ТРУДЫ ИНСТИТУТА МАТЕМАТИКИ И МЕХАНИКИ УрО РАН

Том 27 № 4 2021

УДК 517.977.8

ДИФФЕРЕНЦИАЛЬНАЯ ИГРА С ВОЗМОЖНОСТЬЮ ДОСРОЧНОГО ЗАВЕРШЕНИЯ[1]

Д. В. Хлопин

Рассматривается антагонистическая дифференциальная игра на конечном промежутке, в которой игроки помимо управления траекторией системы влияют также на момент завершения игры. Предполагается, что момент досрочного завершения игры является абсолютно непрерывной случайной величиной, плотность которой задается назначаемой каждым игроком ограниченной измеримой функцией времени (интенсивностью его влияния на завершение игры). Платежная функция при этом может зависеть как от момента завершения игры и положения системы в этот момент, так и от игрока-инициатора завершения игры. Для формализации стратегий применяются неупреждающие случайные процессы с непрерывными справа и имеющими предел слева траекториями. В предположении условия седловой точки в маленькой игре показано существование цены игры. С этой целью исходная игра приближается вспомогательной игрой на основе марковской цепи с непрерывным временем, зависящей от управлений и интенсивностей игроков; и на основе оптимальных в марковской игре стратегий для исходной игры предложена процедура управления со стохастическим поводырем. Показано, что при неограниченном увеличении числа точек в марковской игре такая процедура приводит к сколь угодно близкой к оптимальной стратегии в исходной игре.

Ключевые слова: антагонистические игры, игра Дынкина, дифференциальные игры, стохастический поводырь, экстремальный сдвиг, марковская цепь с непрерывным временем.

D. V. Khlopin. Differential game with the possibility of early termination.

We consider a zero-sum differential game on a finite interval, in which the players not only control the system's trajectory but also influence the terminal time of the game. It is assumed that the early terminal time is an absolutely continuous random variable, and its density is given by bounded measurable functions of time assigned by both players (the intensities of the influence of each player on the termination of the game). The payoff function may depend both on the terminal time of the game together with the position of the system at this time and on the player who initiates the termination. The strategies are formalized by using nonanticipating càdlàg processes. The existence of the game value is shown under the Isaacs condition. For this, the original game is approximated by an auxiliary game based on a Markov chain with continuous time, which depends on the controls and intensities of the players. Based on the strategies optimal in this Markov game, a control procedure with a stochastic guide is proposed for the original game. It is shown that, under an unlimited increase in the number of points in the Markov game, this procedure leads to a near-optimal strategy in the original game.

Keywords: two-person zero-sum game, Dynkin game, differential game, stochastic guide, extremal shift, Markov chain with continuous time.

MSC: 49N70, 60G40, 91A60, 91A23

DOI: 10.21538/0134-4889-2021-27-4-189-214

Введение

Данная работа посвящена развитию методов решения дифференциальных игр со случайным моментом окончания. Рассматривается антагонистическая дифференциальная игра на конечном промежутке $[0;T]$, в которой игроки могут неупреждающим образом влиять на плотность момента досрочного (до момента времени T) завершения игры, задавая в каждый момент времени каждый со своей стороны интенсивность (условную плотность) ее завершения. Функционал качества в рассматриваемой игре зависит как от момента завершения игры и положения системы в этот момент, так и от игрока-инициатора досрочного завершения игры, таким образом этот функционал становится случайной величиной. По этой причине действия

[1]Работа выполнена при поддержке РНФ (проект 17-11-01093).

игроков считаются направленными на оптимизацию математического ожидания этого функционала. Каждый игрок по-прежнему имеет право, назначая по ходу игры свои управление и интенсивность, опираться на имеющееся у него знание о положении системы, но при этом следует также учитывать, что игра между игроками может происходить неоднократно, действия игроков могут меняться в ходе этой серии игр, в частности, они могут оказаться недетерминированными, являясь результатом тех или иных случайных процессов. Вследствие этого в работе предполагается, что, с одной стороны, игроки имеют право задавать не управления, а распределения своих управлений. С другой стороны, собирая статистику по достаточно долгой серии игр, игроки имеют возможность восстановить используемое соперником распределение с любой наперед заданной точностью; в частности, ничего не зная в ходе игры об уже назначенной противником интенсивности досрочного завершения, каждый из них может восстановить управляющий ею вероятностный закон.

В связи с вышесказанным, в данной работе стратегии полагаются случайными процессами, неупреждающим образом зависящими от реализующейся траектории системы, а в случае контрстратегии — и от действий соперника. Подробно соответствующее определение стратегий и контрстратегий дано в разд. 2; отметим здесь лишь, что в целом они не новы, для марковской цепи с непрерывным временем столь общие определения имеются, например в [26], там же подробно прописаны более конструктивные частные случаи такого определения. Например, в нашей постановке каждый игрок имеет право использовать классические детерминированные стратегии: кусочно-непрерывные программные стратегии или кусочно-постоянные позиционные стратегии. Отметим также, что накладываемое в этой статье требование к стратегиям применять лишь càdlàg процессы носит скорее технический характер; в некоторой степени оно может быть оправдано тем, что в теории управления кусочно-непрерывные управления всюду плотны, а в теории дифференциальных игр в рамках классической постановки [20;30;21] кусочно-постоянные позиционные управления игроков реализуют функцию цены с любой наперед заданной точностью.

Обсудим имеющиеся в литературе схожие постановки. Общая постановка игры, в которой игрок управляет моментом выхода из процесса, восходит к классической работе Дынкина [10]. В таких играх (*Dynkin games, stopping games*) обычно предполагается, что имеется некоторый конфликтно-управляемый процесс и хотя бы один из игроков своими действиями должен выбрать удобный для себя момент остановки всей игры. Такие игровые постановки имеют многочисленные приложения, в том числе в финансовой математике, (см., например, [1;7;15]); для них исследуются вопросы существования и описания цены игры, построения оптимальных стратегий, структуры таких стратегий. Начиная с классических работ [5;6], динамика игры при этом обычно описывается при помощи стохастического дифференциального уравнения, но используются также и процессы Леви — Феллера [4], марковские цепи с дискретным (см., например, [24;29]) или непрерывным временем [11;22]. Наконец, дифференциальная игра, в которой распределение момента выхода (а следовательно, и интенсивности) известно игрокам заранее, рассматривалась, например, в работах [13;23].

В данной работе для рассматриваемой игры доказывается существование как цены, так и приближенно оптимальных стратегий, реализующих ее с любой наперед заданной точностью. Их построение опирается на конструкцию стохастического поводыря, идея которой заключается в следующем. По всякому движению y исходной конфликтно-управляемой системы игроком пошагово восстанавливается однозначно некоторое распределение $\mathbb{P}^Z$ рассчитываемых им вспомогательных движений z. При этом в каждый момент t_k некоторого разбиения на следующий промежуток игрок назначает свое управление $\bar{u}$ при помощи реализовавшейся позиции, а фактически случайной величины, $z(t_k)$ правилом

$$\bar{u}(t) \equiv u^{to\,z(t_k)}(t_k, y(t_k)) \quad \forall t \in [t_k; t_{k+1}), \tag{0.1}$$

где $u^{to\,z}$ — управление, максимально сдвигающее позицию исходной системы к точке z. Такая решающая исходную задачу конструкция идейно восходит к классическим методам решения дифференциальных игр [20], предложенным Н. Н. Красовским и А. И. Субботиным. Например, правило $\bar{u}^{to\,z}$ является ключевым в методе экстремального сдвига, а правило (0.1) при подстановке $y \equiv z$ используется в позиционной формализации дифференциальной игры, введенной Н. Н. Красовским и А. И. Субботиным в [20]. Предложенный теми же авторами метод управления с поводырем изначально предполагал однозначную зависимость z от y, но уже в работе Н. Н. Красовского [17] было предложено недетерминированное правило построения z. В [18;19] для построения стохастического поводыря применялось дифференциальное уравнение Ито. В работах [2;3] в качестве z рассматривается траектория марковской игры с непрерывным временем. В нашей статье z также порождено приближенно оптимальной стратегией в марковской игре с непрерывным временем. В частности, вопрос существования цены в исходной игре сводится к аналогичному факту для марковской игры, показанному в [14].

Сама статья организована следующим образом. В разд. 1 формулируется исходная дифференциальная игра и с точностью до определения стратегий основной результат: теорема 1, гарантирующая существование цены в этой игре. В разд. 2 даются некоторые базовые определения теории случайных процессов, после чего вводятся используемые в статье формализации стратегий. В следующем разделе доопределяется динамика исходной игры, а имеющийся целевой функционал представляется в виде интеграла на бесконечном промежутке. Раздел 4 посвящен вспомогательной марковской игре. В разд. 5 рассматривается конструкция стохастического поводыря для первого игрока: вводится случайный процесс (двойная игра), первая компонента которого отвечает состоянию исходной игры, а вторая — состоянию стохастического поводыря, затем при помощи экстремального сдвига прописываются управление первого игрока в каждой компоненте и фиктивное управление за второго в марковской игре. Последний раздел посвящен доказательству теоремы, сведенному к оценкам разности между компонентами случайного процесса — траекториями исходной игры и стохастического поводыря.

