Математическая Теория Игр и её Приложения, т.13, в.4, с. 93–128

УДК 517.977.8
ББК 22.18

ДИФФЕРЕНЦИАЛЬНАЯ ИГРА С ДИСКРЕТНЫМ МОМЕНТОМ ОСТАНОВКИ

Дмитрий В. Хлопин*
Институт математики и механики
им. Н.Н. Красовского УрО РАН
620108, Екатеринбург, ул. С. Ковалевской, 16
e-mail: khlopin@imm.uran.ru

В данной работе рассматривается антагонистическая дифференциальная игра, завершение которой может быть инициировано любым из участников в один из известных заранее моментов времени. Функционал качества может зависеть как от момента завершения игры и положения системы в этот момент, так и от игрока-инициатора завершения игры. При оптимизации математического ожидания этого функционала каждый игрок, на основе имеющейся у него информации о реализовавшейся траектории системы, принимает решения как о своей вероятности завершения игры, так и о собственном управлении конфликтно-управляемой системой; при этом недетерминированные правила–распределения также допустимы. В предположении условия седловой точки в маленькой игре (условие Айзекса) показано существование цены игры. Для построения приближенно оптимальных стратегий игроков использовалась конструкция стохастического поводыря, моделью которого являлась марковская цепь с непрерывным временем.

Ключевые слова: антагонистические игры, игра Дынкина, дифференциальные игры, стохастический поводырь, экстремальный сдвиг,

---

* Работа выполнена при поддержке РНФ (проект 17-11-01093).

марковская цепь с непрерывным временем.

Поступила в редакцию: 30.06.21 *После доработки*: 27.10.21 *Принята к публикации*: 10.12.21

Игровые постановки, в которых каждый игрок (или один из игроков) своими действиями может выбрать удобный для себя момент остановки всей игры, обусловлены многочисленными приложениями, прежде всего в экономике, см., например, [9,15,17]. Игровую постановку, в которой игрок (игроки) управляют моментом выхода из процесса, обычно связывают с основополагающей работой Дынкина [2]. В таких играх обычно динамика считается недетерминированной и может быть задана стохастическим дифференциальным уравнением [13,14] процессом Леви–Феллера [12], марковской цепью с непрерывным [20] или дискретным [7] временем.

В данной работе динамика игры задается детерминированной конфликтно-управляемой системой, а распределение момента завершения не считается известным игрокам заранее, как, например, в работах [6,19] и не считается абсолютно-непрерывной случайной величиной, пусть и зависящей от действий игроков, как в [8]. Каждый игрок может попытаться завершить игру лишь в некоторые дискретные моменты времени, назначив условную вероятность ее завершения (из известных заранее подмножеств компакта $[0;1]$). Наборы моментов времени, в которые игроки могут инициировать завершение игры, не пересекаются и также известны заранее. Вариант, при котором игра так и не будет завершена, при этом не исключается (см. условие (1.4)). Функционал качества является случайной величиной и зависит как от момента завершения игры и положения системы в этот момент, так и от игрока-инициатора досрочного завершения игры; задача игроков – оптимизировать матожидание этой случайной величины. Для этого каждый игрок, на основе имеющейся у него информации о реализовавшейся траектории системы, принимает решения как об условной вероятности завершения игры, так и о собственном управлении конфликтно-управляемой системой. Стратегии при этом фактически предполагаются случайными процессами, неупреждающим образом зависящими от реализующейся траектории системы. В частности, каждый игрок имеет право использовать как классические детерминированные стратегии: кусочно-непрерывные программные стратегии или кусочно-постоянные позиционные стратегии, так и произвольные их распределения, в том числе стратегии со стратегическим поводырем.

Схема управления с поводырем был предложена Н.Н. Красовским и А.И. Субботиным (см. [4]) как метод управления, устойчивый к малым информационным помехам. Чуть позже был предложен ее стохастический вариант [3]. В [5] стохастический поводырь применялся в рамках метода исчезающей нежесткости для уравнений Гамильтона-Якоби. В работах [10,11] был построен стохастический поводырь на основе марковской игры с непрерывным временем. В работе [8] вариант такого стохастического поводыря применялся для построения приближенно оптимальных стратегий в дифференциальной игре на конечном промежутке, досрочное завершение которой являлось пуассоновским процессом с интенсивностью, определяемой игроками.

В предположении условия седловой точки в маленькой игре и несущественности бесконечности (см. (1.4)) будет показано существование цены игры и предложена процедура на основе стохастического поводыря, реализующая с произвольной точностью цену игры. Соответствующая процедура не требует какого-либо знания (накопления знаний) о действиях соперника или памяти обо всей реализованной траектории. В рамках этой процедуры игрок меняет свое управление $\bar{u}$ в некоторые, подчиненные пуассоновскому процессу, моменты времени t_k, при помощи правила

$$\bar{u}(t) \equiv u^{to\, z(t_k)}(t_k, y(t_k)) \qquad \forall t \in [t_k; t_{k+1}),$$

здесь $y(t_k)$ и $z(t_k)$ – реализовавшиеся на этот момент позиции исходной игры и игры-модели, а $u^{to\, z(t_k)}$ – управление, максимально сдвигающее позицию исходной системы к точке $z(t_k)$. Процесс $z(\cdot)$ будет задаваться некоторой конфликтно-управляемой марковской цепью с непрерывным временем, в которой фиктивный игрок-противник насколько возможно отодвигает позицию цепи $z(t)$ от самого позднего доступного ему наблюдения позиции исходной системы, позиции $y(t_k)$, а фиктивный игрок-союзник играет оптимально. Отметим, что его оптимальная стратегия может быть построена точно решением конечной системы уравнений (3.2).

Сама статья организована следующим образом. В первом разделе формализуется исходная дифференциальная игра и формулируется основной результат: существование седловой точки и реализующих ее приближенно оптимальных стратегий со стороны каждого игрока. Во втором разделе для каждого параметра точности подбираются необходимые константы и задается множество состояний для модели на основе марковской игры; там же формулируются удобные для доказательств временные предположения на моменты и вероятности завершения игры. В третьем разделе описываются марковская игра с непрерывным временем и ее оптимальные стратегии, доказывается ряд оценок на расхождение траекторий и значений функционала в этой игре. В четвертом разделе строится стохастический поводырь и доказывается существование цены игры при выполнении временных предположений. В последнем разделе этот результат переносится на общий случай.

1. Исходная дифференциальная игра

Рассмотрим функционирующую в $\mathbb{R}^d$ конфликтно-управляемую систему

$$\frac{d}{dt}y(t) = f(y(t), u(t), v(t)),\ \ y(0) = x_* \in \mathbb{R}^d, \qquad (1.1)$$
$$t \geq 0, \quad u(t) \in \mathbb{U}, \quad v(t) \in \mathbb{V}.$$

Будем далее предполагать, что $\mathbb{U}$ и $\mathbb{V}$ – метрические компакты, а функция $f : \mathbb{R}^d \times \mathbb{U} \times \mathbb{V} \to \mathbb{R}^d$ L-липшицева по x для некоторой константы $L > 1$ и ограничена той же константой. Также потребуем условие седловой точки в маленькой игре (Isaacs condition): для всех $x, w \in \mathbb{R}^d$ выполнено

$$\min_{u\in\mathbb{U}} \max_{v\in\mathbb{V}} \langle w, f(x, u, v)\rangle = \max_{v\in\mathbb{V}} \min_{u\in\mathbb{U}} \langle w, f(x, u, v)\rangle. \qquad (1.2)$$

В рассматриваемой нами постановке игра может завершиться по инициативе любого из игроков, а момент $\theta_{\min}$ окончания игры является недетерминированным и определяется по следующей процедуре. Игроки имеют неограниченно возрастающие наборы моментов времени $(t^I_k)_{k\in\mathbb{N}}$, $(t^{II}_k)_{k\in\mathbb{N}}$ и последовательности отрезков $[\phi^-_k;\phi^+_k]$ и $[\psi^-_k;\psi^+_k]$, вложенных в $[0; 1]$. Будем считать, что наборы моментов времени не пересекаются. Каждый из игроков в свои моменты времени (t^I_k или t^{II}_k соответственно), если игра еще не была завершена, назначает элемент соответствующего отрезка, ϕ_k или ψ_k соответственно, как условную вероятность завершить игру в этот момент по его инициативе. С учетом вышесказанного момент $\theta_{\min}$ завершения игры имеет следующее распределение:

$$\mathbb{P}(\theta_{\min} = t_k^I) = \phi_k \cdot \prod_{i<k}(1-\phi_i) \cdot \prod_{t_k^{II}<t_k^I}(1-\psi_i),$$
$$\mathbb{P}(\theta_{\min} = t_k^{II}) = \psi_k \cdot \prod_{i<k}(1-\psi_i) \cdot \prod_{t_k^{I}<t_k^{II}}(1-\phi_i).$$

Тем самым, игроки кроме ϕ_i, ψ_i не имеют никаких других инструментов влияния на момент завершения игры. В частности, при одних и тех же ϕ_i, ψ_i, не зависят от реализующейся траектории любые вероятности событий о моменте завершения игры и об ее игроке-инициаторе, например если какие-либо ϕ_k^- или ψ_k^- равны единице, то игра в этот момент всегда завершается; если же ϕ_k^+ или ψ_k^+ равны нулю, то в этот момент времени игра закончиться не может.

Обозначим также через θ_1 – случайный элемент, равный моменту окончания игры, если оно было инициировано первым игроком, и равный $+\infty$ в противном случае; аналогично через θ_2 обозначим момент окончания игры, если оно было инициировано вторым игроком, и примем θ_2 равным $+\infty$ в противном случае. Теперь мы также имеем $\theta_{\min} = \min(\theta_1, \theta_2)$ и

$$\mathbb{P}(\theta_1 \geq t_k^I) = \prod_{i<k}(1-\phi_i),\ \mathbb{P}(\theta_2 \geq t_k^{II}) = \prod_{i<k}(1-\psi_i). \tag{1.3}$$

1.1. Стратегии

Будем предполагать, что игроки, помимо общедоступной информации о динамике системы (1.1), начальной позиции игры, возможных моментах θ_I^k, θ_{II}^k завершения игры и допустимых для каждого игрока условных вероятностях завершения игры в эти моменты, обладают также информацией о реализующейся траектории системы (1.1). Они имеют право подать неупреждающим образом зависящее от реализовавшейся траектории как свое управление в систему (1.1), так и условные вероятности завершения игры. Наконец, мы позволим соответствующим правилам быть недетерминированными.

Напомним, что càdlàg функциями называют функции, непрерывные справа и имеющие предел слева, множество таких функций из множества A в множество B будем называть пространством Скорохода (см. [1]).

Обозначим через $\mathfrak{D}^I$ множество всех измеримых отображений U из $D(\mathbb{R}_+, \mathbb{R}^d)$ в пространство вероятностных мер над $D(\mathbb{R}_+, \mathbb{U})$ таких, что для всякого момента $t > 0$ найдется такой момент времени $t' > t$, что для всех $x, x' \in D(\mathbb{R}_+, \mathbb{R}^d)$ из равенства сужений $x|_{[0;t)} = x'|_{[0;t)}$ следует равенство индуцированных $U(x)$ и $U(x')$ вероятностей на $D([0;t'], \mathbb{U})$. Аналогично определим множества $\mathfrak{D}^{II}$. Элементы множеств $\mathfrak{D}^I$, $\mathfrak{D}^{II}$ будем называть допустимыми стратегиями.

Обсудим какими могут оказаться допустимые стратегии, к примеру, у первого игрока. Это могут оказаться константные управления (элементы $\mathbb{U}$); такие управления не зависят ни от времени, ни от реализующейся траектории. Можно рассмотреть программные управления (непрерывные справа и имеющие предел слева управления из $B(\mathbb{R}_+, \mathbb{U})$); такие управления зависят лишь от времени, но никак не зависят от уже реализовавшейся траектории y. В рамках классической формализации теории теории дифференциальных игр обычно рассматриваются позиционные управления вида: $U_t = u(t_i, y(t_i))$ для каждого t из промежутка $[t_i; t_{i+1})$; при этом и отображение $u : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{U}$, и неограниченно возрастающую последовательность моментов t_i можно выбрать произвольно. Отметим, что позиционные стратегии уже используют информацию о реализующейся траектории: более того, для каждого момента времени $t \in [t_i; t_{i+1})$ управление известно хотя бы до момента времени t_{i+1}, таким образом позиционные стратегии также допустимы. Можно использовать управление с поводырем: по реализующейся траектории в реальном времени неупреждающим образом строится движение модели Z, после чего управление выбирается правилом $U_t = u(t_i, Z(t_i))$ для всех $t \in [t_i; t_{i+1})$; так полученная стратегия также допустима. Все перечисленные выше стратегии были детерминированными, однако если выбирать в каждый момент t_i не один элемент из $\mathbb{U}$, а некоторое их распределение, то соответствующее правило будет определять вообще говоря недетерминированную допустимую стратегию. Один из способов задать соответствующие распределения – сделать модельное движение Z случайным процессом, согласованным с какой-либо содержащей исходную фильтрацией, после чего взять маргинальное распределение на $\mathbb{R}^d$ в силу этого процесса.

Примем в качестве стохастического базиса пространство Скорохода $D(\mathbb{R}_+, \mathbb{R}^d)$ (см. [1]), с борелевской сигма-алгеброй $\mathcal{B}(D(\mathbb{R}_+, \mathbb{R}^d))$ и канонической фильтрацией: потоком сигма-алгебр $\mathcal{F}_t \triangleq \sigma(y|_{[0;t]})$, $t \geq 0$. Покажем, что каждая пара допустимых стратегий $(U, V) \in \mathfrak{D}^I \times \mathfrak{D}^{II}$ определяет на этом стохастическом базисе единственную вероятность $\mathbb{P}_{U,V}$, тем самым задавая случайный процесс $\hat{Y}$ на решениях системы (1.1).

