} \right\}, \text{ если } \left\{ \begin{matrix} |x_1^2(t_1)| > L + \dfrac{\bar{u}^2 x_2^2(t_1)}{v^2} \\ \text{в противном случае} \end{matrix} \right\}.$$

Дайте интерпретацию платежной функции. Подробно проанализируйте решение. См. [15].

15-З. Лев и человек находятся на круглой арене, у них одинаковые максимальные скорости. Сможет ли лев настичь человека?

15-И. Атакующий и защищающийся — две точки на плоскости, лежащие вне некоторой области цели. Они двигаются с одинаковой скоростью и могут управлять направлением движения. Когда атакующий и защищающийся находятся достаточно близко друг от друга, защищающийся захватывает атакующего. Защищающийся пытается максимизировать расстояние от точки захвата до области цели. Атакующий пытается подойти как можно ближе к области цели. Предполагая, что захват происходит вне области цели, покажите графически стратегию игры [1].

15-К. В *динамической модели ракетной войны* две стороны A и B, принимают участие в игре в промежутке времени от t_0 до t_1. Фазовыми координатами будут ракеты, остающиеся в каждой стране M_A и M_B, и потери каждой страны — C_A и C_B. Уравнения движения имеют вид

$$\dot{M}_A = -\alpha M_A - \beta M_B \beta' f_B$$
$$\dot{M}_B = -\beta M_B - \alpha M_A \alpha' f_A$$
$$\dot{C}_A = (1 - \beta')\, \beta M_B v_B$$
$$\dot{C}_B = (1 - \alpha')\, \alpha M_A v_A.$$

Управляющие параметры для A — это интенсивность огня α и α' — доля ракет, которые используются для противоракетной обороны. Аналогичные управляющие параметры для B — β и β', причем

$$0 \leqslant \alpha \leqslant \bar{\alpha}, \text{ задано } \bar{\alpha}$$
$$0 \leqslant \alpha' \leqslant 1$$
$$0 \leqslant \beta \leqslant \bar{\beta}, \text{ задано } \bar{\beta}.$$
$$0 \leqslant \beta' \leqslant 1.$$

В уравнениях движения f_B — это эффективность ракет B против ракет A, то есть количество ракет A, уничтоженных одной ракетой B. Аналогично f_A — эффективность ракет A против ракет B, v_B — эффективность ракет B против городов A и v_A — эффективность ракет A против городов B. Таким образом, два члена в уравнении для M_A показывают потери ракет A, включая ракеты, израсходованные на обстрел противника и уничтоженные противоракетной обороной B. Граничные условия таковы:

$$M_A(t_0) = M_{A_0}$$
$$M_B(t_0) = M_{B_0}$$
$$C_A(t_0) = 0$$
$$C_B(t_0) = 0.$$

Предположим, что t_1 задано. Найти оптимальные стратегии интенсивности огня и противоракетной обороны для A и B, считая, что целью для A является минимизация $C_A(t_1) - C_B(t_1)$, а для B — минимизация $C_B(t_1) - C_A(t_1)$.

15-Л. Дифференциальной игрой на выживание называется дифференциальная игра двух участников с нулевой суммой. В ней один участник обязательно выигрывает, а другой проигрывает. Конечную поверхность можно разделить на две поверхности: одну W, на которой выигрывает участник 1, и другую L, на которой он проигрывает. Пространство фазовых координат (некоторое подмножество E^n) также можно разделить на область выигрыша WZ, состоящую из тех точек, откуда игрок 1 наверняка попадает в W, область проигрыша LZ, состоящую из тех точек, откуда игрок 2 наверняка попадает на L, и ничейную область N, в которой ни один из участников не имеет гарантированного успеха.

1. Заданы уравнения движения, множество траекторий управления и граничные условия такие же, как в задаче 15-В. Предположим, что

30—0270

$$W = \{x_1(t_1) \mid x_1(t_1) > 0\}$$
$$L = \{x_1(t_1) \mid x_1(t_1) < 0\}.$$

Показать графически WZ, LZ и N.

2. Снова используя условия задачи 15-В, предположим, что

$$W = \{x_1(t_1) \mid x_1(t_1) \leqslant 1\}$$
$$L = \{x_1(t_1) \mid x_1(t_1) > 1\}.$$

Показать WZ, LZ и N графически.

3. Предположим, что N является гладкой поверхностью. Показать, что нормальный к N вектор $\mathbf{V} = (V_1, \ldots, V_n)$, направленный в сторону WZ, удовлетворяет условию. См. [1, 5].

$$\max_{\mathbf{u}^1 \in \Omega} \min_{\mathbf{u}^2 \in \Omega} [\mathbf{V} \cdot \mathbf{f}(\mathbf{x}, \mathbf{u}^1, \mathbf{u}^2, t] = 0.$$