1. Постановка задачи

1.1. Динамика исходной игры

Пусть задана функционирующая в $\mathbb{R}^d$ конфликтно-управляемая система

$$\frac{dy(t)}{dt} = f(t, y(t), u(t), v(t)), \quad y(0) = x_* \in \mathbb{R}^d, \quad t \in [0;T], \quad u(t) \in \mathbb{U}, \quad v(t) \in \mathbb{V}. \tag{1.1}$$

Предположим, что $\mathbb{U}$ и $\mathbb{V}$ — метрические компакты в некоторых конечномерных евклидовых пространствах, функция $f : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{U} \times \mathbb{V} \to \mathbb{R}^d$ липшицева по всем переменным. Также потребуем условие седловой точки в маленькой игре [20;21]: для всех $x, w \in \mathbb{R}^d$ выполнено

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

Помимо управления движением в $\mathbb{R}^d$ каждый игрок имеет возможность влиять на момент завершения игры, зафиксируем для этого отрезки $[\varphi_-;\varphi_+], [\psi_-;\psi_+] \subset \mathbb{R}_+$ — множества параметров, управляющих интенсивностью завершения игры со стороны каждого игрока. Пусть если первый игрок задал некоторую кусочно-непрерывную функцию $[0;T] \ni t \mapsto \varphi(t) \in [\varphi_-;\varphi_+]$, а вторым выбрана кусочно-непрерывная функция $[0;T] \ni t \mapsto \psi(t) \in [\psi_-;\psi_+]$, то для почти всех $t \in [0;T)$ условная вероятность того, что игра завершится на промежутке $[t; t+\Delta t)$, если в момент t игра еще продолжалась, была равна $(\varphi(t)+\psi(t))\Delta t + o(\Delta t)$. Решая соответствующее дифференциальное уравнение уже для суммарной интенсивности $\varphi + \psi$, получаем, что для любого момента времени $\tau \in [0;T)$ вероятность завершения игры до момента времени τ равна

$$\int_0^\tau (\varphi(t)+\psi(t)) e^{-\int_0^t (\varphi(s)+\psi(s))\,ds}\, dt, \tag{1.3}$$

в частности, при $\varphi(t)+\psi(t)\equiv C$ момент завершения игры имеет то же распределение, что и $\min(\theta,T)$, где θ распределено экспоненциально с показателем C. Воспользовавшись (1.3), на основе соответствующих дифференциальных уравнений уже для вероятностей завершения игры до момента $\tau<T$ по инициативе первого и второго игрока соответственно имеем для них

$$\int\limits_0^\tau \varphi(t)e^{-\int\limits_0^t(\varphi(s)+\psi(s))\,ds}dt,\quad \int\limits_0^\tau \psi(t)e^{-\int\limits_0^t(\varphi(s)+\psi(s))\,ds}dt. \tag{1.4}$$

При этом в качестве следствия получаем, что в случае завершения игры в момент времени $t<T$ ответственным за это будет считаться с вероятностью $\varphi(t)/(\varphi(t)+\psi(t))$ первый игрок, а с вероятностью $\psi(t)/(\varphi(t)+\psi(t))$ — второй. Осталось формализовать формулы (1.3) и (1.4).

Для этого оснастим квадрат $[0;1]^2$ σ-алгеброй $\mathcal{B}([0;1]^2)$ борелевских множеств и рассмотрим на них лебеговскую меру λ. Мы получили вероятностное пространство

$$\big([0;1]^2,\mathcal{B}([0;1]^2),\mathbb{P}_{[0;1]^2}\stackrel{\triangle}{=}\lambda\big),$$

которое задает по случайному числу ω_i для каждого игрока. Теперь для всех борелевских отображений $\varphi:[0;T)\to[\varphi_-;\varphi_+]$, $\psi:[0;T)\to[\psi_-;\psi_+]$ введем случайные величины $\omega\mapsto\theta_1(\omega)$, $\omega\mapsto\theta_2(\omega)$ правилом: для всех $\omega=(\omega_1,\omega_2)\in[0;1]^2$

$$\theta_1(\omega)\stackrel{\triangle}{=}\inf\left\{t\in[0;T)\;\Big|\;\int\limits_0^t\varphi(t)\,dt=-\ln\omega_1\right\}\cup\{T\}, \tag{1.5}$$

$$\theta_2(\omega)\stackrel{\triangle}{=}\inf\left\{t\in[0;T)\;\Big|\;\int\limits_0^t\psi(t)\,dt=-\ln\omega_2\right\}\cup\{T\}. \tag{1.6}$$

Отметим, что при любом $\tau\in[0;T)$ вероятность события $\theta_1<\tau$ совпадает с (1.3) в случае, если $\psi\equiv 0$. Аналогично вероятность события $\theta_2<\tau$ совпадает с (1.3) в случае, если $\varphi\equiv 0$. Таким образом, случайная величина θ_i совпадает с моментом возможного завершения игры по инициативе i-го игрока при нулевой интенсивности его соперника. Более того, поскольку при любых отображениях $\varphi:[0;T)\to[\varphi_-;\varphi_+]$, $\psi:[0;T)\to[\psi_-;\psi_+]$ случайные величины θ_i вслед за ω_i независимы, то для момента досрочного окончания игры, случайной величины

$$\theta_{\min}\stackrel{\triangle}{=}\min(\theta_1,\theta_2),$$

вероятность события $\theta_{\min}<\tau$ находится по формуле (1.3), а событий $\theta_1<\tau$ и $\theta_2<\tau$, — по формуле (1.4), что и требовалось. Отметим также, что всюду далее мы будем считать, что (ω_1,ω_2) выбирается до начала игры, никоим образом не зависит от действий игроков и никакой информацией о (ω_1,ω_2) игроки до момента окончания игры не обладают. Тем самым, зная формулы (1.3), (1.4), зная (или предполагая) свою интенсивность и/или интенсивность соперника, игроки не имеют более никакой информации о моменте $\theta_{\min}$ завершения игры.

Для простоты обозначений введем для первого и второго игроков множества управляющих параметров $\bar{\mathbb{U}}\stackrel{\triangle}{=}\mathbb{U}\times[\varphi_-;\varphi_+]$, $\bar{\mathbb{V}}\stackrel{\triangle}{=}\mathbb{V}\times[\psi_-;\psi_+]$.

1.2. Цели игроков и цена игры

Теперь определим цели игроков. Пусть заданы функции $\sigma_0:\mathbb{R}^d\to\mathbb{R}$, $\sigma_1:[0;T]\times\mathbb{R}^d\to\mathbb{R}$, $\sigma_2:[0;T]\times\mathbb{R}^d\to\mathbb{R}$. Будем считать, что они локально липшицевы по x,t и ограничены по норме числом 1. Пусть $\sigma_0(y(T))$ будет платой первого игрока второму, если досрочного завершения игры не произошло и игра завершилась в момент T. Число $\sigma_1(\theta_1,y(\theta_1))$ пусть станет платой первого второму, если первый инициировал досрочное завершение игры (в момент времени θ_1). Аналогично, число $\sigma_2(\theta_2, y(\theta_2))$ — плата первого второму, если завершение игры (в момент θ_2) инициировано вторым. Тем самым, почти всегда на $[0;1]^2$ при известных траектории $y(\cdot)$ и действиях $(\bar{u},\bar{v})(\cdot)$ игроков определена целевая функция, случайная величина

$$J(y,\bar{u},\bar{v}) \stackrel{\triangle}{=} \begin{cases} \sigma_0(y(T)), & T=\theta_{\min}, \\ \sigma_1(\theta_1, y(\theta_1)), & \theta_1=\theta_{\min}, \\ \sigma_2(\theta_2, y(\theta_2)), & \theta_2=\theta_{\min}. \end{cases}$$

Поскольку исход $\omega\in[0;1]^2$ игрокам неизвестен, они могут лишь оптимизировать математическое ожидание целевой функции J, число

$$\mathbb{E}J(y,\bar{u},\bar{v}),$$

которое, как и траектория $y(\cdot)$ системы (1.1), однозначно восстанавливается по управлениям игроков $\bar{u}(\cdot),\bar{v}(\cdot)$.

Формализации стратегий будут даны в следующем разделе, отметим лишь, что мы разрешим одному из игроков подавать в качестве своего управления càdlàg (непрерывную справа и имеющую предел слева) траекторию некоторого своего случайного процесса, неупреждающим образом зависящего от траектории (1.1). Другому позволим знать не только траекторию, но и подаваемое соперником управление; как следствие его управление как случайный процесс неупреждающе зависит именно от них. Приведем также нужные обозначения.