Отметим сначала, что в начальный момент процесс задан, поскольку положение процесса $\hat{Y}$ известно, а значит известны и соответствующие ему распределения управлений в этот момент. Тогда для некоторого положительного t' известны индуцированные $U(x) \otimes V(x)$ вероятности над $D([0; t'], \mathbb{U}) \times D([0; t'], \mathbb{V})$, поскольку каждая пара элементов из $D([0; t'], \mathbb{U}) \times D([0; t'], \mathbb{U})$ однозначно восстанавливает некоторое решение системы (1.1), то как образ меры в силу этого отображения, однозначно восстанавливается сужение $P_{U,V}$ над $D([0; t'], \mathbb{R})$, а значит и процесс $\hat{Y}$ на промежутке $[0; t']$.

Следовательно непусто множество промежутков, содержащих момент времени 0 и содержащихся в полуоси, на которых нужный нам процесс $\hat{Y}$ определен и определен однозначно. Если это множество содержит полуось, то все показано. Предположим, что это не так. Тогда найдем в нашем множестве наибольший элемент, промежуток $\mathcal{T}$. В силу непрерывности траекторий системы (1.1), а значит и $\hat{Y}|_{\mathcal{T}}$, мы можем однозначно восстановить траектории у $\hat{Y}$ на замыкании $\mathcal{T}$, некотором отрезке $[0; t]$. Тогда найдется такой положительный момент времени t'', что для каждой реализации процесса $\hat{Y}$ на $[0; t]$ известно распределение управлений на $[0; t'')$. Зная эти управления по каждой такой реализации можно построить распределение траекторий системы (1.1) до момента времени t'', тем самым процесс $\hat{Y}$ восстановлен единственным образом уже на интервале $[0; t'')$, содержащем как $\mathcal{T}$, так и его замыкание $[0; t]$. Но это противоречит выбору $\mathcal{T}$ как наибольшего промежутка, на котором процесс $\hat{Y}$ определен и определен однозначно. Следовательно этот процесс определен однозначно на всей полуоси.

1.2. Моменты завершения

Помимо управлений системой (1.1) игроки также управляют вероятностями ϕ_i, ψ_i выхода из игры. Чтобы описать допустимые при этом выборе возможности первого игрока потребуем, чтобы каждое ϕ_k являлось $\mathcal{F}_{\theta^I_k}\otimes\mathcal{F}^{\mathbb{U}}_{\theta^I_k}$-измеримой случайной величиной (на $D(\mathbb{R}_+,\mathbb{R}^m)\times D(\mathbb{R}_+,\mathbb{U})$) со значениями в $[\phi^-_k;\phi^+_k]$, в частности неупреждающим образом зависело от пары (траектория, управление) из $D(\mathbb{R}_+,\mathbb{R}^m)\times D(\mathbb{R}_+,\mathbb{U})$.

В частности, ими окажутся как аналоги программных управлений случайные величины со значениями в $[\phi^-_k;\phi^+_k]$, не зависящие от реализующейся траектории. Допустимыми последовательностями будет также произвольные удовлетворяющие тем же правилам вероятности, зависящие от состояния процесса в момент t^I_k (или более ранний). Наконец, ничего не мешает задать удовлетворяющие этим правилам вероятности через траектории некоторого другого случайного процесса Z, согласованного с содержащей исходную фильтрацией и отвечающего за стратегию U.

Отметим, что правило (1.3) каждой такой последовательности ϕ_k сопоставляет распределение(вероятность) θ_1 на $\{\theta^I_k \,|\, i\in\mathbb{N}\}\cup\{+\infty\}$. Более того, при этом можно считать, что имеются некоторое стандартное вероятностное пространство Ω^I (например, в [8] для этого использовался экземпляр $([0;1],\mathcal{B}([0;1]),\lambda)$), и для каждого элемента из $D(\mathbb{R}_+,\mathbb{R}^m)\times D(\mathbb{R}_+,\mathbb{U})$ его значение у θ_1 являются распределением некоторой случайной величины, заданной в этом пространстве и обозначаемой для простоты той же буквой. Будем называть такие зависящие от $D(\mathbb{R}_+,\mathbb{R}^m)\times D(\mathbb{R}_+,\mathbb{U})$ случайные величины (на Ω^I) со значениями в $\{\theta^I_k \,|\, i\in\mathbb{N}\}\cup\{+\infty\}$ допустимыми моментами выхода первого игрока. Поместим все такие случайные величины в множество $\mathfrak{Q}^I$. Аналогично для второго игрока зафиксируем стандартное вероятностное пространство Ω^{II} и каждой последовательности $\mathcal{F}_{\theta^I_k}\otimes\mathcal{F}^{\mathbb{V}}_{\theta^I_k}$-измеримых случайных величин со значениями в $[\phi^-_k;\phi^+_k]$ сопоставим заданную на Ω^{II} случайную величину θ_2, распределение которой подчиняется (1.3). Такие зависящие от $D(\mathbb{R}_+,\mathbb{R}^m)$ случайные величины (на Ω^{II}) со значениями в $\{\theta^I_k \,|\, i\in\mathbb{N}\}\cup\{+\infty\}$ будем называть допустимыми моментами выхода второго игрока. Образуем из них множество $\mathfrak{Q}^{II}$.

Теперь на вероятностном пространстве $\Omega \stackrel{\triangle}{=} \Omega^I \times \Omega^I I$ корректно определяется зависящая от $D(\mathbb{R}_+, \mathbb{R}^m) \times D(\mathbb{R}_+, \mathbb{U}) \times D(\mathbb{R}_+, \mathbb{V})$ случайная величина $\theta_{\min} = \min(\theta_1, \theta_2)$. Более того, всякая пара допустимых стратегий (U, V) однозначно задает также вероятность $P_{\Omega,U,V}$ на $\Omega^I \times \Omega^I I \times D(\mathbb{R}_+, \mathbb{R}^m)$.

1.3. Цели игроков

Теперь определим цели игроков. Пусть заданы функции $\sigma_1 : \mathbb{R}_+ \times \mathbb{R}^d \to [-1; 1]$ и $\sigma_2 : \mathbb{R}_+ \times \mathbb{R}^d \to [-1; 1]$. Будем также считать, увеличив при необходимости L, что они L-липшицевы по x, t. Число $\sigma_1(\theta_1, y(\theta_1))$ станет платой первого второму, если первый инициировал завершение игры в момент времени θ_1. Число $\sigma_2(\theta_2, y(\theta_2))$ станет платой первого второму, если второй инициировал завершение игры в момент времени θ_2.

Теперь в случае завершения игры в конечный момент времени платежная функция определена. Поскольку существует, возможно и нулевая, вероятность того, что игра никогда не будет завершена, потребуем выполнение хотя бы одного из трех условий:

1. Платежные функции σ и управляемая система (1.1) таковы, что $\sigma_1(t, y(t))$ и $\sigma_2(t, y(t))$ стремятся к общему пределу при $t \to \infty$, причем равномерно по всем возможным траекториям y системы (1.1): для всякого положительного ε найдется такое натуральное N, что для всех $t > N$ и всех траекторий y системы (1.1) выполнено
$$|\sigma_1(t, y(t)) - \sigma_2(t, y(t))| + |\sigma_1(N, y(N)) - \sigma_1(t, y(t))| < \varepsilon.$$
2. Хотя бы один из наборов $(\phi_k^-)_{k\in\mathbb{N}}$ и $(\psi_k^-)_{k\in\mathbb{N}}$ содержит единицу.
3. Хотя бы один из рядов $\sum\limits_{k\in\mathbb{N}} \phi_k^-$ и $\sum\limits_{k\in\mathbb{N}} \psi_k^-$ расходится.

Как второе, так и третье условие гарантирует, что с вероятностью 1, вне зависимости от выбора допустимых стратегий U, V и моментов выхода θ_1, θ_2 игроков, игра завершится за конечное время. Более того, в силу выполнения любого из трех условий выше найдется такое достаточно большое натуральное N_ε, что при каждом $i \in \{1, 2\}$, при любых допустимых стратегиях U,V и моментов выхода θ_1,θ_2 игроков момент окончания игры $\theta_{\min}=\min(\theta_1,\theta_2)$ удовлетворяет

$$\mathbb{P}_{\Omega,U,V}(\theta_{\min}\geq N_\varepsilon)\sup_{t\geq N_\varepsilon}\mathbb{E}_{U,V}\big(\sigma_i(\theta_{\min},y(\theta_{\min}))\,\big|\,\theta_{\min}\geq t\big)-\varepsilon<$$
$$<\mathbb{P}_{\Omega,U,V}(\theta_{\min}\geq N_\varepsilon)\mathbb{E}_{\Omega,U,V}\big(J(y,\theta_1,\theta_2)\,\big|\,\theta_{\min}\geq N_\varepsilon\big)<$$
$$<\mathbb{P}_{\Omega,U,V}(\theta_{\min}\geq N_\varepsilon)\inf_{t\geq N_\varepsilon}\mathbb{E}_{\Omega,U,V}\big(\sigma_i(\theta_{\min},y(\theta_{\min}))\,\big|\,\theta_{\min}\geq t\big)+\varepsilon. \quad (1.4)$$

Действительно, при выполнении первого условия в этой формуле мало отличаются друг от друга все матожидания, второе условие при достаточно большом N_ε зануляет вероятности в этой формуле, третье условие позволяет хотя бы одно из выражений

$$-\sum_{t_k^{II}<N_\varepsilon}\ln(1-\phi_k^-),\qquad -\sum_{t_k^{II}<N_\varepsilon}\ln(1-\psi_k^-)$$

сделать сколь угодно большим, тем самым делая вероятности в этой формуле также стремящимися нулю.

Теперь, в случае, если игра не завершена за конечное время и вероятность такого события положительна, корректно писать

$$\sigma_1(+\infty,y(+\infty))=\sigma_2(+\infty,y(+\infty)),$$

а значит и плата первого второму в этом случае также определена: $\sigma_1(\theta_1,y(\theta_1))=\sigma_2(\theta_2,y(\theta_2))$. Но тогда при выполнении любого из трех условий выше плата первого второму почти всюду равна

$$J(y,\theta_1,\theta_2)=\begin{cases}\sigma_1(\theta_1,y(\theta_1)), & \theta_1=\theta_{\min},\\ \sigma_2(\theta_2,y(\theta_2)), & \theta_2=\theta_{\min}.\end{cases}$$

1.4. Цена игры

Пусть игроки выбрали некоторые допустимые стратегии U,V. Как показано выше однозначно восстанавливается распределение процесса $\hat{Y}$ на решениях (1.1), тем самым на множестве всевозможных решений (1.1) однозначно восстанавливается вероятность $\mathbb{P}_{U,V}$, что, в свою очередь, задает и соответствующее математическое ожидание $\mathbb{E}_{U,V}$. Теперь с каждым выбором допустимых моментов выхода θ_1,θ_2 можно рассмотреть распределение случайной величины $J(y,\theta_1,\theta_2)$.

Будем считать, что игроки своими действиями стараются оптимизировать ее матожидание, число

$$\mathbb{E}_{\Omega,U,V} J(y, \theta_1, \theta_2). \tag{1.5}$$

В зависимости от того, какой именно из игроков информационно дискриминирован, мы получаем два значения игры

$$\mathcal{V}^+ \triangleq \inf_{U \in \mathfrak{D}^I, \theta_1 \in \mathfrak{Q}^I} \sup_{V \in \mathfrak{D}^{II}, \theta_2 \in \mathfrak{Q}^{II}} \mathbb{E}_{\Omega,U,V} J(y, \theta_1, \theta_2),$$
$$\mathcal{V}^- \triangleq \sup_{V \in \mathfrak{D}^{II}, \theta_2 \in \mathfrak{Q}^{II}} \inf_{U \in \mathfrak{D}^I, \theta_1 \in \mathfrak{Q}^I} \mathbb{E}_{\Omega,U,V} J(y, \theta_1, \theta_2).$$

Основная цель статьи – доказать следующую теорему:

Теорема 1.1. *При выполнении* (1.4) *имеет место равенство* $\mathcal{V}^- = \mathcal{V}^+$. *Более того, для каждого положительного* ε *найдутся такие допустимые стратегии* $U^\varepsilon \in \mathfrak{D}^I$, $V^\varepsilon \in \mathfrak{D}^{II}$ *игроков и моменты выхода* $\theta_1^\varepsilon \in \mathfrak{Q}^I, \theta_2^\varepsilon \in \mathfrak{Q}^{II}$, *что*

$$-\varepsilon + \sup_{V \in \mathfrak{D}^{II}, \theta_2 \in \mathfrak{Q}^{II}} \mathbb{E}_{\Omega,U^\varepsilon,V} J(y, \theta_1^\varepsilon, \theta_2) \leq \mathcal{V}^+ =$$
$$= \mathcal{V}^- \leq \varepsilon + \inf_{U \in \mathfrak{D}^I, \theta_1 \in \mathfrak{Q}^I} \mathbb{E}_{\Omega,U,V^\varepsilon} J(y, \theta_1, \theta_2^\varepsilon).$$

Опишем план ее доказательства. Поскольку $\mathcal{V}^- \leq \mathcal{V}^+$, то, из симметрии игроков, при каждом положительном ε требуется описать лишь для первого игрока нужные допустимые стратегию U_ε и момент выхода θ_1^ε, удовлетворяющий

$$-\varepsilon + \sup_{V \in \mathfrak{D}^{II}, \theta_2 \in \mathfrak{Q}^{II}} \mathbb{E}_{\Omega,U^\varepsilon,V} J(y, \theta_1^\varepsilon, \theta_2) \leq \mathcal{V}^+. \tag{1.6}$$

Для этого мы сначала сделаем удобные для доказательства временные предположения на допустимые моменты выхода игроков, введем нужные константы, в частности перейдя от t_k^I, t_k^{II} к $t_n = nh$; затем рассмотрим в качестве модельной марковскую игру с непрерывным временем, укажем для нее ряд оценок и построим оптимальные стратегии в этой игре. На основе этих оптимальных стратегий далее мы опишем поведение модели-поводыря и, на ее основе, конструкцию стратегии для исходной игры, покажем близость в среднеквадратичном соответствующих им траекторий, а значит и значений, в среднем, платежной функции. В последнем разделе мы снимем временные предположения на допустимые моменты выхода игроков.