Пусть $\mathfrak{S}^I,\mathfrak{S}^{II}$ обозначают множества допустимых стратегий игроков как распределений случайных процессов, неупреждающе зависящих от траектории $y(\cdot)$; множествам допустимых контрстратегий игроков сопоставим обозначения $\mathfrak{Q}^I_{x_*}$, $\mathfrak{Q}^{II}_{x_*}$. Через $\mathfrak{R}_{x_*}(\mathbb{P}^{I+II})$ обозначим множество всевозможных распределений $\mathbb{P}^{all}_{x_*}$ случайных процессов игры, соблюдающих совместную стратегию $\mathbb{P}^{I+II}$ игроков. При этом, естественно, всякие пошаговые стратегии (тем более программные стратегии) будут допустимыми стратегиями, а всякая допустимая стратегия — контрстратегией.

В зависимости от того, какой именно из игроков информационно дискриминирован, т. е. заранее сообщил сопернику о применяемой им стратегии, мы получаем два значения игры

$$V^+ \stackrel{\triangle}{=} \inf_{\mathbb{P}^I\in\mathfrak{S}^I_{x_*}} \sup_{\mathbb{Q}^{II}\in\mathfrak{Q}^{II}_{x_*}} \sup_{\mathbb{P}^{all}_{x_*}\in\mathfrak{R}_{x_*}(\mathbb{Q}^{II}[\mathbb{P}^I])} \mathbb{E}^{all}_{x_*}J(y,\bar{u},\bar{v}),$$

$$V^+ \stackrel{\triangle}{=} \sup_{\mathbb{P}^{II}\in\mathfrak{S}^{II}_{x_*}} \inf_{\mathbb{Q}^{I}\in\mathfrak{Q}^{I}_{x_*}} \inf_{\mathbb{P}^{all}_{x_*}\in\mathfrak{R}_{x_*}(\mathbb{Q}^{I}[\mathbb{P}^{II}])} \mathbb{E}^{all}_{x_*}J(y,\bar{u},\bar{v}).$$

Оказывается, имеет место

Теорема 1. *Выполнено равенство $V^-=V^+$. Более того, для каждого положительного ε у игроков найдутся допустимые стратегии $\mathbb{P}^I_\varepsilon\in\mathfrak{S}^I_{x_*}$, $\mathbb{P}^{II}_\varepsilon\in\mathfrak{S}^{II}_{x_*}$, гарантирующие*

$$-\varepsilon+\sup_{\mathbb{Q}^{II}\in\mathfrak{Q}^{II}_{x_*}} \sup_{\mathbb{P}^{all}_{x_*}\in\mathfrak{R}_{x_*}(\mathbb{Q}^{II}[\mathbb{P}^I_\varepsilon])} \mathbb{E}^{all}_{x_*}J(y,\bar{u},\bar{v}) \le V^+=V^- \le \varepsilon+\inf_{\mathbb{Q}^{I}\in\mathfrak{Q}^{I}_{x_*}} \inf_{\mathbb{P}^{all}_{x_*}\in\mathfrak{R}_{x_*}(\mathbb{Q}^{I}[\mathbb{P}^{II}_\varepsilon])} \mathbb{E}^{all}_{x_*}J(y,\bar{u},\bar{v}).$$

Саму теорему мы докажем в последнем разделе. Перед этим сначала мы напомним некоторые определения из теории вероятностных процессов и дадим, наконец, обещанные ранее определения стратегий.

2. Стратегии как распределения случайных процессов

2.1. Общие сведения из теории случайных процессов

Напомним, что для любой пары вероятностных пространств $(\Omega',\mathcal{F}',\mathbb{P}')$, $(\Omega'',\mathcal{F}'',\mathbb{P}'')$ можно рассмотреть σ-алгебру $\mathcal{F}'\otimes\mathcal{F}''$ подмножеств $\Omega'\times\Omega''$ как минимальную σ-алгебру, содержащую все произведения вида $A\times B$, где $A\in\mathcal{F}', B\in\mathcal{F}''$, и однозначно восстановить на ней вероятность $\mathbb{P}' \otimes \mathbb{P}''$ правилом $(\mathbb{P}' \otimes \mathbb{P}'')(A \times B) = \mathbb{P}'(A)\mathbb{P}''(B)$ для всех $A \in \mathcal{F}', B \in \mathcal{F}''$. Так получается произведение исходных вероятностных пространств — $(\Omega' \times \Omega'', \mathcal{F}' \otimes \mathcal{F}'', \mathbb{P}' \otimes \mathbb{P}'')$. Кроме того, если имеется отображение g из вероятностного пространства $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ в измеримое пространство $(\Omega_2, \mathcal{F}_2)$ и это отображение измеримо, т. е. $g^{-1}(A) \in \mathcal{F}_1$ для всех $A \in \mathcal{F}_2$, то это отображение задает *образ меры*, вероятность $\mathbb{P}_2$ для $(\Omega_2, \mathcal{F}_2)$, правилом $\mathbb{P}_2(A) \triangleq \mathbb{P}_1(g^{-1}(A))$ для всех $A \in \mathcal{F}_2$. В частности, у введенной выше для $\Omega' \times \Omega''$ вероятности маргинальные вероятности $\mathbb{P}'$ и $\mathbb{P}''$ для Ω' и Ω'' совпадают с образами меры $\mathbb{P}' \otimes \mathbb{P}''$ при отображениях $(\omega', \omega'') \mapsto \omega'$ и $(\omega', \omega'') \mapsto \omega''$ соответственно.

Зафиксируем пространство состояний $\mathbb{X}$, некоторое замкнутое подмножество конечномерного евклидового пространства. Снабдим его σ-алгеброй борелевских множеств $\mathcal{B}(\mathbb{X})$, получая измеримое пространство $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$. Зафиксируем также временной промежуток $\mathcal{T}$.

Рассмотрим теперь пространство Скорохода $D\big(\mathcal{T}, \mathbb{X}\big)$ [8;16] всевозможных càdlàg функций из $\mathcal{T}$ в $\mathbb{X}$ (непрерывных справа и имеющих предел слева функций из $\mathcal{T}$ в $\mathbb{X}$). Это пространство польское; оснащая его соответствующей ему борелевской σ-алгеброй $\mathcal{B}\big(D(\mathcal{T}, \mathbb{X})\big)$, мы получим измеримое пространство. На каждом пространстве Скорохода по умолчанию зафиксируем *каноническую фильтрацию*

$$\Big(\mathcal{A}^{\mathbb{X}}_t \triangleq \mathcal{B}\big(D((-\infty; t] \cap \mathcal{T}, \mathbb{X})\big) \otimes \{\varnothing, D\big((t; +\infty) \cap \mathcal{T}, \mathbb{X})\}\Big)_{t \in \mathcal{T}},$$

получая *стохастический базис* $\big(D\big(\mathcal{T}, \mathbb{X}\big), \mathcal{B}\big(D(\mathcal{T}, \mathbb{X})\big), (\mathcal{A}^{\mathbb{X}}_t)_{t \in \mathcal{T}}\big)$.

Следуя [25, IV.1.O2], *случайным процессом с временным промежутком* $\mathcal{T}$ *и пространством состояний* $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ назовем систему из некоторого вероятностного пространства $(\Omega, \mathcal{F}, \mathbb{P})$ и некоторого семейства $(X_t)_{t \in \mathcal{T}}$ измеримых отображений из Ω в $\mathbb{X}$. При этом каждое X_t назовем *состоянием процесса в момент* t, а отображение $t \mapsto X_t(\omega)$ — *траекторией процесса*, соответствующей данному исходу $\omega \in \Omega$. Случайный процесс назовем *càdlàg процессом*, если $\mathbb{P}$-почти все его траектории таковы. Каждому процессу можно сопоставить *первый канонический процесс* [25], заменяя каждый ω соответствующей траекторией $t \mapsto X_t(\omega)$; в частности, все càdlàg процессы с пространством состояний $\mathbb{X}$ и временным промежутком $\mathcal{T}$ однозначно восстанавливаются некоторой вероятностной мерой (*распределением процесса* [9, § 17.1]), заданной на измеримом пространстве $(D(\mathcal{T}, \mathbb{X}), \mathcal{B}(D(\mathcal{T}, \mathbb{X})))$. Далее большинство càdlàg процессов мы будем задавать в виде первых канонических процессов, фиксируя лишь их распределения в подходящем пространстве Скорохода.

Будем говорить, что $(X_t)_{t \in \mathcal{T}}$ *согласован с* $(\mathcal{F}_t)_{t \in \mathcal{T}}$, если для каждого $t \in \mathcal{T}$ $\mathcal{F}_t$-измерима случайная величина $\omega \mapsto X_t(\omega)$. Случайную величину τ, принимающую значения в $\mathcal{T}$, называют *моментом остановки* (относительно фильтрации $(\mathcal{F}_t)_{t \in \mathcal{T}}$), если для всех $t \in \mathcal{T}$ событие $\{\tau \leq t\} = \{\omega \mid \tau(\omega) \leq t\}$ лежит в $\mathcal{F}_t$. Отметим, что в случае, если в качестве фильтрации $(\mathcal{F}_t)_{t \in \mathcal{T}}$ для $\Omega \triangleq D(\mathcal{T}, \mathbb{X})$ взята каноническая фильтрация $(\mathcal{A}^{\mathbb{X}}_t)_{t \in \mathcal{T}}$, то всякий заданный на такой базисе процесс является $\mathcal{A}^{\mathbb{X}}_t$-измеримым, а всевозможные моменты остановки τ — в точности те отображения $\tau : D(\mathcal{T}, \mathbb{X}) \mapsto \mathcal{T}$, для которых при любых $x, x' \in D(\mathcal{T}, \mathbb{X})$ из $x(t) = x'(t)$ для всех $\tau(x) \geq t$ следует $\tau(x) = \tau(x')$.