2. Временные предположения и введение констант

Возьмем некоторое положительное число $\varepsilon < \min(1/16L\pi, 6/\sqrt{d})$, а с ним и некоторое натуральное число $N_{\varepsilon/12}$ (см. (1.4)), тогда можно найти и некоторый подынтервал $[T_-;T_+) \subset [N_{\varepsilon/12};+\infty)$, не содержащий элементов неограниченно возрастающих последовательностей $(t^I_k)_{k\in\mathbb{N}}$, $(t^{II}_k)_{k\in\mathbb{N}}$. Поскольку в $[0;T_+)$ элементов этих последовательностей конечное число, найдется такое положительное $h < \frac{T_+-T_-}{2}$, для которого при всех натуральных k, l из $t^I_k, t^{II}_l \le T_+$ следует

$$t^I_k > 2h, \quad t^{II}_l > 2h, \quad |t^I_k - t^{II}_l| > 2h, \quad t^I_k/h \notin \mathbb{Z}, \quad t^{II}_l/h \notin \mathbb{Z}.$$

При этом, уменьшая при необходимости h, при $t^I_k \le T_+$ и $\phi^-_k > 1-h$ можно гарантировать и $\phi^-_k = 1$. Более того, можно то же свойство обеспечить также и для пар (t^{II}_k, ψ^+_k).

Назначим в качестве T наименьший элемент $h\mathbb{Z}$, больший T_-. Теперь $T \in [T_-;T_+)\cap h\mathbb{Z}$, а интервал $(T-h;T+h)$ не содержит элементов последовательностей $(t^I_k)_{k\in\mathbb{N}}$, $(t^{II}_k)_{k\in\mathbb{N}}$.

Сделаем теперь следующие временные предположения на T, h, а также последовательности $(t^I_k)_{k\in\mathbb{N}}$, $(t^{II}_k)_{k\in\mathbb{N}}$ и соответствующие им отрезки $[\phi^-_k;\phi^+_k]$, $[\psi^-_k;\psi^+_k]$:

1. для всех $k \in \mathbb{N}$ из $t^I_k \le T$ следует $\phi^+_k < 1-h$, а из $t^{II}_k \le T$ следует $\psi^+_k < 1-h$;

2. все не превосходящие T элементы последовательностей $(t^I_k)_{k\in\mathbb{N}}$, $(t^{II}_k)_{k\in\mathbb{N}}$ лежат в $h\mathbb{N}$.

Первое предположение гарантирует, что вероятность незавершения игры к моменту времени T отделена от нуля константой, не зависящей от действий игроков. Второе предположение означает, что всевозможные моменты окончания игры до момента T включительно лежат в некоторой решетке.

Помимо временных предположений выше, взяв h при необходимости равным h/k при достаточно большом $k \in \mathbb{N}$, мы всюду далее будем считать обеспеченным следующее неравенство:

$$h < \min\Big(\frac{\varepsilon^2 e^{-8LT}}{288L^2 d}, \frac{\varepsilon}{12T(1+16^{1/3}T^{-2/3})^2}, \frac{\varepsilon}{6(L+4)}\Big),$$

что с учетом $L > 1$ и $\varepsilon < \min(1/16\pi L, 6/\sqrt{d}) < 192(\sqrt[3]{2}-1)$ гарантирует также

$$6h(L+4) < \varepsilon, \quad h < \frac{1}{32\pi L(L+1)} < \frac{1}{L(2L/3+5)}, \quad Lh\sqrt{d} < 1, \tag{2.1}$$

$$3h < 32\pi h(hL\sqrt{d} + L^2) < 1, \quad Le^{4LT}\sqrt{2d}\sqrt{h} < \varepsilon/12, \tag{2.2}$$

$$h < \sqrt[3]{32T} - \sqrt[3]{16T}, \quad hT(1+16^{1/3}T^{-2/3})^2 < \varepsilon/12. \tag{2.3}$$

По построению, для всякого натурального n найдется не более одного элемента последовательностей $(t_k^I)_{k\in\mathbb{N}}$ и $(t_k^{II})_{k\in\mathbb{N}}$ на промежутке $[nh; nh+h)$. Если на этом промежутке находится некоторый t_k^I, присвоим $\Phi_n \stackrel{\triangle}{=} [\phi_k^-/h; \phi_k^+/h]$; в противном случае примем $\Phi_n \stackrel{\triangle}{=} \{0\}$. Если там находится некоторый t_k^{II}, присвоим $\Psi_n \stackrel{\triangle}{=} [\psi_k^-; \psi_k^+]$, в противном случае – $\Psi_n \stackrel{\triangle}{=} \{0\}$.

Заметим, что в силу ограниченности динамики системы (1.1) найдется компакт $K_< \subset \mathbb{R}^d$ за пределы которого, на промежутке $[0;T]$ не может выйти никакое решение y уравнения (1.1) с начальным условием $y(0) = x_*$. Увеличив при необходимости этот компакт до $K_>$, мы можем считать, что любое движение (1.1) в любой момент из $[0;T]$ удалено от границ компакта $K_>$ хотя бы на расстояние L.

Зафиксируем некоторую бесконечно гладкую монотонно невозрастающую скалярную функцию $a : \mathbb{R} \to [0;1]$ так, чтобы $a(0) = 1$, $a(L) = 0$, а ее константа Липшица не превосходила 1. Введем также

$$\hat{f}(x,u,v) \stackrel{\triangle}{=} a(\mathrm{dist}(x; K_<))f(x,u,v)$$

для всех $(x,u,v) \in \mathbb{R}^d \times \mathbb{U} \times \mathbb{V}$. Оставаясь всюду непрерывной эта функция вне $K_>$ равна нулю. Теперь, вслед за f, функция $\hat{f}|_{K_>\times\mathbb{U}\times\mathbb{V}}$ липшицева по x с константой $2L$, а ее норма также не больше L.

Лишь теперь примем, наконец, $\mathcal{Z} \stackrel{\triangle}{=} h((\mathbb{N}\cup\{0\}) \times \mathbb{Z}^{d+1}) \subset \mathbb{R}^{d+2}$ и $\mathcal{Z}_< \stackrel{\triangle}{=} \mathcal{Z} \cap (\mathbb{R} \times \mathbb{K}_> \times ((-\infty;-1] \cup [1;+\infty) \cup \{0\}) \subset \mathbb{R}^{d+2}$.

Обозначим базис в $\mathbb{R}^{d+2}$ через $(\pi_0, \pi_1, \ldots, \pi_d, \pi_{d+1})$. Для краткости, координаты вектора $z \in \mathbb{R}^{d+2}$ будем записывать в виде

$$(\pi_0 z, \pi_1 z, \ldots, \pi_d z, \pi_{d+1} z).$$

При этом нулевая координата каждой точки будет соответствовать времени, последняя предназначена для фиксации момента окончания игры. Поскольку координаты $(\pi_1 z, \dots, \pi_d z)$ будут отслеживать траекторию исходной системы, мы будем считать, что $\mathbb{R}^d$ совпадает с подмножеством

$$\{z \in \mathbb{R}^{d+2} \,|\, \pi_0 z = \pi_{d+1} z = 0\} \subset \mathbb{R}^{d+2},$$

следовательно мы можем вычитать элементы $\mathbb{R}^d$ и $\mathbb{R}^{d+2}$ друг из друга, получая элемент $\mathbb{R}^{d+2}$. Более того, всякий элемент $z \in \mathbb{R}^{d+2}$ проецируется на $\mathbb{R}^d$ правилом:

$$z \mapsto z - (\pi_0 z)\pi_0 - (\pi_{d+1} z)\pi_{d+1}.$$

Тогда мы можем применять скалярное произведение и норму на $\mathbb{R}^d$ к векторам $\mathbb{R}^{d+2}$, обозначая их через $\langle \cdot, \cdot \rangle_d$ и $|| \cdot ||_d$ соответственно.

3. Модельная игра

В качестве поводыря будут применяться траектории специально подобранной модельной игры: конфликтно-управляемой марковской цепи с непрерывным временем с фазовым пространством $\mathcal{Z}_<$.

3.1. Рандомизированные стратегии в марковской игре

В качестве стратегий игроков возьмем стационарные рандомизированные стратегии.

Под стационарной рандомизированной стратегией $\bar{\mu}$ первого игрока будем понимать пару (μ, ϕ_μ) отображений, сопоставляющих каждому $w = (nh, x, s) \in \mathcal{Z}_<$ некоторую вероятностную меру $\mu[w]$ на $\mathbb{U}$ и число $\phi_\mu[w] \in \Phi_n$. Под рандомизированной стратегией $\bar{\nu}$ второго игрока будем понимать пару (ν, ψ_ν) отображений, сопоставляющих каждому $w = (nh, x, s) \in \mathcal{Z}_<$ некоторую вероятностную меру $\nu[w]$ на $\mathbb{V}$ и число $\psi_\nu[w] \in \Psi_n$. Обозначим множество всех стационарных стратегий первого и второго игрока через $\check{\mathbb{U}}_\varsigma$ и $\check{\mathbb{V}}_\varsigma$. Через $\mathbb{U}_\varsigma$ и $\mathbb{V}_\varsigma$ будем обозначать их проекции: множества отображений $w \mapsto \mu[w]$ и $w \mapsto \nu[w]$ соответственно.

Определим отображение $\check{f} : \mathcal{Z}_< \times \mathbb{U}_\varsigma \times \mathbb{V}_\varsigma \to \mathbb{R}^d$ правилом: для всякой пары стратегий $(\mu, \nu) \in \mathbb{U}_\varsigma \times \mathbb{V}_\varsigma$ и точки $w = (t, x, s) \in \mathcal{Z}_<$

$$\check{f}(w, \mu, \nu) \triangleq 1_{\{0\}}(s) \int_{\mathbb{U}} \int_V \hat{f}(x, u, v)\, \mu[w](du)\nu[w](dv).$$

Зададим для всякого $i \in [1:d]$ его проекции $\pi_i^+\check{f}$, $\pi_i^-\check{f}$ правилами:

$$\pi_i^+\check{f}(w,\mu,\nu) \stackrel{\triangle}{=} 1_{\{0\}}(s)\int_{\mathbb{U}}\int_V \max(0,\pi_i\hat{f}(x,u,v)\,\mu[w](du)\nu[w](dv),$$
$$\pi_i^-\check{f}(w,\mu,\nu) \stackrel{\triangle}{=} 1_{\{0\}}(s)\int_{\mathbb{U}}\int_V \min(0,\pi_i\hat{f}(x,u,v))\,\mu[w](du)\nu[w](dv).$$

Нам также потребуются зависящие от времени стратегии, отображения из $D(\mathbb{R}_+,\check{\mathbb{U}}_\varsigma)$ и $D(\mathbb{R}_+,\check{\mathbb{V}}_\varsigma)$. Обозначим их множества через $\check{\mathbb{U}}_\varpi$ и $\check{\mathbb{V}}_\varpi$ соответственно. На эти множества отображения $\check{f},\pi_i^+\check{f},\pi_i^-\check{f}$ продолжаются по тем же формулам.