2.2. Множества стратегий

В этом разделе $\mathcal{T}$ — замкнутый слева интервал, а t_0 — его нижняя грань.

Пусть помимо пространства Скорохода $D\big(\mathcal{T}, \mathbb{X}\big)$ даны некоторые компакты Υ_1, Υ_2 в некоторых конечномерных евклидовых пространствах — множества управляющих параметров игроков; примем также $\Upsilon \triangleq \Upsilon_1 \times \Upsilon_2$. *Программными стратегиями первого, второго игроков* и их *совместными программными стратегиями* назовем всевозможные кусочно-непрерывные справа и имеющие предел слева отображения — элементы $D(\mathcal{T}, \Upsilon_1)$, $D(\mathcal{T}, \Upsilon_2)$ и $D(\mathcal{T}, \Upsilon)$ соответственно. Как и раньше, снабдим их σ-алгебрами борелевских множеств с канонической фильтрацией.

Для фазового пространства $\mathbb{X}$ и множества управляющих параметров Υ назовем отображение

$$\mathcal{P}(\mathbb{X}) \times D\big(\mathcal{T}, \Upsilon\big) \ni (\varrho, \upsilon(\cdot)) \mapsto \mathbb{P}^G_{\upsilon(\cdot)}[\varrho] \in \mathcal{P}\big(D\big(\mathcal{T}, \mathbb{X}\big)\big)$$

допустимой динамикой $\mathbb{P}^G$ с начальным распределением $\varrho \in \mathcal{P}(\mathbb{X})$, если оно удовлетворяет начальному условию: $\mathbb{P}^G_{\upsilon(\cdot)}[\varrho](X(t_0) \in A) \equiv \varrho(A)$ для всех $A \in \mathcal{B}(\mathbb{X})$, $\upsilon(\cdot) \in D\big(\mathcal{T}, \Upsilon\big)$, а также является неупреждающим, т. е. для всех $\upsilon'(\cdot), \upsilon''(\cdot) \in D\big(\mathcal{T}, \Upsilon\big)$ и момента времени $t \in \mathcal{T}$ из $\upsilon'|_{[t_0;t)} = \upsilon''|_{[t_0;t)}$ следует, что выполнено $\mathbb{P}^G_{\upsilon'(\cdot)}[\varrho](A) = \mathbb{P}^G_{\upsilon''(\cdot)}[\varrho](A)$ для всех $A \in \mathcal{A}^{\mathbb{X}}_t$. Отметим, что всякая вероятность $\mathbb{P}^G_{\upsilon(\cdot)}[\varrho]$ регулярна, поскольку $D\big(\mathcal{T}, \mathbb{X}\big)$ — полное сепарабельное метрическое пространство.

Назовем *стратегией* $\mathbb{P}^i$ *i-го игрока* отображение, сопоставляющее каждому элементу из $x \in D\big(\mathcal{T}, \mathbb{X}\big)$ распределение $\mathbb{P}^i_{x(\cdot)}$ на $D(\mathcal{T}, \Upsilon_i)$ (а значит, и первый канонический процесс $(\upsilon^i_{x(\cdot)}(t))_{t\in\mathcal{T}}$ со значениями в Υ_i) такое, что для всех $t \in \mathcal{T}$ выполнено условие неупреждаемости: если $x'(s) = x''(s)$ для всех $s \in \mathcal{T}$, не больших t, то $\mathbb{P}^i_{x'(\cdot)}(A) = \mathbb{P}^i_{x''(\cdot)}(A)$ для всех $A \in \mathcal{A}^{\Upsilon_i}_t$. Аналогично вводится понятие *совместной стратегии* $\mathbb{P}^{I+II}$ игроков, отображающей траектории из $D\big(\mathcal{T}, \mathbb{X}\big)$ в распределения на $D(\mathcal{T}, \Upsilon)$.

Отметим, что любая программная стратегия $\upsilon_i(\cdot) \in D(\mathcal{T}, \Upsilon_i)$ как не зависящее от траектории $x(\cdot)$ отображение $x(\cdot) \mapsto \upsilon_i(\cdot)$ из $D\big(\mathcal{T}, \mathbb{X}\big)$ в $D(\mathcal{T}, \Upsilon_i)$ является стратегией i-го игрока.

Для всякого начального условия $\varrho \in \mathcal{P}(\mathbb{X})$ назовем случайный процесс $(X_t)_{t\in\mathcal{T}}$ со значениями в $\mathbb{X}$ *реализацией совместной стратегии* $\mathbb{P}^{I+II}$ *при допустимой динамике* $\mathbb{P}^G$, если имеется процесс $(\upsilon_t, X_t)_{t\in\mathcal{T}}$, заданный некоторым распределением $\mathbb{P}^{all}_\varrho$ на $D(\mathcal{T}, \Upsilon) \times D\big(\mathcal{T}, \mathbb{X}\big)$, для которого отображения $\upsilon(\cdot) \mapsto \mathbb{P}^G_{\upsilon(\cdot)}[\varrho]$ и $x(\cdot) \mapsto \mathbb{P}^{I+II}_{x(\cdot)}$ являются относительно $\mathbb{P}^{all}_\varrho$ производными Радона—Никодима маргинальных распределений $\mathbb{P}^\chi_\varrho$ и $\mathbb{P}^\Upsilon_\varrho$ (на $D\big(\mathcal{T}, \mathbb{X}\big)$ и $D(\mathcal{T}, \Upsilon)$ соответственно), т. е.

$$\mathbb{P}^\Upsilon_\varrho(A) \stackrel{\triangle}{=} \mathbb{P}^{all}_\varrho(\upsilon_. \in A) = \int\limits_{D(\mathcal{T},\mathbb{X})} \mathbb{P}^{I+II}_{x(\cdot)}(A)\,\mathbb{P}^\chi(dx) \quad \forall A \in \mathcal{B}(D(\mathcal{T}, \Upsilon)),$$

$$\mathbb{P}^\chi_\varrho(C) \stackrel{\triangle}{=} \mathbb{P}^{all}_\varrho(X_. \in C) = \int\limits_{D(\mathcal{T},\Upsilon)} \mathbb{P}^G_{\upsilon(\cdot)}[\varrho](C)\,\mathbb{P}^\Upsilon(d\upsilon) \quad \forall C \in \mathcal{B}(D(\mathcal{T}, \mathbb{X})).$$

Будем обозначать множество всевозможных реализаций для стратегии $\mathbb{P}^{I+II}$ через $\mathfrak{R}_\varrho(\mathbb{P}^{I+II})$.

Будем называть стратегию $\mathbb{P}^i$ i-го игрока *допустимой* для начального условия $\varrho \in \mathcal{P}(\mathbb{X})$, если у пары стратегий $(\mathbb{P}^i, \delta_{\upsilon_{3-i}})$ имеется соответствующая реализация для любой программной стратегии $\upsilon_{3-i} \in D(\mathcal{T}, \Upsilon_{3-i})$ его соперника. В частности, сами программные стратегии допустимы; действительно, при любом выборе программных стратегий (υ_1, υ_2) игроков в качестве реализации их совместной стратегии можно взять $\delta_{(\upsilon_1,\upsilon_2)} \otimes \mathbb{P}^G_{(\upsilon_1,\upsilon_2)(\cdot)}[\varrho]$.

Назовем *контрстратегией* $\mathbb{Q}^i[\mathbb{P}^{III-i}]$ *i-го игрока* отображение, сопоставляющее каждой допустимой стратегии его соперника $\mathbb{P}^{III-i}$ некоторую совместную стратегию $x(\cdot) \mapsto \mathbb{Q}^i[\mathbb{P}^{III-i}]_{x(\cdot)}$ на $D(\mathcal{T}, \Upsilon_1 \times \Upsilon_2)$ (а значит, и случайный процесс $(\upsilon_{x(\cdot)}(t))_{t\in\mathcal{T}}$ со значениями в $\Upsilon_1 \times \Upsilon_2$) такую, что соответствующее маргинальное распределение на Υ_{3-i} совпадает с $\mathbb{P}^{III-i}_{x(\cdot)}$. Назовем контрстратегию *допустимой*, если стратегия $x(\cdot) \mapsto \mathbb{Q}^i[\mathbb{P}^{III-i}]_{x(\cdot)}$ допустима при любой допустимой стратегии соперника $\mathbb{P}^{III-i}$. Понятно, что всякая допустимая стратегия $\mathbb{P}^i$, как сопоставляющая стратегии соперника $\mathbb{P}^{III-i}$ некоторую совместную стратегию $x(\cdot) \mapsto \mathbb{P}^I_{x(\cdot)} \otimes \mathbb{P}^{II}_{x(\cdot)}$, является также и допустимой контрстратегией.

Обозначим через $\mathfrak{S}^I_{x_*}, \mathfrak{S}^{II}_{x_*}$ множества допустимых (для начального условия $\varrho = \delta_{x_*}$) стратегий первого и второго игроков. Аналогично через $\mathfrak{Q}^I_{x_*}, \mathfrak{Q}^{II}_{x_*}$ обозначим множества допустимых (для того же условия) контрстратегий игроков. Из вышесказанного следует цепочка вложений $D(\mathcal{T}, \Upsilon_i) \hookrightarrow \mathfrak{S}^i_{x_*} \hookrightarrow \mathfrak{Q}^i_{x_*}$.

Построим пошаговую стратегию для первого, к примеру, игрока; заодно покажем, что всякая такая стратегия также является допустимой. Пусть задано некоторое разбиение $\Delta = (t_k)_{k\in\mathbb{N}}$ промежутка времени $\mathcal{T}$ и для каждого $k \in \mathbb{N}$, на каждом промежутке $[t_{k-1}; t_k)$, задана некоторая допустимая (например программная) стратегия $\mathbb{P}^{I:k}$ первого игрока. В частности, на этом промежутке для любой программной стратегии $\upsilon_k \in D([t_{k-1}; t_k), \Upsilon_2)$ второго игрока имеется реализация $\mathbb{P}^{all:k}$ их совместной стратегии $\delta_{\upsilon_k} \otimes \mathbb{P}^{I:k}$. Теперь каждому начальному (в момент t_{k-1}) распределению ϱ_{k-1} можно сопоставить вероятность $\mathbb{P}^{all:k}_{\varrho_{k-1}}(X(\cdot) \in B)$ для всех борелевских подмножеств B множества $D([t_{k-1}; t), \mathbb{X})$, а значит, и зависящую от ϱ маргинальную вероятность $\varrho_k[\varrho](B) \overset{\triangle}{=} \mathbb{P}^{\chi:k}_{\varrho_{k-1}}(X(t_k) \in B)$. Таким образом, построена зависящая от ϱ_0 последовательность переходных вероятностей $\varrho_k : \mathcal{P}(\mathbb{X}) \to \mathcal{P}(\mathbb{X})$, после чего корректно ввести распределение $\mathbb{P}^{all}_{\varrho_0}$ как вероятность $\otimes_{k\in\mathbb{N}}\mathbb{P}^{all:k}_{\varrho_{k-1}}$ на борелевских подмножествах $\prod_{k\in\mathbb{N}} D([t_{k-1}; t_k), \Upsilon \times \mathbb{X})$. Итак, реализация совместной стратегии $\mathbb{P}^{all}_{\varrho_0}$ построена, и всякая так построенная стратегия, *пошаговая* стратегия, допустима.

2.3. Допустимость исходной динамики

Покажем, что исходная игра задает допустимую динамику с $\mathbb{X} \overset{\triangle}{=} \mathbb{R}^{d+1}$, $\mathcal{T} \overset{\triangle}{=} [0; T]$, $\Upsilon \overset{\triangle}{=} \bar{\mathbb{U}} \times \bar{\mathbb{V}}$. Действительно, при фиксировании программных стратегий — некоторых càdlàg управлений $\bar{u} = (u, \varphi), \bar{v} = (v, \psi)$ — соответствующий им детерминированный процесс $t \mapsto y(t)$ однозначно определяет дираковскую меру $\mathbb{P}_{(u,v)}[\varrho] = \delta_y$ над множеством своих траекторий из $D(\mathcal{T}, \mathbb{R}^d)$ при дираковском начальном условии ϱ. За счет линейности эти распределения $\mathbb{P}_{(u,v)}[\varrho]$ определены для всевозможных начальных условий ϱ. Далее, каждая пара (φ, ψ) формулами (1.5), (1.6) задает моменты остановки θ_i, а с ними и $\theta_{\min}$. Пусть случайная величина s равна нулю, если игра не завершена ($T = \theta_{\min}$), равна $1 + \theta_{\min}$, если $\theta_1 = \theta_{\min} \in [0; T)$, и равна $-1 - \theta_{\min}$ в случае $\theta_2 = \theta_{\min} \in [0; T)$. Сопоставим каждой паре

$$(x(\cdot), s) \in D([0; T], \mathbb{R}^d) \times ([-T - 1; -1] \cup \{0\} \cup [1; T + 1])$$

траекторию $y_{(x(\cdot),s)}(\cdot) \in D([0; T], \mathbb{X})$ равенствами

$$t \mapsto \big(t, x(\min(t, |s| - 1)), s1_{[0;t]}(|s| - 1)\big) \text{ при } |s| \neq 0 \text{ и } t \mapsto (t, x(t), 0) \text{ иначе.} \tag{2.1}$$

Тогда образ меры $\mathbb{P}_{(u,v)}[\varrho] \otimes \mathbb{P}_{(\varphi,\psi)}$ относительно этого отображения задает распределение $\mathbb{P}^G_{(\bar{u},\bar{v})}[\varrho]$ на $D(\mathcal{T}, \mathbb{X})$. В частности, последняя координата теперь переключается с нулевого значения на значение случайной величины s в момент досрочного окончания игры. Теперь распределение $\mathbb{P}^{all}_{\varrho}$ на $D(\mathcal{T}, \bar{\mathbb{U}} \times \bar{\mathbb{V}}) \times D(\mathcal{T}, \mathbb{X})$ однозначно восстанавливается правилом: для всех борелевских подмножеств $B \subset D(\mathcal{T}, \bar{\mathbb{U}} \times \bar{\mathbb{V}})$, $D \subset D(\mathcal{T}, \mathbb{X})$

$$\mathbb{P}^{all}_{\varrho}(B \times D) \overset{\triangle}{=} \mathbb{P}^G_{(\bar{u},\bar{v})}[\varrho](D) 1_B(\bar{u}, \bar{v}). \tag{2.2}$$

Поскольку в исходной игре для произвольной совместной программной стратегии — пары càdlàg управлений $\bar{u} = (u, \varphi)$, $\bar{v} = (v, \psi)$ — реализация построена, то динамика исходной игры допустима. Следовательно, пошаговые стратегии в исходной игре также допустимы.

3. Достроенная исходная игра

Хотя игровой процесс, соответствующий исходной игре, полностью определен, нам будет удобнее доопределить его так, чтобы он был задан уже на неограниченном промежутке времени и в расширенном фазовом пространстве. Для этого назначим $\mathcal{X} \overset{\triangle}{=} \mathbb{R}_+ \times \mathbb{R}^{d+1}$. Зададим единичные орты этого пространства: $\pi_0 \overset{\triangle}{=} (1, 0, \ldots, 0)$, $\pi_1 \overset{\triangle}{=} (0, 1, \ldots, 0)$, $\ldots$ $\pi_d \overset{\triangle}{=} (0, 0, \ldots, 1, 0)$,

$\pi_{d+1} \overset{\triangle}{=} (0, 0, \ldots, 1)$. Теперь для любого вектора $z \in \mathcal{X}$ его i-ю координату можно записать в виде $\pi_i z \overset{\triangle}{=} \langle \pi_i, z \rangle$. Наконец, через $||\cdot||_d$ будем обозначать норму вектора без учета последней координаты, т. е. $||z||_d \overset{\triangle}{=} ||z - (\pi_{d+1} z)\pi_{d+1}||$.

3.1. Необходимые константы

Заметим, что в силу ограниченности динамики системы найдется компакт $K_< \subset \mathbb{R}^d$, за пределы которого на промежутке $[0;T]$ не может выйти никакое решение y системы (1.1) с начальным условием из единичного шара с центром в x_*. Увеличив при необходимости этот компакт до $K_>$, мы можем считать, что любое решение системы (1.1) (с начальным условием x_*) в любой момент из $[0;T]$ удалено от границ компакта $K_>$ на расстояние, большее единицы. Теперь мы можем также найти такое достаточно большое число $L \geq 2$, что функции $f|_{[0;T]\times K_> \times \mathbb{U}\times\mathbb{V}}$ и $\sigma_0|_{K_>}$, $\sigma_1|_{[0;T]\times K_>}$, $\sigma_2|_{[0;T]\times K_>}$ липшицевы с константой L. Можно также считать, что норма $f|_{[0;T]\times K_> \times \mathbb{U}\times\mathbb{V}}$, как и числа φ_+, ψ_+ также не больше $L-1$. Напомним, что по условию σ_0, σ_1, σ_2 по модулю уже не превосходят единицу; для определенности продолжим также на $[T;\infty)\times\mathbb{R}^d$ отображения $\sigma_1|_{[0;T]\times\mathbb{R}^d}$, $\sigma_2|_{[0;T]\times\mathbb{R}^d}$ с сохранением как их константы Липшица, так и нормы в равномерной метрике; это возможно в силу [28, Theorem 9.58].

Зафиксируем некоторую гладкую монотонно невозрастающую 2-липшицевую скалярную функцию $a : \mathbb{R} \to [0;1]$, для которой $a(0) = 1$, $a(1) = 0$, а функция $r \mapsto a(r-1/2)-1/2$ является нечетной. Введем также $a_K(x) \overset{\triangle}{=} 1 - a(\operatorname{dist}(x; \mathbb{R}^d \setminus K_>))$ для всех $x \in \mathbb{R}^d$.