3.2. Динамика марковской игры

Для всех точек $w=(t,x,s)\in\mathcal{Z}_<$ и стационарных рандомизированных стратегий $\bar{\mu},\bar{\nu}$ игроков определим меру Леви $\check{\eta}(w,\bar{\mu},\bar{\nu};\cdot)$: для каждого подмножества $A\subset\mathcal{Z}_<$ приняв $\check{\eta}(w,\bar{\mu},\bar{\nu};A)$ равным

$$\frac{1}{h}\delta_{\pi_0}(A)+\frac{1}{h}\sum_{i=1}^{d}\left[\pi_i^+\check{f}(w,\mu,\nu)\delta_{h\pi_i}(A)+\pi_i^-\check{f}(w,\mu,\nu)\delta_{-h\pi_i}(A)\right]+$$
$$+\frac{1_{\{0\}}(\pi_{d+1}w)}{h}\Big[\frac{\phi_\nu[w]}{1-\phi_\nu[w]}\delta_{(s+1)\pi_{d+1}}(A)+\frac{\psi_\nu[w]}{1-\psi_\nu[w]}\delta_{-(s+1)\pi_{d+1}}(A)\Big].$$

Отметим, что при фиксированных стратегиях μ,ν отображения $w\mapsto\check{f}(w,\mu,\nu)$, $w\mapsto(\pi_i^+\check{f}(w,\mu,\nu),\pi_i^-\check{f}(w,\mu,\nu))$ наследуют от функции f ее $2L$-липшицевость по x и ограниченность ее нормы константой L. Тем самым показаны также следующие неравенства: для всех $w\in\mathcal{Z}_<$

$$\Big\|\int_{\mathbb{R}^{d+2}}\sum_{i=1}^{d}\pi_i y\,\eta(w,\bar{\mu},\bar{\nu};dy)\Big\|\leq L,$$
$$\sum_{i=1}^{d}\Big|\pi_i^+\check{f}(w,\mu,\nu)+\pi_i^-\check{f}(w,\mu,\nu)\Big|\leq L\sqrt{d}. \qquad (3.1)$$

Для всех рандомизированных стратегий $\bar{\mu},\bar{\nu}$ для такой меры $\check{\eta}$ рассмотрим цепь Маркова с непрерывным временем, соответствующую матрице Колмогорова $(\bar{Q}_{wy}(\bar\mu,\bar\nu))_{w,y\in\mathcal{Z}_<}$, введенной правилом:

$$\begin{cases} \frac{1}{h}, & y = w + h\pi_0,\\ \frac{1}{h}\pi_i^+\check f(w,\mu,\nu), & y = w + h\pi_i, i\in[1:d],\\ \frac{1}{h}\pi_i^-\check f(w,\mu,\nu), & y = w - h\pi_i, i\in[1:d],\\ \frac{1_{\{0\}}(s)}{h}\cdot\frac{\phi_\nu[w]}{1-\phi_\nu[w]}, & y = w + (1+\pi_0 w)\pi_{d+1},\\ \frac{1_{\{0\}}(s)}{h}\cdot\frac{\psi_\nu[w]}{1-\psi_\nu[w]}, & y = w - (1+\pi_0 w)\pi_{d+1},\\ -\frac{1}{h} - \frac{1}{h}\sum_{i=1}^d(\pi_i^+\check f(w,\mu,\nu) + \pi_i^-\check f(w,\mu,\nu)) & \\ -\frac{1_{\{0\}}(s)}{h}\big(\frac{\phi_\nu[w]}{1-\phi_\nu[w]} + \frac{\psi_\nu[w]}{1-\psi_\nu[w]}\big), & y = w,\\ 0, & \text{иначе.} \end{cases}$$

Замечание 3.1. По построению марковской цепи каждый из моментов остановки

$$\tau_k \stackrel{\triangle}{=} \{t \,|\, \pi_0\check Y(t) = kh\}$$

почти всюду конечен, более того, являясь суммой независимых в совокупности, экспоненциально с параметром $1/h$ распределенных случайных величин, каждый из них никак не зависит от действий игроков, в частности от прошлых и/или будущих значений μ, ν, ϕ, ψ.

Замечание 3.2. Отметим также, что по построению марковской цепи вероятность прыжка по последней координате (в $\pi_{d+1}\check Y$) между моментами остановки τ_{k-1} и τ_k для интенсивностей (ϕ,ψ) совпадает с вероятностью окончания исходной игры в момент $t_k = kh$ при тех же интенсивностях игроков.

3.3. Цена игры и оптимальные стратегии в марковской игре

Поскольку матрицы $\bar Q^h_{wy}$, как и длины прыжков, равномерно ограничены, выполнены все предположения [16, Remark 4.2(b)]. Тогда, как показано в [16, Proposition 3.1(a)], для каждой пары зависящих от времени рандомизированных стратегий $(\bar\mu,\bar\nu)\in\check{\mathbb{U}}_\varpi\times\check{\mathbb{V}}_\varpi$ и начальных условий z_0 (в момент времени 0) имеется порожденный ими процесс $(\check Y(t))_{t\geq 0}$, а значит и вероятность $\check{\mathbb{P}}_{\bar\mu,\bar\nu}[z_0]$ на всевозможных càdlàg отображениях из $\mathbb{R}_+$ в $\mathcal{Z}_<$.

Теперь введем текущий платеж будущей марковской игры. Для всякого $w=(t,x,s)\in\mathbb{R}^{d+1}\times([-1-T;-1]\cup\{0\}\cup[1;1+T])$ определим

$$\check\sigma(w)=\begin{cases}\frac{1}{2}(\sigma_1(T,x)+\sigma_2(T,x)), & s=0;\\ \sigma_1(s-1,x), & s\geq 1;\\ \sigma_2(-s-1,x), & s\leq -1.\end{cases}$$

Как и σ_i, эта функция ограничена единицей и L-липшицева. Для каждой начальной позиции $z_0\in\mathcal{Z}_<$ в момент времени 0 игроки могут обеспечить одно из значений в зависимости от того, кто из них информационно дискриминирован:

$$\check{\mathcal{V}}^+(z_0)=\inf_{\bar\mu\in\check{\mathbb{U}}_\varpi}\sup_{\bar\nu\in\check{\mathbb{V}}_\varpi}\check{\mathbb{E}}_{\bar\mu,\bar\nu}[z_0]\int_0^\infty he^{-ht}\check\sigma(z(t))\,dt,$$
$$\check{\mathcal{V}}^-(z_0)=\sup_{\bar\nu\in\check{\mathbb{V}}_\varpi}\inf_{\bar\mu\in\check{\mathbb{U}}_\varpi}\check{\mathbb{E}}_{\bar\mu,\bar\nu}[z_0]\int_0^\infty he^{-ht}\check\sigma(z(t))\,dt.$$

Теперь, согласно [16, Theorem 5.1] и [21, Theorem 2], система уравнений (см., [16, (5.4)] и [21, (11)])

$$\inf_{\bar\mu[w]}\sup_{\bar\nu[w]}\sum_{y\in\mathcal{Z}_<}\bar Q_{wy}(\bar\mu,\bar\nu)\check{\mathcal{V}}(y)=h(\check{\mathcal{V}}(w)-e^{-h\pi_0 w}\check\sigma(w))=$$
$$=\sup_{\bar\nu[w]}\inf_{\bar\mu[w]}\sum_{y\in\mathcal{Z}_<}\bar Q_{wy}(\bar\mu,\bar\nu)\check{\mathcal{V}}(y),\quad w\in\mathcal{Z}_< \tag{3.2}$$

имеет в силу условия седловой точки в маленькой игре единственное решение и это решение совпадает с $\check{\mathcal{V}}\stackrel{\triangle}{=}\check{\mathcal{V}}^-=\check{\mathcal{V}}^+$. Более того, любые стационарные стратегии $\bar\mu^{opt},\bar\nu^{opt}$ игроков, реализующие седловую точку в этой системе при всех $w\in\mathcal{Z}_<$, являются оптимальными стратегиями в этой задаче (см., например [21,(6)] и [16,(5.5)-(5.7)]). Выберем какие-либо $\bar\mu^{opt}$, $\bar\nu^{opt}$ с таким свойством.

3.4. Оценки на целевой функционал марковской игры

Сначала отметим, что, как и в [8,(4.4)], можно показать неравенство

$$\check{\mathbb{E}}\max_{t\in[0;T+r],t/h\in\mathbb{N}}|\pi_0\check Y(t)-t|^2\leq 4\check{\mathbb{E}}|\pi_0\check Y(T+r)-T-r|^2=8h(T+r)$$

для всех натуральных r/h. Теперь, в силу неравенства Маркова, вероятность того, что максимальное на этом промежутке значение

$|\pi_0\check{Y}(t)-t|^2$ окажется больше r^2, не превосходит $8h(T+r)/r^2$. Тогда и вероятность того, что соответствующий моменту времени T момент остановки $\tau_{T/h}$ окажется больше $T+r$, не превосходит $8h(T+r)/r^2$.

Рассмотрим произвольную траекторию $\check{Y}$ марковской игры. Отметим, что по построению, с момента остановки $\tau_{T/h}$ эта траектория меняется лишь по нулевой координате, в частности $\check{\sigma}(\check{Y}(t)) = \check{\sigma}(\check{Y}(\tau_{T/h}))$ для всех $t \geq \tau_{T/h}$. В силу ограниченности $\check{\sigma}$, для всех траекторий $\check{Y}$ из $\tau_{T/h} \leq T+r$ следует

$$\Big|\int_0^\infty he^{-ht}\check{\sigma}(\check{Y}(t))\,dt - \check{\sigma}(\check{Y}(\tau_{T/h}))\Big| \leq h\int_0^{\tau_{T/h}} e^{-hT}dt \leq h(T+r),$$

а в случае $\tau_{(T+r)/h} > T+r$, этот модуль оценивается сверху числом 2. Поскольку вероятность события $\tau_{T/h} > T+r$, как замечено выше, не превосходит $8h(T+r)/r^2$, нами показано, что

$$\check{\mathbb{E}}\Big|\int_0^\infty he^{-ht}\check{\sigma}(\check{Y}(t))\,dt - \check{\sigma}(\check{Y}(\tau_{T/h}))\Big| \leq h(T+r)(1+16/r^2). \qquad (3.3)$$

Правая часть в (3.3) убывает по r при $r \leq \sqrt[3]{32T}$. Возьмем натуральное число из промежутка $[\sqrt[3]{16T}/h; \sqrt[3]{16T}/h+1)$ в качестве r/h, в силу (2.3) обеспечено $r \leq \sqrt[3]{32T}$. Теперь, оценив в (3.3) правую часть через

$$h(T+r)(1+16/r^2) \leq hT(1+\sqrt[3]{16T}/T)^2 \overset{(2.3)}{\leq} \varepsilon/12,$$

получим независящую от действий игроков оценку

$$\check{\mathbb{E}}\Big|\int_0^\infty he^{-ht}\check{\sigma}(\check{Y}(t))\,dt - \check{\sigma}(\check{Y}(\tau_{T/h}))\Big| \leq \varepsilon/12. \qquad (3.4)$$

3.5. Оценки на траектории марковской игры

Зафиксируем некоторую пару рандомизированных стратегий $\bar{\mu} = (\mu,\phi_\mu)$, $\bar{\nu} = (\nu,\psi_\nu)$, а значит распределение $\check{\mathbb{P}} \overset{\triangle}{=} \check{\mathbb{P}}_{\bar{\mu},\bar{\nu}}[z_0]$ и меру Леви $\eta(z;\cdot) \overset{\triangle}{=} \check{\eta}(z,\bar{\mu},\bar{\nu};\cdot)$. Такой мере Леви соответствует генератор Леви–Хинчина (см., например, [18,(5.1)] и [22,(2.14)]), сопоставляющий каждой функции $g \in C_c^2(\mathcal{Z}_<)$ отображение $x \mapsto \Lambda g(x)$ правилом

$$\check{\Lambda}g(x) = \int_{\mathbb{X}}[g(x+y)-g(x)]\eta(x;dy) \qquad \forall x \in \mathcal{Z}_<. \qquad (3.5)$$

При этом имеет место формула Дынкина [22, Proposition 2.3]:

$$\check{\mathbb{E}}\Big(g(\hat{Y}(t'')) - \int_{t'}^{t''} \Lambda g(\check{Y}(s))\,ds \,\Big|\, \check{Y}(t')\Big) = g(\check{Y}(t')) \tag{3.6}$$

для всех $g \in C_c^2(\mathcal{Z}_<)$ и всех моментов времени t', t'' $(t'' \geq t')$. Тогда является мартингалом процесс

$$g(\check{Y}(t'')) - \int_{t'}^{t''} \Lambda g(\check{Y}(s))\,ds.$$

Рассмотрим поведение марковской цепи если $t' = 0$ и известно начальное состояние $w' \stackrel{\triangle}{=} \check{Y}(0)$. Примем $M \stackrel{\triangle}{=} \sqrt{hL\sqrt{d} + L^2}$. Применяя для мартингала выше в качестве g функцию

$$w \mapsto ||w - w'||_d^2 \stackrel{\triangle}{=} \sum_{i=1}^{d} (\pi_i w - \pi_i w')^2 \qquad \forall w \in \mathbb{R}^{d+2},$$

из оценок (3.1) можно вывести (см. [8,(4.14)]).

$$\check{\mathbb{E}}||\check{Y}(t) - \check{Y}(0)||_d^2 \leq hL\sqrt{d}t + \frac{4LM}{3}(e^t - 1)^{3/2}.$$

Теперь, в силу $e^r - 1 \leq re^r$ при $r \geq 0$, имеет место $(e^r - 1)^{3/2} \leq \sqrt{r}(e^{3r/2} - e^{r/2})$ и

$$\check{\mathbb{E}}||\check{Y}(t) - \check{Y}(0)||_d^2 \leq hL\sqrt{d}t + \frac{4LM\sqrt{t}}{3}(e^{3t/2} - e^{t/2}).$$

В частности, для момента остановки $\tau = \inf\{t \geq 0 \,|\, \pi_0\check{Y}(t) \neq \pi_0\check{Y}(0)\}$, в силу его экспоненциальной, с показателем $1/h$, плотности, а также с учетом равенства $\int_0^\infty \sqrt{r}e^{-\nu r}dr = \sqrt{\frac{\pi}{4\nu^3}}$, имеем

$$\begin{aligned}\check{\mathbb{E}}||\check{Y}(\tau) - \check{Y}(0)||_d^2 &\leq hL\sqrt{d}\check{\mathbb{E}}\tau + \frac{4LM}{3}\check{\mathbb{E}}\int_0^\infty \sqrt{s}(e^{3s/2} - e^{s/2})e^{-s/h}\,ds = \\ &= h^2L\sqrt{d} + \frac{2LM\sqrt{\pi}}{3}\big((1/h - 3/2)^{-3/2} - (1/h - 1/2)^{-3/2}\big) \leq \\ &\leq h^2L\sqrt{d} + LM\sqrt{\pi}(1/h - 3/2)^{-5/2} \leq h^2L(\sqrt{d} + 1);\end{aligned}$$