Зафиксируем некоторое положительное число $h < \min(1, e^{-2}/T, 56L^2Te^{-12LT}/d)$ такое, что $e^{-T/8h} \leq hT$ и $1/h \in \mathbb{N}$. Примем также

$$\gamma \overset{\triangle}{=} \sqrt{-2hT\ln(hT)}, \quad \Gamma \overset{\triangle}{=} T + \gamma = T + \sqrt{-2hT\ln(hT)}.$$

Теперь имеют место неравенства

$$-\ln(hT) \geq 2, \quad 2\sqrt{hT} \leq \gamma, \quad \gamma < T/2, \quad h^2 d/2 \leq 7L^2 e^{-12LT}\gamma^2. \tag{3.1}$$

3.2. Доопределенная динамика исходной системы

Напомним, что согласно предложенной в предыдущем разделе формализации завершение игры по инициативе первого игрока в момент времени θ заключается в сдвиге положения системы по дополнительной $(d+1)$-й координате на вектор $(1+\theta)\pi_{d+1}$, по инициативе второго игрока — на вектор $-(1+\theta)\pi_{d+1}$. Теперь расширенное пространство прежней динамической системы (1.1) совпадает с $\mathbb{R}^{d+1} \times \{s = 0\}$, в множестве $\mathbb{R}^{d+1} \times \{s \geq 1\}$ находятся все мыслимые финальные положения системы, если завершение игры произошло по инициативе первого игрока; аналогично, в множестве $\mathbb{R}^{d+1} \times \{s \leq -1\}$ будут все финальные положения системы при завершении игры по инициативе второго игрока. По этой причине распространим динамику (1.1) на $\mathcal{X} = \mathbb{R}_+ \times \mathbb{R}^{d+1}$.

Для этого каждым $w = (t, x, s) \in \mathcal{X}$, $(u, \varphi) \in \bar{\mathbb{U}}$, $(v, \psi) \in \bar{\mathbb{V}}$ сопоставим

$$\hat{f}(w, \bar{u}, \bar{v}) \overset{\triangle}{=} (1, 1_{[0;T)}(t) a_K(x) f(t, x, u, v), 0). \tag{3.2}$$

Легко проверить, что функция $\hat{f}$ ограничена по норме числом L и является $3L$-липшицевой в каждой из областей $[0;T)\times\mathbb{R}^{d+1}$ и $[T;\infty)\times\mathbb{R}^{d+1}$. Также $\hat{f}(w, \bar{u}, \bar{v})$ совпадает с $(1, f(t, x, u, v), 0)$ при $w = (t, x, s) \in [0;T) \times K_< \times \{0\}$, $(u, \varphi) \in \bar{\mathbb{U}}$, $(v, \psi) \in \bar{\mathbb{V}}$. Более того, за пределами множества $\mathcal{X}_< \overset{\triangle}{=} \mathbb{R}_+ \times K_> \times [-2-T; 2+T]$ функция $\hat{f}$ равна нулю по всем координатам, кроме нулевой.

Рассмотрим конфликтно-управляемую систему

$$\frac{dy(t)}{dt} = \hat{f}(y(t), \bar{u}(t), \bar{v}(t)), \ y(0) = (0, x_*, 0) \text{ при } t \in \mathbb{R}_+, \quad y \in \mathcal{X}, \quad \bar{u}(t) \in \bar{\mathbb{U}}, \quad \bar{v}(t) \in \bar{\mathbb{V}}. \tag{3.3}$$

В силу так введенных $K_<, K_>, \mathcal{X}_<$ при динамике (3.2) траектории не выйдут за пределы $\mathcal{X}_<$, если только их начальные условия были там же; кроме того, для исходных начальных условий при любых управлениях игроков соответствующие координаты у траекторий системы (3.3) совпадают с координатами траекторий системы (1.1), а время совпадает с нулевой координатой решения системы (3.3). Также отметим, что с момента завершения игры, т. е. при $\pi_0 y(t) = t \geq T$ или при $|\pi_{d+1}y(t)| \geq 1$, у решения системы (3.3) изменяется лишь отвечающая за время нулевая координата. Наконец, заметим, что в силу ограниченности правой части для любого решения y с динамикой (3.2) выполнено

$$||y(t) - y(t')||_d \leq L(t - t') \quad \forall t' > t \geq 0. \tag{3.4}$$

Для описания прыжка по последней координате потребуем, чтобы условная вероятность завершения игры по инициативе первого (второго) игрока в течение малого промежутка $[t; t+\Delta t)$ была равна $a(|\pi_{d+1}y|)1_{[0;T)}(\pi_0 y)\varphi(t)\Delta t + o(\Delta t)$ (соответственно $a(|\pi_{d+1}y|)1_{[0;T)}(\pi_0 y)\psi(t)\Delta t + o(\Delta t)$). Тогда моменты положительных и отрицательных прыжков по последней координате — моменты остановки θ_1, θ_2 — описываются правилами: $\theta_1 \overset{\triangle}{=} \inf\{t \,|\, \pi_{d+1}y(t) > \pi_{d+1}y(0)\} \cup \{T\}$, $\theta_2 \overset{\triangle}{=} \inf\{t \,|\, \pi_{d+1}y(t) < \pi_{d+1}y(0)\} \cup \{T\}$, что в случае исходных начальных условий будет соответствовать правилам (1.5), (1.6) при нулевой интенсивности соперника. В общем случае в силу равенства $t = \pi_0 y(t)$ недетерминированный переход системы, соответствующий досрочному завершению игры, а значит, и значение $\theta_{\min} = \min(\theta_1, \theta_2) < T$, может быть описан зависящей от $z, \bar{u} = (u, \varphi), \bar{v} = (v, \psi)$ мерой Леви

$$\hat{\eta}(y, \bar{u}, \bar{v}; A) \overset{\triangle}{=} a(|\pi_{d+1}y|)1_{[0;T)}(\pi_0 y)(\varphi\delta_{(1+\pi_0 y)\pi_{d+1}} + \psi\delta_{-(1+\pi_0 y)\pi_{d+1}})(A) \quad \forall A \in \mathcal{B}(\mathcal{X}).$$

Покажем, что так доопределенная динамика допустима. Снова для $\mathcal{T} \overset{\triangle}{=} \mathbb{R}_+$ каждой паре управлений $\bar{u} = (u, \varphi)$, $\bar{v} = (v, \psi)$ сопоставим дираковские меры $\hat{\mathbb{P}}_{(\bar{u},\bar{v})}[\varrho]$, сосредоточенные на порожденных этими управлениями на всю полуось решениях y уже с динамикой (3.2). Определим по интенсивностям (φ, ψ) моменты остановки θ_1, θ_2, а значит, и $\theta_{\min}$. Так же, как и в предыдущем разделе, определим по θ_1, θ_2 случайную величину s и сопоставим каждой паре $(x(\cdot), s) \in D(\mathbb{R}_+, \mathbb{X}) \times ([-1-T; -1] \cup \{0\} \cup [1; 1+T])$ путь $y_{(x(\cdot),s)}(\cdot) \in D(\mathbb{R}_+, \mathcal{X})$ правилом (2.1). Снова образ меры $\hat{\mathbb{P}}_{(\bar{u},\bar{v})} \otimes \mathbb{P}_{(\varphi,\psi)}$ относительно этого отображения задает распределение $\hat{\mathbb{P}}^G_{(\bar{u},\bar{v})}[\varrho]$ на $D(\mathcal{T}, \mathcal{X})$, после чего распределение $\hat{\mathbb{P}}^{all}_{\varrho}$ на $D(\mathcal{T}, \bar{\mathbb{U}} \times \bar{\mathbb{V}}) \times D(\mathcal{T}, \mathcal{X})$ однозначно восстанавливается аналогом (2.2) — правилом

$$\hat{\mathbb{P}}^{all}_{\varrho}(B \times D) \overset{\triangle}{=} \hat{\mathbb{P}}^G_{(\bar{u},\bar{v})}[\varrho](D)1_B(\bar{u}, \bar{v}) \;\; \forall B \in \mathcal{B}(D(\mathcal{T}, \bar{\mathbb{U}} \times \bar{\mathbb{V}})), \quad D \in \mathcal{B}(D(\mathcal{T}, \mathcal{X})).$$

Итак, в достроенной игре для произвольной пары càdlàg управлений $\bar{u} = (u, \varphi)$, $\bar{v} = (v, \psi)$ построена их реализация, а значит, достроенная динамика допустима. Вследствие этого, как и ранее, имеет место допустимость программных и пошаговых стратегий.

3.3. Платежная функция в виде интеграла

Определим функцию $W : \mathcal{X} \to \mathbb{R}$ правилом: при любых $z = (t, x, s) \in \mathcal{X}$

$$W(z) \overset{\triangle}{=} a_K(x)\big(a(|s|)\sigma_0(x) + a(1-s)\sigma_1(s-1, x) + a(s-1)\sigma_2(1-s, x)\big).$$

Заметим, что функция W по норме не превосходит 1, не зависит от нулевой координаты, а в силу $L \geq 2$ также является $3L$-липшицевой.