на последнем шаге мы также использовали условие $3h < 2^5M^2\pi h < 1$ (см. (2.2)). Аналогично, из равенства $\int_0^\infty r^{3/2}e^{-\nu r}dr = \frac{3}{4}\sqrt{\frac{\pi}{\nu^5}}$ также следует

$$\check{\mathbb{E}}\int_0^\tau ||\check{Y}(t)-\check{Y}(0)||_d^2\,dt \leq$$
$$\leq \int_0^\infty \int_0^s \Big[hL\sqrt{d}t + \frac{4LM\sqrt{t}}{3}(e^{3t/2}-e^{t/2})\Big]\,dt\,e^{-s/h}\,ds \leq$$
$$\leq hL\sqrt{d}\int_0^\infty \frac{s^2}{2}e^{-s/h}\,ds + \frac{4LM}{3}\int_0^\infty s^{3/2}(e^{3s/2}-e^{s/2})e^{-s/h}\Big]\,dt\,ds \leq$$
$$\leq h^3L\sqrt{d} + LM\sqrt{\pi}((1/h-3/2)^{-5/2}-(1/h-1/2)^{-5/2}) \leq$$
$$\leq h^3L\sqrt{d} + \frac{5LM\sqrt{\pi}}{2}(1/h-3/2)^{-7/2} \leq h^3L(\sqrt{d}+5).$$

Поскольку стратегия стационарна, а рассматриваемый процесс является сильно марковским, те же оценки имеют место, если рассматривать вместо начального момента времени произвольный момент остановки τ', в качестве τ взять любой момент остановки $\tau'' \geq \tau'$, не превосходящий

$$\inf\{t \geq 0 \,|\, \pi_0\check{Y}(t) \neq \pi_0\check{Y}(\tau'), t \geq \tau'\},$$

и, наконец, заменить математическое ожидание на условное матожидание в момент остановки τ'. Таким образом,

$$\check{\mathbb{E}}_{\tau'}||\check{Y}(\tau'')-\check{Y}(\tau')||_d^2 \leq h^2L(\sqrt{d}+1),$$
$$\check{\mathbb{E}}_{\tau'}\int_{\tau'}^{\tau''} ||\check{Y}(t)-\check{Y}(\tau')||_d^2\,dt \leq h^3L(\sqrt{d}+5). \qquad (3.7)$$

4. Схема поводыря и двойная игра

4.1. Нацеливание

При построении схемы с поводырем нам нужно уметь нацеливаться в заданном направлении (и отклоняться от него). Введем необходимые для этого функции сначала в детерминированной системе. Для этого заметим, что, благодаря (1.2), для любых векторов $x, z \in \mathbb{R}^d$ найдутся управления $u^{to\ z}(x) \in \mathbb{U}$, $v^{from\ z}(x) \in \mathbb{V}$ такие, что

$$\min_{u\in\mathbb{U}}\max_{v\in\mathbb{V}}\langle x-z, \hat{f}(x,u,v)\rangle = \max_{v\in\mathbb{V}}\langle x-z, \hat{f}(t,x,u^{to\ z}(x),v)\rangle =$$
$$= \max_{v\in\mathbb{V}}\min_{u\in\mathbb{U}}\langle x-z, \hat{f}(x,u,v)\rangle = \min_{u\in\mathbb{U}}\langle x-z, \hat{f}(x,u,v^{from\ z}(x))\rangle.$$

Для удобства примем также $u^{to\ w'} \equiv u^{to\ z'}$ для всех $w' = (t, z', s) \in \mathbb{R}^{m+2}$. Аналогично, для каждого $x' \in \mathbb{R}^d$ снабдим второго игрока стационарной рандомизированной стратегией $\bar{\nu}^{from\ x'}$:

$$\mathcal{Z}_< \ni w \stackrel{\triangle}{=} (t, z, s) \mapsto \bar{\nu}^{from\ x'}(w) \stackrel{\triangle}{=} (v^{from\ x'}(z), \psi^{opt}[w]).$$

При этом для всех $x, z \in \mathbb{R}^d$, $w = (\tau, z, 0) \in \mathbb{R}^{d+2}$, в силу $2L$-липшицевости $\hat{f}$ и $\check{f}$, выполнено

$$\begin{gathered}\langle x - w, f(x, u^{to\ w}(x), v) - \check{f}(w, \mu^{opt}(w), \nu^{from\ x}(w))\rangle_d \leq \\ \leq \max_u \left\langle x - w, f(x, u, v) - f(x, u, v^{from\ x}(z))\right\rangle_d + \\ + \max_{\bar{\mu}} \left\langle x - w, \check{f}(w, \delta_{u^{to\ w}(x)}, \nu^{from\ x}(w)) - \check{f}(w, \bar{\mu}, \nu^{from\ x}(w))\right\rangle_d + \\ +2L||x - w||_d^2 \leq 2L||x - w||_d^2. \end{gathered} \tag{4.1}$$

Зафиксируем некоторые $x' \in \mathbb{R}^d$, $w' \in \mathbb{R}^{d+2}$. Рассмотрим независящее от нулевых и последних координат аргументов отображение $R_{x',w'} : \mathbb{R}^{d+2} \times \mathbb{R}^{d+2} \to \mathbb{R}$, заданное правилом: для всех $x \in \mathcal{Z}_<$

$$\hat{R}_{x',w'}(x) = \sum_{i=1}^{d} (\pi_i x' - \pi_i w')(\pi_i x - \pi_i x') = \langle x' - w', x - x'\rangle_d; \tag{4.2}$$

в качестве аргумента у этой функции мы можем использовать и элементы из $\mathbb{R}^d$ в силу вложения $\mathbb{R}^d \subset \mathbb{R}^{d+2}$. Прямым подсчетом генератора, воспользовавшись $2L$-липшицевостью $\check{f}$, получаем (см. [8,(5.2)]) что в случае $x \in \mathbb{R}^d$, $w \in \mathbb{R}^{d+2}$ выполнено

$$\begin{gathered}\check{\Lambda}[\bar{\mu}^{opt}, \bar{\nu}^{from\ x'}]R_{w',x'}(w) = -\langle x' - w', \check{f}(w, \mu^{opt}, \nu^{from\ x'})\rangle \leq \\ \leq -\langle x' - w', \check{f}(w', \mu^{opt}, \nu^{from\ x'})\rangle_d + 2L||x' - w'||_d \cdot ||w - w'||_d \leq \\ \leq -\langle x' - w', \check{f}(w', \mu^{opt}, \nu^{from\ x'})\rangle_d + L||x' - w'||_d^2 + L||w - w'||_d^2. \end{gathered} \tag{4.3}$$

Такое представление нам будет полезно при подстановке генератора в формулу Дынкина (3.6).

Рассматривая случайный процесс уже в исходной игре, можно проверить (см. [8,(5.4)]), что при произвольной стратегии второго игрока v, для генератора

$$K_< \ni x \mapsto \hat{\Lambda}[u^{to\ w'}(x'), v]R_{x',w'}(x) = \langle x' - w', f(x, u^{to\ w'}(x'), v)\rangle_d \tag{4.4}$$

имеет место аналог формулы Дынкина: для любых моментов остановки τ', τ'' ($\tau' \leq \tau''$)

$$R_{x',w'}(y(\tau')) = \mathbb{E}_{\tau'}\Big[R_{x',w'}(y(\tau'')) - \int_{\tau'}^{\tau''} \hat{\Lambda}[u^{to\ w'}(x'), v] R_{x',w'}(y(t))\,dt\Big]. \tag{4.5}$$

В частности, применяя к (4.4) $2L$-липшицевость f, оценивая произведение полусуммой квадратов и воспользовавшись (4.1), получаем

$$\hat{\Lambda}[\bar{u}^{to\ w'}(x'), \bar{v}(t)] R_{x',w'}(x) \leq$$
$$\leq \langle x' - w', f(x', u^{to\ w'}(x'), v[t])\rangle_d + 2L||x - x'||_d \cdot ||x' - w'||_d \overset{(4.1)}{\leq}$$
$$\overset{(4.1)}{\leq} \langle x' - w', \check{f}(w', \mu^{opt}, v^{from\ x'})\rangle_d + 3L||x' - w'||_d^2 + L||x - x'||_d^2. \tag{4.6}$$

4.2. Конструкция поводыря для первого игрока

Сейчас нами будет построена конструкция случайного процесса с càdlàg траекториями в фазовом пространстве $K_< \times \mathcal{Z}_<$; мы будем строить этот процесс вплоть до момента окончания игры. При этом мы будем считать, что некоторые допустимые стратегия V и момент выхода θ_2 выбраны вторым игроком, но неизвестны первому. Для каждого выбора V на траекториях в фазовом пространстве $K_< \times \mathcal{Z}_<$ будет получаться свое вероятностное распределение, фиксируя выбор V, мы фиксируем и распределение $\widetilde{\mathbb{P}}$, и соответствующее ему матожидание $\widetilde{\mathbb{E}}$. Проекция $\hat{Y}$ этого распределения с траекториями со значениями $K_< \times \mathcal{Z}_<$ на первую компоненту даст распределение $\mathbb{P}_{U,V}$ траекторий y исходной игры и, действительно, как мы увидим позже, будет порождаться некоторой допустимой стратегией U первого игрока. Проекция $\check{Y}$ этого распределения на вторую компоненту будет ответственна за распределение поводыря и будет строиться первым игроком на основе распределений $\check{\mathbb{P}}$, порожденных стационарными стратегиями $(\mu^{opt}, \nu^{from\ y(t_k)})$.

Итак, пусть в заранее заданные моменты времени $(t_k = kh)_{k\in\mathbb{N}}$ первый игрок считывает позицию движения $y(\cdot)$ реальной игры, можно также считать, что $t_0 = 0$. Определим для случайного процесса $\check{Y}$ моменты остановки $\tau_k \overset{\triangle}{=} \min\{t\,|\,\pi_0\check{Y}(t) = kh\}$. Как отмечено в Замечании 3.1, по построению марковской цепи каждый из этих моментов остановки почти всюду конечен, более того, каждый, являясь суммой независимых в совокупности, распределенных экспоненциально с параметром $1/h$ случайных величин, эти моменты никак не зависят от действий игроков. Более того, всю их последовательность можно даже считать сгенерированной до начала игры. Тем самым нам уже известна нулевая координата модели: $\pi_0\check{Y}(t) \overset{\triangle}{=} |\{i \in \mathbb{N} \cup \{0\} \mid \tau_i \le t\}|h$.

Для генерации до момента остановки $\check{\theta}_{\min}$ с первой по d-ю координату модели будем пользоваться проекцией той же меры $\check{\eta}$, но лишь на первые d координат:

$$\eta(w, \bar{\mu}, \bar{\nu}; A) \overset{\triangle}{=} \frac{1}{h}\sum_{i=1}^{d}\left[\pi_i^+ \check{f}(w, \mu, \nu)\delta_{h\pi_i} + \pi_i^- \check{f}(w, \mu, \nu)\delta_{-h\pi_i}\right],$$

что соответствует матрице Колмогорова $(Q_{wy}(\bar{\mu}, \bar{\nu}))_{w,y\in\mathcal{Z}_<}$, введенной правилом

$$\begin{cases} \pi_i^+ \check{f}(w, \mu, \nu)/h, & y = w + h\pi_i, i \in [1:d], \\ \pi_i^- \check{f}(w, \mu, \nu)/h, & y = w - h\pi_i, i \in [1:d], \\ -\sum_{i=1}^{d}(\pi_i^+ \check{f}(w, \mu, \nu) + \pi_i^- \check{f}(w, \mu, \nu))/h & y = w, \\ 0, & \text{иначе.} \end{cases}$$

Опишем, наконец, для первого игрока на каждом промежутке $[t_{k-1}; t_k)$ процедуру управления исходной конфликтно-управляемой системой, если игра в момент t_k еще не завершена и им уже построен случайный процесс Y до момента остановки τ_k.

В момент времени $t_0 = 0$ игрок инициализирует случайный процесс Y, присваивая $\hat{Y}(0) = y(0) = x_*$, $\check{Y}(0) = w_0$, где w_0 – ближайшая к $(0, y(0), 0)$ точка с решетки $h\mathbb{R}^{d+2}$, и назначает на промежутке $[0; h)$ управление $u^{to\ w_0}$. Затем первый игрок рассчитывает траекторию $\check{Y}$ процесса с мерой Леви η в силу пары стратегий $(\mu^{opt}, \nu^{from\ y(0)})$ вплоть до момента остановки τ_1, после чего назначает $w_1 = \check{Y}(\tau_1)$. Поскольку, по условию, выход из игры в начальный момент времени невозможен, то единственное возможное для первого (и для второго) игрока значение при этом для ϕ_0 (и для ψ_0) равно нулю.

Как управляется второй игрок на промежутке $[0; h)$ первому игроку неизвестно, но если игра еще не закончилась, то в момент времени $t_1 = h$ первый игрок, зная позицию $y(t_1) = \hat{Y}(t_1)$, в исходной игре на промежутке $[h; 2h)$ назначает $u^{to\ w_1}$ в качестве своего управления, а для модели $\check{Y}$ на промежутке $[\tau_1; \tau_2)$ за обоих игроков назначает стационарные стратегии $(\mu^{opt}, \nu^{from\ y(t_1)})$. Он также примет $\phi_1 = \phi^{opt}(\check{Y}(\tau_1)) \in \Phi_1$ в качестве вероятности инициированного им завершения исходной игры.

Итак, пусть построены управление в исходной игре и траектория марковского процесса до момента времени t_k и момента остановки τ_k соответственно. Покажем, как управляется первый игрок до момента времени t_{k+1} и момента остановки τ_{k+1} в том случае, если исходная игра не завершилась на промежутке $[t_{k-1}; t_k)$ и траектория исходной игры известна до момента t_k включительно.

В момент времени $t_k = kh$ первый игрок, зная позиции $y(t_k) = \hat{Y}(t_k)$ и $w_k \stackrel{\triangle}{=} \check{Y}(\tau_k)$, назначает в исходной игре на промежутке $[t_k; t_{k+1})$ функцию $u^{to\ w_k}$ в качестве своего управления, а для модели на промежутке $[\tau_k; \tau_{k+1})$ за обоих игроков назначает стационарные стратегии $(\mu^{opt}, \nu^{from\ y(t_k)})$. Он также примет $\phi_k = \phi^{opt}(Y(\tau_k))$. Как и раньше, он будет считать, что только выход в исходной игре на протяжении $[t_k; t_{k+1})$ означает выход из модельной игры, тем самым он попрежнему не нуждается в информации о вероятности инициированного вторым игроком выхода ни для расчета марковской цепи вплоть до момента остановки τ_{k+1}, ни для траектории исходной игры до момента времени t_{k+1}. Траектории до момента времени t_{k+1} и момента остановки τ_{k+1} построены, а следовательно предположение индукции показано.

Итак, нами описано как построение траектории $y(\cdot) = \hat{Y}(\cdot)$ исходной игры вплоть до момента ее окончания $\theta_{\min}$, так и построение всех, кроме последней, координат модели $\check{Y}$ до момента времени $\check{\theta}_{\min}$. Сделаем это и для последней координаты.

Как бы не развивалась исходная игра, она может быть закончена досрочно только в один из моментов t_n^I, t_n^{II}, а на каждом промежутке $[t_k; t_{k+1})$ их не более одного. Тогда можно считать, что исходная игра завершилась в момент $\theta_{\min}$ из $[t_k; t_{k+1})$ в точности тогда, когда модельная игра будет закончена на промежутке $[\tau_k; \tau_{k+1})$, то есть осуществится одно из событий $\check{\theta}_i \in [\tau_k; \tau_{k+1})$. В силу Замечания 3.2, из одинаковых со стороны каждого игрока вероятностей выхода в этих играх, вероятность этого события совпадёт с вероятностью переноса по последней координате в марковской игре. Тогда на каждом промежутке времени $[0;\tau_k)$ мы имеем возможность восстановить последнюю координату модели: $\pi_{d+1}\check{Y}(t) \stackrel{\triangle}{=} 0$ вплоть до момента $\check{\theta}_{\min}$ завершения модельной игры, $\pi_{d+1}\check{Y}(t) \stackrel{\triangle}{=} 1+\check{\theta}_{\min}$ для $t \geq \check{\theta}_{\min} = \check{\theta}_1$ и $\pi_{d+1}\check{Y}(t) \stackrel{\triangle}{=} -1-\check{\theta}_{\min}$ для $t \geq \check{\theta}_{\min} = \check{\theta}_2$. С другой стороны, первый игрок, задавая вероятности $\phi_k = \phi^{opt}(\check{Y}(\tau_k))$, тем самым определил не только момент остановки $\check{\theta}_1$ марковской игры, но, в силу того же замечания, задал и момент выхода θ_1 в исходной игре.

Таким образом для каждого выбора стратегии и момента выхода вторым игроком на траекториях процесса $(\hat{Y}, \check{Y})$ в фазовом пространстве $K_< \times \mathcal{Z}_<$ построено вероятностное распределение $\tilde{\mathbb{P}}$, и соответствующее ему матожидание $\tilde{\mathbb{E}}$. Дезынтегрируя его сначала по $\bar{\phi}^{opt}(\check{Y}|_{[0;\tau_{k-1}]})$, а затем по $y|_{[0;t_{k-1}]} = \hat{Y}|_{[0;t_{k-1}]}$ и $u^{\to y()}|_{[0;t_{k-1}]}$, мы получаем распределение управлений первого игрока вплоть до момента времени t_{k+1}, тем самым предъявленная в этом параграфе стратегия допустима. Аналогично, каждая вероятность ϕ_k выхода первым(вторым) игроком, определяется как маргинальная вероятность при уже известной, реализовавшейся части траектории и управления, также является допустимой.

Отметим, что указанная процедура построения допустимых стратегии U и момента выхода θ_1 нигде не использовала временных предположений. Действительно, нас нигде не интересовало в какой именно момент из промежутка $[t_k = kh; t_{k+1} = kh+h)$ может закончиться исходная игра, нас интересовало лишь закончилась ли она на этом промежутке. Далее, поданная нами вероятность $\phi^{opt}(Y(\tau_k))$ заведомо попадает в нужный компакт по конструкции Φ_k, а поскольку при этом используется лишь информация, имеющаяся в момент времени t_k, а значит заведомо имеющаяся в любой момент из промежутка $[t_k; t_{k+1})$, то полученный при этом момент остановки θ_1 допустим. Наконец, указанная процедура использовала информацию о том, остановилась ли игра, лишь в конце каждого промежутка $[t_k; t_{k+1})$. В частности, она не запрещала второму игроку принять решение о своей очередной вероятности завершить игру в течение промежутка $[t_k; t_{k+1})$ лишь в момент t_{k+1}. Тем самым, эта процедура допускает за второго игрока не только все допустимые моменты остановки, но и все моменты остановки θ_2 второго игрока, для которых при всех

$n \in \mathbb{N}$, $s \in [0;1)$ вероятности событий $\{\theta_2 < nh + sh\}$ лишь $\mathcal{F}_{nh}$-измеримы. Зафиксируем момент остановки θ_2 с таким свойством.

4.3. Расхождение траекторий исходной и марковской игр

Введем моменты остановки $\theta_{\min} = \min(\theta_1, \theta_2)$ и $\check{\theta}_{\min} = \min(\check{\theta}_1, \check{\theta}_2)$ в исходной и в марковской игре соответственно.

Зафиксируем некоторое $k \in \mathbb{N}$, а с ним моменты $t_{k-1} = (k-1)h, t_k = kh \le T$ из разбиения. Теперь мы можем определить и моменты остановки

$$t' \stackrel{\triangle}{=} \min(t_{k-1}, \theta_{\min}),\ t'' \stackrel{\triangle}{=} \min(t_k, \theta_{\min}).$$

$$\tau' \stackrel{\triangle}{=} \min(\tau_{k-1}, \check{\theta}_{\min}),\ \tau'' \stackrel{\triangle}{=} \min(\tau_k, \check{\theta}_{\min}).$$

Обозначим через x' позицию $y(t')$, через w' – точку с координатами $(\pi_1\check{Y}(\tau'), \ldots, \pi_d\check{Y}(\tau'))$. В частности, для всех t из $\tau' \le t < \tau''$ следует $\pi_0\check{Y}(t) = (k-1)h$; кроме того, имеют место оценки (3.7).

Дезынтегрируем вероятность $\tilde{\mathbb{P}}$ по t', τ'. Обозначим условное матожидание относительно сигма-алгебры $\mathcal{F}_{t'} \otimes \mathcal{F}_{\tau'}$ через $\widetilde{\mathbb{E}}_{t',\tau'}$; отметим, что пара (w', x') $\mathcal{F}_{t'} \otimes \check{\mathcal{F}}_{\tau'}$-измерима.

Из определений $R_{x',w'} = \langle \cdot - x', x' - w' \rangle_d, R_{w',x'}$ и равенства

$$||w - x||_d^2 - ||w - x + x' - w'||_d^2 = 2R_{x',w'}(x) + 2R_{w',x'}(w) + ||x' - w'||_d^2$$

для всех $w, x \in \mathbb{R}^{d+2}$ следуют неравенства

$$\frac{||x - w||_d^2 - ||x' - w'||_d^2}{2} \le ||x - x'||_d^2 + ||w - w'||_d^2 + R_{x',w'}(x) + R_{w',x'}(w).$$

и

$$\begin{aligned} \widetilde{\mathbb{E}}_{t',\tau'}||y(t'') - \check{Y}(\tau'')||_d^2 - ||x' - w'||_d^2 \le 2\widetilde{\mathbb{E}}_{t',\tau'}||\check{Y}(\tau'') - w'||_d^2 + \\ +2\widetilde{\mathbb{E}}_{t',\tau'}||y(t'') - x'||_d^2 + 2\widetilde{\mathbb{E}}_{t',\tau'}R_{x',w'}(y(t'')) + 2\widetilde{\mathbb{E}}_{t',\tau'}R_{w',x'}(\check{Y}(\tau'')). \quad (4.7) \end{aligned}$$

Первое слагаемое правой части уже оценено в (3.7):

$$2\widetilde{\mathbb{E}}_{t',\tau'}||\check{Y}(t'') - w'||_d^2 = 2\check{\mathbb{E}}_{\tau'}||\check{Y}(\tau'') - w'||_d^2 \le 2h^2L(\sqrt{d} + 1). \quad (4.8)$$

Второе слагаемое в правой части (4.7), в силу ограниченности динамики (1.1), легко оценивается:

$$2\widetilde{\mathbb{E}}_{t',\tau'}||y(t'') - x'||_d^2 \le 2L^2\widetilde{\mathbb{E}}_{t',\tau'}(t'' - t')^2 \le 2L^2h^2. \quad (4.9)$$

Для третьего слагаемого в правой части (4.7), воспользовавшись формулой Дынкина (4.5), из (4.6), $||y(s) - x'||_d^2 \leq L(s - t')$ и $t'' - t' \leq h$ получаем, что

$$2\widetilde{\mathbb{E}}_{t',\tau'} R_{x',w'}(y(t'')) = 2\widetilde{\mathbb{E}}_{t',\tau'}[R_{x',w'}(y(t'')) - R_{x',w'}(y(t'))] \leq$$
$$\leq 2h\widetilde{\mathbb{E}}_{t',\tau'}\langle x' - w', \check{f}(x', \mu^{opt}, v^{from\ x'})\rangle_d +$$
$$+6L||x' - w'||_d^2 + 4L\mathbb{E}_{t'} \int_{t'}^{t''} ||y(s) - x'||_d^2\, ds \leq$$
$$\leq 2\langle x' - w', \check{f}(x', \mu^{opt}, v^{from\ x'})\rangle_d h + 6L||x' - w'||_d^2 h + 4L^3h^3/3. \quad (4.10)$$

Последнее слагаемое в (4.7), в силу формулы Дынкина (3.6), оценок (4.3), (3.7) и $\widetilde{\mathbb{E}}_{t',\tau'}(\tau'' - \tau') \leq \check{\mathbb{E}}_{\tau'}(\tau_k - \tau_{k-1}) = h$, может быть оценено через

$$2\widetilde{\mathbb{E}}_{t',\tau'} R_{w',x'}(\check{Y}(\tau'')) \leq 2\widetilde{\mathbb{E}}_{t',\tau'} \int_{\tau'}^{\tau''} \hat{\Lambda}[u^{to\ w'}(x'), v(s)] R_{x',w'}(\check{Y}(s))\, ds \leq$$
$$\leq 2(L||x' - w'||_d^2 - \langle x' - w', \check{f}(x', \mu^{opt}, \nu^{from\ x'})\rangle_d)\check{\mathbb{E}}_{\tau'}(\tau'' - \tau') +$$
$$+2L \int_{\tau'}^{\tau''} \check{\mathbb{E}}_{\tau'}||\check{Y}(s) - w'||_d^2\, ds \overset{(3.7)}{\leq}$$
$$\overset{(3.7)}{\leq} 2L||x' - w'||_d^2 h - 2\langle x' - w', \check{f}(x', \mu^{opt}, \nu^{from\ x'})\rangle_d h + 2h^3L^2(\sqrt{d} + 5). \quad (4.11)$$

Подставляя оценки (4.9)–(4.11) в (4.7), воспользовавшись $L\sqrt{d}h < 1$ и $Lh(2L/3 + 5) < 1$ (см. (2.1)), получаем, что

$$\widetilde{\mathbb{E}}_{t',\tau'}||y(t'') - \check{Y}(\tau'')||_d^2 \leq$$
$$\leq (1 + 8Lh)||x' - w'||_d^2 + 2h^2L(\sqrt{d} + 1 + 2L^2h/3 + L\sqrt{d}h + 5Lh) \overset{(2.1)}{\leq}$$
$$\overset{(2.1)}{\leq} ||y(t') - \check{Y}(\tau')||_d^2(1 + 8Lh) + 6h^2L(1 + \sqrt{d}).$$

Введем для краткости при всех натуральных i моменты остановки: $t_i^\theta \overset{\triangle}{=} \min(t_i, \theta_{\min})$, $\tau_i^\theta \overset{\triangle}{=} \min(\tau_i, \check{\theta}_{\min})$. Теперь из полученного следует при $i \leq T/h$

$$\widetilde{\mathbb{E}}_{t_{i-1}^\theta, \tau_{i-1}^\theta}||y(t_i^\theta) - \check{Y}(\tau_i^\theta)||_d^2 \leq ||y(t_{i-1}^\theta) - \check{Y}(\tau_{i-1}^\theta)||_d^2\, e^{8Lh} + 6Lh^2(1 + \sqrt{d}).$$

Итерируя (4.3) и при всех значениях i, меньших $k \triangleq T/h$, получаем

$$\widetilde{\mathbb{E}}_{t^\theta_{k-1},\tau^\theta_{k-1}}||y(t^\theta_k)-\check{Y}(\tau^\theta_k)||^2_d \le$$
$$\le \widetilde{\mathbb{E}}_{t^\theta_{k-2},\tau^\theta_{k-2}}||y(t^\theta_{k-1})-\check{Y}(\tau^\theta_{k-1})||^2_d\, e^{2\cdot 8Lh} + 6Lh^2(1+\sqrt{d})(1+e^{8Lh}) \le$$
$$\le \widetilde{\mathbb{E}}||y(t_0)-\check{Y}(\tau_0)||^2_d\, e^{8kLh} + 6Lh^2(1+\sqrt{d})\sum_{i=0}^{k-1} e^{8iLh} \le$$
$$\le \Big(\frac{hd}{2} + \frac{6Lh^2(1+\sqrt{d})}{8Lh}\Big)e^{8LT} \le 2hde^{8LT}.$$

Здесь, помимо $||y(t_0)-\check{Y}(\tau_0)||^2_d = ||x_0-x_*||^2 \le hd/2$, на последнем шаге использовалось также $1+\sqrt{d}\le 2d$. Таким образом, показано

$$\widetilde{\mathbb{E}}||y(\min(T,\theta_{\min}))-\check{Y}(\min(\tau_{T/h},\check{\theta}_{\min}))||_d \le e^{4LT}\sqrt{2d}\sqrt{h}. \qquad (4.12)$$

4.4. Оценка разности платежей

Теперь наша задача оценить разность $J(y,\theta_1,\theta_2)$ и $\check{\sigma}(\check{Y}(\tau_{T/h}))$. Рассмотрим для этого три случая: $\theta_{\min}=\theta_1\le T$, $\theta_{\min}=\theta_2\le T$ и $\theta_{\min}>T$.