Всякому пути $y \in D\big(\mathbb{R}_+, \mathcal{X}\big)$ сопоставим платежную функцию

$$\bar{\sigma}(y) \overset{\triangle}{=} \int_0^{\infty} he^{h(T+2\gamma-t)}1_{[T+2\gamma;\infty)}(\pi_0 y(t))W(y(t))\,dt. \tag{3.5}$$

Покажем следующее представление математического ожидания платежной функции: для всякого случайного процесса с распределением $\hat{\mathbb{P}}^G_{\bar{u},\bar{v}}[\delta_{(0,x_0,0)}] \in \mathcal{P}(D(\mathbb{R}_+,\mathcal{X}))$ и его траектории $t \mapsto y(t) = (t, x(t), s(t))$ с начальными условиями $x(0) \in K_<$, $s(0) = 0$ выполнено

$$\mathbb{E}J(x,\bar{u},\bar{v}) \overset{\triangle}{=} \hat{\mathbb{E}}^G_{\bar{u},\bar{v}}\bar{\sigma}(y) = \hat{\mathbb{E}}^G_{\bar{u},\bar{v}} \int\limits_0^\infty he^{h(T+2\gamma-t)}1_{[T+2\gamma;\infty)}(\pi_0 y(t))W(y(t))\,dt. \tag{3.6}$$

Действительно, рассмотрим произвольную траекторию $t \mapsto y(t) = (t, x(t), s(t))$ процесса. Эта траектория определена на всей полуоси, но, начиная с некоторого момента остановки τ_T, все ее координаты кроме нулевой перестанут меняться. Это произойдет либо по исчерпанию времени игры, тогда $\tau_T = T$ и $W(y(t)) = \sigma_0(x(T))$ для всех $t \geq \tau_T$; либо по инициативе первого игрока, тогда $\tau_T = \theta_1$ — первый момент с $s(t) > 0$, при этом $z(\tau_T) = (\tau_T, x(\tau_T), \tau_T + 1)$ и $W(y(t)) = \sigma_1(\tau_T, x(\tau_T))$ для всех $t \geq \tau_T$; либо по инициативе второго игрока, тогда $\tau_T = \theta_2$ — первый момент с $s(t) < 0$, при этом $W(y(t)) = \sigma_2(\tau_T, x(\tau_T))$ для всех $t \geq \tau_T$. Поскольку до момента $T + 2\gamma$ вдоль той же траектории $y(\cdot)$ имеем $e^{h(T+2\gamma-t)}1_{[T+2\gamma;\infty)}(\pi_0 y(t))W(y(t)) \equiv 0$, то $\bar{\sigma}(y)$ равно в точности $\sigma_0(x(T))$, $\sigma_1(\theta_1, x(\theta_1))$, $\sigma_2(\theta_2, x(\theta_2))$ в каждом из этих случаев. Итак, равенство (3.6) показано.

Поскольку в силу (3.6) математические ожидания платежных функций совпадают, для так доопределенной динамики построенная игра при начальном условии $(0, x_*, 0)$ совпадает с исходной.

4. Вспомогательная марковская игра

Теперь начнем строить аппроксимирующую марковскую игру, зависящую от шага h. Для этого рассмотрим целочисленную решетку $\mathcal{Z} \overset{\triangle}{=} h\mathbb{Z}^{d+2} \cap \mathcal{X}$ шага h размерности $d+2$ с неотрицательной нулевой координатой. В качестве множества состояний возьмем $\mathcal{Z}_< \overset{\triangle}{=} \mathcal{Z} \cap \mathcal{X}_<$. Пространством путей станет пространство Скорохода $D\big(\mathbb{R}_+, \mathcal{Z}_<\big)$, оснастим его σ-алгеброй борелевских множеств и снабдим канонической фильтрацией.

4.1. Рандомизированные стратегии в марковской игре

В качестве стратегий игроков возьмем рандомизированные стратегии.

Под *рандомизированной стратегией* $\bar{\mu}$ *первого игрока* будем понимать пару (μ, φ_μ) отображений, сопоставляющих каждому $w \in \mathcal{Z}_<$ некоторые зависящую от времени вероятностную меру $\mu[w]$ на борелевских подмножествах $\mathbb{U}$ и измеримую по Борелю функцию $\varphi_\mu[w]$ со значениями в $[\varphi_-; \varphi_+]$. Под *рандомизированной стратегией* $\bar{\nu}$ *второго игрока* будем понимать пару (ν, ψ_ν) отображений, сопоставляющих каждому $w \in \mathcal{Z}_<$ некоторые зависящую от времени вероятностную меру $\nu[w]$ на борелевских подмножествах $\mathbb{V}$ и измеримую по Борелю функцию $\varphi_\nu[w]$ со значениями в $[\psi_-; \psi_+]$. Обозначим множества всех рандомизированных стратегий первого и второго игроков через $\check{\mathbb{U}}_\varpi$ и $\check{\mathbb{V}}_\varpi$ соответственно. В случае, если элемент из $\check{\mathbb{U}}_\varpi$ (или $\check{\mathbb{V}}_\varpi$) не зависит от времени, будем называть его *стационарной стратегией*. Обозначим множество всех стационарных стратегий первого и второго игроков 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 $\check{f} : \mathcal{Z}_< \times \check{\mathbb{U}}_\varsigma \times \check{\mathbb{V}}_\varsigma \to \mathcal{Z}_<$ следующим образом: для всякой пары стратегий $(\bar{\mu}, \bar{\nu}) \in \check{\mathbb{U}}_\varsigma \times \check{\mathbb{V}}_\varsigma$ и точки $w \in \mathcal{Z}_<$

$$\check{f}(w, \bar{\mu}, \bar{\nu}) \overset{\triangle}{=} \int\limits_U \int\limits_V \hat{f}(w, u, \varphi_\mu, v, \psi_\nu)\, \mu[w](du)\nu[w](dv).$$

Определим также для всякого $i \in [0 : d+1]$ проекции $\pi_i^+\check{f}$, $\pi_i^-\check{f}$ отображения $\check{f}$ правилами

$$\pi_i^+\check{f}(w,\bar{\mu},\bar{\nu}) \stackrel{\triangle}{=} \int\limits_U \int\limits_V \max(0, \pi_i \hat{f}(w,u,\varphi_\mu[w],v,\psi_\nu[w]))\,\mu[w](du)\nu[w](dv),$$

$$\pi_i^-\check{f}(w,\bar{\mu},\bar{\nu}) \stackrel{\triangle}{=} \int\limits_U \int\limits_V \max(0, -\pi_i \hat{f}(w,u,\varphi_\mu[w],v,\psi_\nu[w]))\,\mu[w](du)\nu[w](dv).$$

Отметим, что в $\check{f}$ (и не только) на место стационарных рандомизированных стратегий мы будем подставлять и значения зависящих от времени рандомизированных стратегий $\bar{\mu},\bar{\nu}$, но при этом для простоты обозначений будем опускать отвечающий за время аргумент.

4.2. Динамика марковской игры

Для всех точек $w \in \mathcal{Z}_<$ и рандомизированных стратегий $\bar{\mu},\bar{\nu}$ игроков определим меру Ито $\check{\eta}(w,\bar{\mu},\bar{\nu};\cdot)$: для всех борелевских подмножеств $A \subset \mathcal{Z}_<$

$$\check{\eta}(w,\bar{\mu},\bar{\nu};A) \stackrel{\triangle}{=} \frac{1}{h}\sum_{i=0}^{d}\left[\pi_i^+\check{f}(w,\bar{\mu},\bar{\nu})\delta_{h\pi_i}(A) + \pi_i^-\check{f}(w,\bar{\mu},\bar{\nu})\delta_{-h\pi_i}(A)\right]$$
$$+ 1_{[0;\Gamma)}(\pi_0 w)a(|\pi_{d+1}w|)\left[\bar{\varphi}[w]\delta_{(\pi_0 w+1)\pi_{d+1}}(A) + \bar{\psi}[w]\delta_{-(\pi_0 w+1)\pi_{d+1}}(A)\right].$$

Для этой меры легко проверить нужные нам далее для построения генератора Леви — Хинчина равенства: для всех $w \in \mathcal{Z}_<$

$$\int\limits_{\mathcal{X}}[y - (\pi_{d+1}y)\pi_{d+1}]\,\check{\eta}(w,\bar{\mu},\bar{\nu};dy) = \int\limits_{\mathcal{X}}\sum_{i=0}^{d}\pi_i y\,\check{\eta}(w,\bar{\mu},\bar{\nu};dy) = \check{f}(w,\bar{\mu},\bar{\nu}), \tag{4.1}$$

$$\int\limits_{\mathcal{X}}||y||_d^2\,\check{\eta}(w,\bar{\mu},\bar{\nu};dy) = h\sum_{i=0}^{d}\left(\pi_i^+\check{f}(w,\bar{\mu},\bar{\nu}) + \pi_i^-\check{f}(w,\bar{\mu},\bar{\nu})\right). \tag{4.2}$$