В первом случае $\theta_{\min}=\theta_1\le T$, тогда выполнено $J(y,\theta_1,\theta_2) = \sigma_1(y(\theta_{\min}))$ и $\check{\theta}_{\min}\le\tau_{T/h}$, с учетом

$$\check{\sigma}(\check{Y}(T/h)) = \check{\sigma}(\check{Y}(\theta_{\min})) = \sigma_1(\pi_0\check{Y}(\theta_{\min}),\ldots,\pi_d\check{Y}(\theta_{\min})),$$

имеем

$$|J(y,\theta_1,\theta_2)-\check{\sigma}(\check{Y}(\tau_{T/h}))| = |\sigma_1(y(\theta_{\min}))-\sigma_1(\pi_1\check{Y}(\theta_{\min}),\ldots,\pi_d\check{Y}(\theta_{\min}))|.$$

Теперь, в силу L-липшицевости σ_1 из (4.12) получаем

$$\widetilde{\mathbb{E}}\big(|J(y,\theta_1,\theta_2)-\check{\sigma}(\check{Y}(\tau_{T/h}))|\,\big|\,\theta_{\min}=\theta_1\le T\big) \le Le^{4LT}\sqrt{2d}\sqrt{h}.$$

Во втором случае аналогичная оценка следует из L-липшицевости σ_2.

Для третьего случая $\theta_{\min}>T$ отметим, что из условия (1.4) для σ_i, в силу выбора $T>N_{\varepsilon/12}$, следуют уже для $\check{\sigma}$ неравенства

$$\widetilde{\mathbb{P}}(\theta_{\min}\ge T)\,\widetilde{\mathbb{E}}\big(\check{\sigma}(T,y(T),0)\,\big|\,\theta_{\min}\ge T\big) - \varepsilon/12 <$$
$$< \widetilde{\mathbb{P}}(\theta_{\min}\ge T)\,\widetilde{\mathbb{E}}\big(J(y,\theta_1,\theta_2)\,\big|\,\theta_{\min}\ge T\big) <$$
$$< \widetilde{\mathbb{P}}(\theta_{\min}\ge T)\,\widetilde{\mathbb{E}}\big(\check{\sigma}(T,y(T),0)\,\big|\,\theta_{\min}\ge T\big) + \varepsilon/12.$$

Теперь из L-липшицевости σ_i имеем

$$\widetilde{\mathbb{P}}(\theta_{\min} \geq T)\widetilde{\mathbb{E}}\big(|J(y,\theta_1,\theta_2) - \check\sigma(\check Y(\tau_{T/h}))| \,\big|\, \theta_{\min} \geq T\big) < \\ < \varepsilon/12 + Le^{4LT}\sqrt{2d}\sqrt{h} \overset{(2.2)}{\leq} \varepsilon/6.$$

Объединяя все эти случаи, получаем

$$\widetilde{\mathbb{E}}|J(y,\theta_1,\theta_2) - \check\sigma(\check Y(\tau_{T/h}))| \leq \varepsilon/6 + Le^{4LT}\sqrt{2d}\sqrt{h} \overset{(2.2)}{\leq} \varepsilon/4.$$

Осталось отметить, что из (3.4) теперь следует

$$\check{\mathbb{E}}\Big|\int_0^\infty he^{-ht}\check\sigma(\check Y(t))\,dt - J(y,\theta_1,\theta_2)\Big| \leq \varepsilon/4 + \varepsilon/12 = \varepsilon/3.$$

Поскольку эта оценка не зависит от выбранной вторым игроком стратегии, поскольку вне зависимости от стратегии второго при выбранных первым игроком стратегиях μ^{opt}, ϕ^{opt} математическое ожидание $\widetilde{\mathbb{E}}\int_0^\infty he^{-ht}\check\sigma(\check Y(t))\,dt$ не больше $\check{\mathcal{V}}(0,z_0,0)$, то нами показано, что

$$\widetilde{\mathbb{E}}J(y,\theta_1,\theta_2) \geq \check{\mathcal{V}}(0,z_0,0) - \varepsilon/3.$$

Отметим, что $J(y,\theta_1,\theta_2)$ определена в исходной игре, тогда

$$\mathbb{E}_{\Omega,U,V}J(y,\theta_1,\theta_2) \geq \check{\mathcal{V}}(0,z_0,0) - \varepsilon/3.$$

Поскольку допустимые стратегия V и момент выхода θ_2 второго игрока были произвольны, то показано

$$\mathcal{V}_- \geq \check{\mathcal{V}}(0,z_0,0) - \varepsilon/3.$$

Рассматривая на основе той же марковской игры поводыря для второго игрока, получаем симметричную оценку $\mathcal{V}_+ \leq \check{\mathcal{V}}(0,z_0,0) + \varepsilon/3$, откуда

$$|\mathcal{V}_- - \mathcal{V}_+| \leq 2\varepsilon/3; \tag{4.13}$$

теперь построенная в разделе 4.2 пара (U,θ_1) является $2\varepsilon/3$-оптимальной в верхней игре, в частности для нее имеет место и (1.6). Итак, нужная нам теорема доказана, если помимо условия (1.4) выполнены и временные предположения.

5. Избавление от временных предположений

Теперь покажем как свести исходную игру к игре с такими предположениями. Для этого нам потребуется

Лемма 5.1. *Пусть у первого и второго игроков помимо множества допустимых стратегий,* $\mathfrak{U}$ *и* $\mathfrak{V}$ *соответственно, имеется по два класса моментов выхода,* $\mathfrak{Q}^I_\flat$, $\mathfrak{Q}^I_\sharp$ *и* $\mathfrak{Q}^{II}_\flat$, $\mathfrak{Q}^{II}_\sharp$ *соответственно. Пусть имеются отображения*

$$\mathfrak{Q}^I_\sharp \ni \theta_1 \mapsto \theta_1^\flat \in \mathfrak{Q}^I_\flat, \qquad \mathfrak{Q}^{II}_\flat \ni \theta_2 \mapsto \theta_2^\sharp \in \mathfrak{Q}^{II}_\sharp$$

и неотрицательные числа r', r'' *такие, что при любом выборе моментов выхода* $\theta_1 \in \mathfrak{Q}^I_\sharp$, $\theta_2 \in \mathfrak{Q}^{II}_\flat$ *и любом выборе стратегий игроков вероятности событий* $|\theta_1 - \theta_1^\flat| > r'$, $|\theta_2 - \theta_2^\sharp| > r'$ *не превосходят* r''.

Тогда выполнено неравенство $|\mathcal{V}^\flat_- - \mathcal{V}^\sharp_-| \leq 4r'' + Lr'$, *где* $\mathcal{V}^\sharp$ *– нижняя цена в игре, где допустимы моменты выхода из классов* $\mathfrak{Q}^I_\sharp$ *и* $\mathfrak{Q}^{II}_\sharp$, *а* $\mathcal{V}^\flat$ *– нижняя цена в игре, где допустимы моменты выхода из классов* $\mathfrak{Q}^I_\flat$ *и* $\mathfrak{Q}^{II}_\flat$.

Доказательство. Отметим сначала, что по условию при любом выборе допустимых стратегий U, V и допустимых моментов выхода $\theta_1 \in \mathfrak{Q}^I_\sharp$, $\theta_2 \in \mathfrak{Q}^{II}_\flat$ вероятность события $|\min(\theta_1, \theta_2^\sharp) - \min(\hat{\theta}_1^\flat, \hat{\theta}_2)| > r'$ не больше $2r''$. Поскольку J ограничено по модулю числом 1, то порожденное стратегиями U, V математическое ожидание у $|J(y, \theta_1, \hat{\theta}_2^\sharp) - J(y, \theta_1^\flat, \hat{\theta}_2)|$ не превосходит $4r'' + Lr'$.

Пусть допустимая стратегия U с моментом выхода $\theta_1 \in \mathfrak{Q}^I_\sharp$ является s-оптимальной парой в нижней игре с $\mathfrak{Q}^I_\sharp$ и $\mathfrak{Q}^{II}_\sharp$, то есть

$$\mathbb{E}_{\Omega,U,V} J(y, \theta_1, \theta_2) \leq \mathcal{V}^\sharp_- + s$$

для всякой допустимой стратегии V и допустимого момента выхода $\theta_2 \in \mathfrak{Q}^{II}_\sharp$. Тогда для всякой допустимой стратегии V и допустимого момента выхода $\theta_2 \in \mathfrak{Q}^{II}_\flat$ выполнено

$$\mathbb{E}_{\Omega,U,V} J(y, \theta_1^\flat, \theta_2) - 4r'' - Lr' \leq \mathbb{E}_{\Omega,U,V} J(y, \theta_1, \theta_2^\sharp) \leq \mathcal{V}^\sharp_- + s;$$

таким образом, поскольку $\theta_1^\flat \in \mathfrak{Q}^I_\flat$, то показано $\mathcal{V}^\flat_- \leq \mathcal{V}^\sharp_- + 4r'' + Lr' + s$. В силу произвольности положительного s неравенство $\mathcal{V}^\flat_- - \mathcal{V}^\sharp_- \leq 4r'' + Lr'$ также доказано. Симметрия $(\flat, 1) \leftrightarrow (\sharp, 2)$ влечет противоположное неравенство. Лемма показана. $\square$

5.1. Избавление от первого предположения

Ранее показано, что при временных предположениях имеет место (4.13) и для некоторых $2\varepsilon/3$-оптимальных U и θ_1 выполнено (1.6). Покажем, что если для исходной игры нашлась пара (T, h), удовлетворяющая только второму временному предположению, то построенная для нее в разделе 4.2 пара (U, θ_1) является $2\varepsilon/3 + 4h$-оптимальной в верхней игре, в частности выполнено (1.6), а кроме того

$$\mathcal{V}_- \geq \mathcal{V}_+ - 2\varepsilon/3 - 4h \overset{(2.1)}{\geq} \mathcal{V}_+ - \varepsilon. \tag{5.1}$$

Рассмотрим модифицированную игру, игру с той же динамикой, с той же целевой функцией и теми же последовательностями t^I_k и t^I_k, но для которой выполнено

$$\hat{\phi}^-_k \triangleq \min(1-h, \phi^-_k),\ \hat{\phi}^+_k \triangleq \min(1-h, \phi^+_k),$$
$$\hat{\psi}^-_l \triangleq \min(1-h, \psi^-_l),\ \hat{\psi}^+_k \triangleq \min(1-h, \psi^+_l)$$

в случае $t^I_k \leq T$, $t^{II}_l \leq T$ и $\hat{\phi}^-_k \triangleq \phi^-_k$, $\hat{\phi}^+_k \triangleq \phi^+_k$, $\hat{\psi}^-_k \triangleq \psi^-_k$, $\hat{\psi}^+_k \triangleq \psi^+_k$ в противном случае. Эта игра удовлетворяет первому предположению. Обозначим через $\mathfrak{Q}^I_\sharp$ и $\mathfrak{Q}^{II}_\sharp$ классы допустимых в модифицированной игре моментов выхода первого и второго игрока соответственно. Примем также $\mathfrak{Q}^I_\flat \triangleq \mathfrak{Q}^I$, $\mathfrak{Q}^{II}_\flat \triangleq \mathfrak{Q}^{II}$.

Рассмотрим теперь произвольный момент выхода $\theta_2 \in \mathfrak{Q}^{II} = \mathfrak{Q}^{II}_\flat$ второго игрока, допустимый в исходной игре. Он порожден принимающими значения из $[\psi^-_k; \psi^+_k]$ $\mathcal{F}_{t^{II}_{k+1}}$-измеримыми случайными величинами $\psi[k]$ на $D(\mathbb{R}_+, \mathbb{R}^d)$. Тогда для всех $k \in \mathbb{N}$ примем $\hat{\psi}[k] \triangleq \min(1-h, \psi[k])$. Это гарантирует значения случайных величин $\hat{\psi}[k]$ в $[\hat{\psi}^-_k; \hat{\psi}^+_k]$, при этом $|\psi[k] - \hat{\psi}[k]| \leq h$ Теперь правила $\hat{\psi}[k]$ восстанавливают уже допустимый в модифицированной игре момент выхода $\theta^\sharp_2 \in \mathfrak{Q}^{II}_\sharp$.

Теперь рассмотрим допустимый в модифицированной игре момент выхода $\hat{\theta}_1 \in \mathfrak{Q}^I_\sharp$ первого игрока и соответствующую последовательность случайных величин $\hat{\phi}[k]$. Заметим, что значение $\hat{\phi}[k]$ вне промежутка $[\phi^-_k; \phi^+_k]$ возможно только при $\hat{t}^I_{k+1} < T$ и либо $1-h \leq \phi^-_k$, либо $h \geq \phi^-_k$; отсюда, во выбору h, выполнено $\phi^-_k = \phi^+_k = 1$, $\hat{\phi}^-_k = 1-h$. Тогда примем $\phi[k] = \hat{\phi}[k]$ при $\hat{\phi}[k] \in [\phi^-_k; \phi^+_k]$ и $\phi[k] = \phi^+_k$ иначе; это гарантирует $|\hat{\phi}[k]-\phi[k]| \leq h$. Эти правила восстанавливают допустимый в исходной игре для первого игрока момент выхода $\hat{\theta}_1^\flat \in \mathfrak{Q}_\flat^I$.

В силу одинаковой динамики множества $\mathfrak{D}^I$, $\mathfrak{D}^{II}$ являются множествами всех допустимых стратегий и в исходной, и в модифицированной игре. В частности, любой выбор допустимой пары $(U,V) \in \mathfrak{D}^I \times \mathfrak{D}^{II}$ задает одну и ту же вероятность на всевозможных траекториях в каждой из игр. По построению, вероятности событий $\theta_1 \neq \theta_1^\flat$, $\theta_2 \neq \theta_2^\sharp$, вслед за событиями $|\psi[k]-\hat{\psi}[k]| \leq h$, не больше h.

Все условия леммы выполнены для $r'' = h$ и $r' = 0$, следовательно нижние цены исходной и модифицированной игр отличаются не более чем на $4h$; теперь если исходная игра удовлетворяет второму предположению, то модифицированная игра удовлетворяет обоим, а для ее нижней цены имеет место (4.13); значит для нижней цены исходной игры доказано (5.1).

Более того, если исходная игра удовлетворяет второму предположению, то прописанные в разделе 4.2 стратегия U и момент выхода θ_1 совпадают с построенными по той же процедуре и в модифицированной игре. В частности, она допустимы и $2\varepsilon/3+4h$-оптимальны для верхней модифицированной игры. Следовательно в исходной задаче они $2\varepsilon/3+8h$-оптимальны, в частности, удовлетворяют (1.6). Итак, (5.1) и (1.6) показаны, если исходная игра удовлетворяет второму предположению.

5.2. Избавление от второго предположения

Нам для начала потребуются на $\mathbb{R}_+$ две унарные операции: $\lceil\cdot\rceil$ и $\lfloor\cdot\rfloor$. Пусть при каждом $R \in [0;T/h]$ $\lceil R\rceil$ означает наименьший элемент решетки $h\mathbb{N}$, больший R; $\lfloor R\rfloor$ означает наибольший элемент решетки $h\mathbb{N}$, небольший R. При каждом при $R \notin [0;T/h]$ положим $\lceil R\rceil = \lfloor R\rfloor = R$. Более того, поскольку любой подынтервал отрезка $[0;T]$ длины $2h$ не содержит более одного элемента последовательностей $(t_k^I)_{k\in\mathbb{N}}$, $(t_k^{II})_{k\in\mathbb{N}}$, на $[T-h;T+h]$ таких элементов также нет, то из $0 < |\lceil t\rceil - t|, \lfloor t\rfloor - t < h$ следует, что эти унарные операции сохраняют порядок между элементами t_k^I и t_l^{II}, более того порядок не нарушается, если операция применяется лишь к одному из элементов.

Введем теперь две вспомогательные игры, с той же динамикой, с той же целевой функцией и теми же компактами $[\phi_k^-;\phi_k^+]$, $[\psi_k^-;\psi_k^+]$, но отличающиеся от исходной возможными моментами выхода. В

$\lceil\rfloor$-игре эти последовательности будут образованы элементами $\lceil t_k^I \rceil$ для первого и элементами $\lfloor t_k^{II} \rfloor$ для второго игрока. В $\lfloor\rceil$-игре будет наоборот, $\lfloor t_k^I \rfloor$ – для первого игрока, и $\lceil t_k^I \rceil$ – для второго. Обозначим соответствующие классы допустимых для игроков моментов выхода через $\lceil \mathcal{Q}^I \rceil$, $\lfloor \mathcal{Q}^{II} \rfloor$ и $\lfloor \mathcal{Q}^I \rfloor$, $\lceil \mathcal{Q}^{II} \rceil$ для $\lceil\rfloor$-игры и $\lfloor\rceil$-игры. Отметим, что тогда унарные операции $\lceil\cdot\rceil$ и $\lfloor\cdot\rfloor$ переводят $\lfloor \mathcal{Q}^I \rfloor$ и $\lceil \mathcal{Q}^I \rceil$, а также $\lfloor \mathcal{Q}^{II} \rfloor$ и $\lceil \mathcal{Q}^{II} \rceil$ в друг в друга.

Рассмотрим теперь произвольный допустимый в исходной игре момент выхода θ_i одного из игроков, в частности для всякого t события $\theta_i \le t$ $\mathcal{F}_t$-измеримы. Тогда у момента выхода $\lceil \theta_i \rceil$ для всякого t события $\lceil \theta_i \rceil \le t$ также $\mathcal{F}_t$-измеримы. Следовательно как обладающие подходящими множествами значений момент выхода $\lceil \theta_1 \rceil$ будет допустим в $\lceil\rfloor$-игре, а момент $\lceil \theta_2 \rceil$ – в $\lfloor\rceil$-игре. Те же рассуждения дают тот же вывод, если в качестве θ_1 берется момент выхода, допустимый в $\lfloor\rceil$-игре, а в качестве θ_2 – допустимый в $\lceil\rfloor$-игре момент выхода.

Снова, в силу общей динамики, общими для всех рассматриваемых игр являются множества допустимых стратегий, и каждая такая допустимая стратегия создает на траекториях y одно и то же для всех игр распределение. По построению, вероятности событий $|\theta_i - \lceil \theta_1 \rceil| > h$, $|\theta_i - \lfloor \theta_1 \rfloor| > h$ равны нулю. В силу леммы при $r' = h$ и $r'' = 0$, для нижних цен $\lceil\rfloor$-игры и $\lfloor\rceil$-игры выполнено неравенство $|\mathcal{V}_-^{\lceil\rfloor} - \mathcal{V}_-^{\lfloor\rceil}| \le Lh$. Теперь, воспользовавшись показанными вложениями для моментов выхода в исходной игре, имеем

$$\mathcal{V}_- = \sup_{V \in \mathfrak{D}^{II}, \theta_2 \in \mathfrak{Q}^{II}} \inf_{U \in \mathfrak{D}^I, \theta_1 \in \mathfrak{Q}^I} \mathbb{E}_{\Omega,U,V} J(y, \theta_1, \theta_2) \le$$
$$\le Lh + \sup_{V \in \mathfrak{D}^{II}, \theta_2 \in \mathfrak{Q}^{II}} \inf_{U \in \mathfrak{D}^I, \theta_1 \in \mathfrak{Q}^I} \mathbb{E}_{\Omega,U,V} J(y, \lceil \theta_1 \rceil, \lfloor \theta_2 \rfloor) \le$$
$$\le Lh + \sup_{V \in \mathfrak{D}^{II}, \theta_2' \in \lfloor \mathfrak{Q}^{II} \rfloor} \inf_{U \in \mathfrak{D}^I, \theta_1 \in \lceil \mathfrak{Q}^I \rceil} \mathbb{E}_{\Omega,U,V} J(y, \theta_1', \theta_2') = Lh + \mathcal{V}_-^{\lceil\rfloor}.$$

Аналогично показывается неравенство $\mathcal{V}_- \ge Lh - \mathcal{V}_-^{\lfloor\rceil}$. Итак, доказано

$$|\mathcal{V}_-^{\lfloor\rceil} - \mathcal{V}_-| \le 2Lh.$$

Поскольку второе предположение выполнено для построенной нами $\lceil\rfloor$-игры, то построенные для нее в разделе 4.2 стратегия U и момент выхода $\theta_1^{\lceil\rfloor}$ удовлетворяет (1.6). Кроме того, для нее выполнено

(5.1), откуда для нижней цены $\mathcal{V}_-$ исходной игры следует

$$\mathcal{V}_- \geq \mathcal{V}_+ - 2\varepsilon/3 - 2h(L+2) \overset{(2.1)}{\geq} \mathcal{V}_+ - \varepsilon.$$

Осталось отметить, что стратегия U не меняется при переходе к исходной игре, а момент выхода θ_1, построенный в разделе 4.2 для исходной игры, связан с $\theta_1^{\lceil\rfloor}$ соотношением $\theta_1 = \lceil\theta_1^{\lceil\rfloor}\rceil$. Теперь $2\varepsilon/3+4h$-оптимальность пары $(U,\theta_1^{\lceil\rfloor})$ для верхней $\lceil\rfloor$-игры гарантирует $2\varepsilon/3+2h(L+4)$-оптимальность (U,θ_1) для исходной игры. В силу (2.1) показана и ее ε-оптимальность, то есть (1.6) также показано. Итак, построенные в разделе 4.2 для этой игры стратегия U и момент выхода θ_1 удовлетворяют (1.6) уже без каких-либо отличных от свойства (1.4) предположений на исходную игру.

Теорема доказана.

СПИСОК ЛИТЕРАТУРЫ

1. Биллингсли П. *Сходимость вероятностных мер.* М.: Наука, 1977.

2. Дынкин Е.Б. *Игровой вариант задачи об оптимальной остановке* // Доклады АН СССР. 1969. Т. 185, № 1. С. 16–19.

3. Красовский Н.Н. *Игра сближения-уклонения со стохастическим поводырем* // Доклады АН СССР. 1977. Т. 237, № 5. С. 1020–1023.

4. Красовский Н.Н., Субботин А.И. *Позиционные дифференциальные игры.* М.: Наука, 1974.

5. Красовский Н.Н., Котельникова А.Н. *Дифференциальная игра сближения–уклонения. Стохастический поводырь* // Труды ИММ УрО РАН. 2009. Т. 15, № 4. С. 146–166

6. Петросян Л.А., Шевкопляс Е.В. *Кооперативные дифференциальные игры со случайной продолжительностью* // Вестник Санкт-Петербургского Университета. Серия 1: Математика, Механика, Астрономия. 2000. № 4. С. 14—18.

7. Серегина Т.В., Ивашко А.А., Мазалов В.В. *Стратегии оптимальной остановки в игре “The Price Is Right”* // Труды ИММ УрО РАН. 2019. Т. 25, № 3. С. 217–231

8. Хлопин Д.В. *Дифференциальная игра с возможностью досрочного завершения* // Труды ИММ УрО РАН. 2021. Т. 27, № 4.

9. Amir R., Evstigneev I.V., Schenk-Hoppé K.R. *Asset market games of survival: a synthesis of evolutionary and dynamic games* // Annals of Finance. 2013. V. 9, N. 2. P. 121–144.

10. Averboukh Y. *Approximate solutions of continuous-time stochastic games* // SIAM J. Control Optim. 2016. V. 54, N. 5. P. 2629–2649.

11. Averboukh Y. *Approximate Public-Signal Correlated Equilibria For Nonzero-Sum Differential Games* // SIAM J. Control Optim. 2019. V. 57, N. 1. P. 743–772.

12. Basu A., Stettner L. *Zero-sum Markov games with impulse controls* // SIAM J. Control Optim. 2020. V. 58, N. 1. P. 580–604.

13. Bensoussan A., Friedman A. *Nonlinear variational inequalities and differential games with stopping times* // Journal of Functional Analysis. 1974. Vol. 16, N. 1. P. 305–352.

14. Bensoussan A., Friedman A. *Nonzero-sum stochastic differential games with stopping times and free boundary problems* // Transactions of the American Mathematical Society. 1977. V. 231, N. 2. P. 275–327.

15. Bielecki T. R., Crépey, S., Jeanblanc M., Rutkowski M. *Arbitrage pricing of defaultable game options with applications to convertible bonds* // Quantitative Finance. 2008. V. 8, N. 8. P. 795–810.

16. Guo X., Hernández-Lerma O. *Zero-sum continuous-time Markov games with unbounded transition and discounted payoff rates* // Bernoulli. 2005. V. 11, N. 6. P. 1009–1029.

17. Hamadéne S. *Mixed zero-sum stochastic differential game and American game options* // SIAM J. Control Optim. 2006. V. 45, N.2. P. 496–518.

18. Kolokoltsov V.N. *Markov processes, semigroups and generators.* De Gruyter Studies in Mathematics 38. Berlin: De Gryuter, 2011.

19. Laraki R., Solan E. *The value of zero-sum stopping games in continuous time* // SIAM J. Control Optim. 2005. V. 43, N. 5. P. 1913–1922.

20. Marin-Solano J., Shevkoplyas E. *Non-constant discounting and differential games with random time horizon* // Automatica. 2011. Vol. 47, N. 12. P. 2626–2638.

21. Neyman A. *Continuous-time stochastic games* // Games and Economic Behavior. 2017. V. 104. P. 92–130.

22. Prieto-Rumeau T, Hernández-Lerma O. *Selected Topics on Continuous-Time Controlled Markov Chains and Markov Games.* Vol. 5. ICP Advanced Texts in Mathematics. London: Imperial College Press, 2012.

DIFFERENTIAL GAME WITH DISCRETE STOPPING TIME

Dmitry V. Khlopin, Krasovskii Institute of Mathematics and Mechanics of the Ural Branch of the Russian Academy of Sciences, Cand.Sc. (khlopin@imm.uran.ru).

Abstract: We consider a two agents zero-sum differential game with discrete stopping times. The payoff function may depend both on the stopping time, on the state of the system at this time, and on the player who constitutes this stopping time. To formalize strategies, non-anticipating càdlàg stochastic processes are used. Under the Isaacs condition, the existence of the game value is established. The corresponding near-optimal strategy is constructed with stochastic guide based on an auxiliary model game governed by continuous-time Markov chain.

Keywords: zero-sum two-person games, Dynkin's stopping game, differential games, strategy with guide, near optimal strategies, extremal shift.