Введенной мере $\check{\eta}$ можно сопоставить цепь Маркова с непрерывным временем, заданную матрицей Колмогорова $(\bar{Q}_{wy}(\bar{\mu},\bar{\nu}))_{w,y\in\mathcal{Z}_<}$: для всех зависящих от времени рандомизированных стратегий $\bar{\mu},\bar{\nu}$

$$\bar{Q}_{wy}(\bar{\mu},\bar{\nu}) \stackrel{\triangle}{=} \begin{cases} 1/h, & y = w + h\pi_0,\\ \pi_i^+\check{f}(w,\bar{\mu},\bar{\nu})/h, & y = w + h\pi_i, i\in[1:d],\\ \pi_i^-\check{f}(w,\bar{\mu},\bar{\nu})/h, & y = w - h\pi_i, i\in[1:d],\\ 1_{[0;\Gamma)}(\pi_0 w)a(|\pi_{d+1}w|)\varphi_\mu[w], & y = w + (1+\pi_0 w)\pi_{d+1},\\ 1_{[0;\Gamma)}(\pi_0 w)a(|\pi_{d+1}w|)\psi_\nu[w], & y = w - (1+\pi_0 w)\pi_{d+1},\\ -\sum_{i=0}^{d}(\pi_i^+\check{f}(w,\bar{\mu},\bar{\nu}) + \pi_i^-\check{f}(w,\bar{\mu},\bar{\nu}))/h & \\ \ - 1_{[0;\Gamma)}(\pi_0 w)a(|\pi_{d+1}w|)(\varphi_\mu[w]+\psi_\nu[w]), & y = w,\\ 0, & \text{иначе.}\end{cases}$$

Отметим, что для каждого $w \in \mathcal{Z}_<$ как мера $\check{\eta}(w,\bar{\mu},\bar{\nu};\cdot)$ на булеане $\mathcal{Z}_<$, так и соответствующий $\bar{Q}_{wy}$ оператор линейно зависят от $\bar{\mu}$ и $\bar{\nu}$. Теперь из теоремы о минимаксе для любого отображения $q : \mathcal{Z}_< \to \mathbb{R}$ имеем равенство

$$\min_{\bar{\mu}\in\check{\mathbb{U}}_\varsigma}\max_{\bar{\nu}\in\check{\mathbb{V}}_\varsigma}\sum_{y\in\mathcal{Z}_<}\bar{Q}_{wy}(\bar{\mu},\bar{\nu})q(y) = \max_{\check{\nu}\in\check{\mathbb{V}}_\varsigma}\min_{\bar{\mu}\in\check{\mathbb{U}}_\varsigma}\sum_{y\in\mathcal{Z}_<}\bar{Q}_{wy}(\bar{\mu},\bar{\nu})q(y). \tag{4.3}$$

Для любого отображения $q : \mathcal{Z}_< \to \mathbb{R}$ введем (и зафиксируем) стационарные стратегии $\bar{\mu}^{\downarrow q}$, $\bar{\nu}^{\uparrow q}$, реализующие этот минимакс в каждой точке.

4.3. Целевая функция и цена марковской игры

Поскольку все матрицы $\bar{Q}_{wy}$, как и длины прыжков равномерно ограничены, выполнены все предположения [14, Remark 4.2(b); 26, Theorem 5.1; 27, Assumptions 2.1,2.2]. Теперь в силу [14, Proposition 3.1(a); 27, § 2.3] 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 ϱ (вероятности над $\mathcal{Z}_<$ для момента времени 0) имеется порожденный ими процесс $(\check{Y}(t))_{t\geq 0}$, а значит, и его распределение $\check{\mathbb{P}}^G_{\bar{\mu},\bar{\nu}}[\varrho] \in \mathcal{P}(D(\mathbb{R}_+, \mathcal{Z}_<))$. Тогда каждая рандомизированная стратегия является допустимой, и ей можно сопоставить процесс $(\check{Y}(t))_{t\geq 0}$ с распределением

$$\check{\mathbb{P}}^{all}_\varrho \triangleq \delta_{\bar{\mu},\bar{\nu}} \otimes \check{\mathbb{P}}^G_{\bar{\mu},\bar{\nu}}[\varrho] \in \mathcal{P}(D(\mathbb{R}_+, \check{\mathbb{U}}_\varpi \times \check{\mathbb{V}}_\varpi) \times D(\mathbb{R}_+, \mathcal{Z}_<)).$$

Снова каждому элементу пространства путей $z \in D\big(\mathbb{R}_+, \mathcal{Z}_<\big) \subset D\big(\mathbb{R}_+, \mathcal{X}_<\big)$ сопоставим платеж $\bar{\sigma}(z)$ правилом (3.5). Для каждой начальной позиции $z_0 \in \mathcal{Z}_<$ в момент времени 0 игроки могут выбором рандомизированной стратегии обеспечить одно из следующих значений в зависимости от того, кто из них информационно дискриминирован:

$$\check{\mathcal{V}}^+(z_0) = \inf_{\bar{\mu}\in\check{\mathbb{U}}_\varpi} \sup_{\bar{\nu}\in\check{\mathbb{V}}_\varpi} \check{\mathbb{E}}^{all}_{z_0} \int\limits_0^\infty h e^{h(T+2\gamma-t)} 1_{[T+2\gamma;\infty)}(\pi_0 z(t)) W(z(t))\, dt,$$

$$\check{\mathcal{V}}^-(z_0) = \sup_{\bar{\nu}\in\check{\mathbb{V}}_\varpi} \inf_{\bar{\mu}\in\check{\mathbb{U}}_\varpi} \check{\mathbb{E}}^{all}_{z_0} \int\limits_0^\infty h e^{h(T+2\gamma-t)} 1_{[T+2\gamma;\infty)}(\pi_0 z(t)) W(z(t))\, dt.$$

Более того, как следует из [14, Theorem 5.1; 26, Theorem 2], система уравнений (см. [14, (5.4); 26, (11)]),

$$\inf_{\bar{\mu}[w]} \sup_{\bar{\nu}[w]} \sum_{y\in\mathcal{Z}_<} \bar{Q}_{wy}(\bar{\mu},\bar{\nu})\check{\mathcal{V}}(y) = h(\check{\mathcal{V}}(w) - e^{h(T+2\gamma)} 1_{[T+2\gamma;\infty)}(\pi_0 w) W(w))$$
$$= \sup_{\bar{\nu}[w]} \inf_{\bar{\mu}[w]} \sum_{y\in\mathcal{Z}_<} \bar{Q}_{wy}(\bar{\mu},\bar{\nu})\check{\mathcal{V}}(y), \quad w \in \mathcal{Z}_< ,$$

имеет в силу равенства (4.3) единственное решение, и это решение совпадает с $\check{\mathcal{V}} \triangleq \check{\mathcal{V}}^- = \check{\mathcal{V}}^+$. Отметим, что в каждом уравнении системы имеется не более $2d+4$ ненулевых слагаемых по построению матрицы $\bar{Q}$. Более того, в случае $\pi_0 y \geq \Gamma$ значение $\check{\mathcal{V}}(y)$ считается явно и равно $e^{h(T+2\gamma-\pi_0 w)} W(w)$. Таким образом, эта система сводится к системе конечного числа уравнений.

4.4. Оптимальные стационарные стратегии

Из [26, (6); 14, (5.5)–(5.7)] следует, что любые стационарные стратегии $\bar{\mu}^{opt}, \bar{\nu}^{opt}$ игроков, разрешающие для всех $w \in \mathcal{Z}_<$ системы

$$h(\check{\mathcal{V}}(w) - e^{h(T+2\gamma)} 1_{[T+2\gamma;\infty)}(\pi_0 w) W(w)) = \max_{\bar{\nu}\in\check{\mathbb{V}}_\varsigma} \sum_{y\in\mathcal{Z}_<} \bar{Q}_{wy}(\bar{\mu}^{opt},\bar{\nu})\check{\mathcal{V}}(y),$$
$$h(\check{\mathcal{V}}(w) - e^{h(T+2\gamma)} 1_{[T+2\gamma;\infty)}(\pi_0 w) W(w)) = \min_{\bar{\mu}\in\check{\mathbb{U}}_\varsigma} \sum_{y\in\mathcal{Z}_<} \bar{Q}_{wy}(\bar{\mu},\bar{\nu}^{opt})\check{\mathcal{V}}(y),$$

являются оптимальными стратегиями в этой задаче. Таким образом, теперь можно взять $\bar{\mu}^{opt} \equiv \bar{\mu}^{\downarrow\check{\mathcal{V}}}$, $\bar{\nu}^{opt} \equiv \bar{\nu}^{\uparrow\check{\mathcal{V}}}$ в качестве оптимальных стратегий, более того, выбрать их принимающими значения среди мер Дирака над $\hat{\mathbb{U}}$ и $\hat{\mathbb{V}}$ соответственно. В этом случае вторая компонента в $\bar{\mu}^{opt} = (u^{opt}, \varphi^{opt})$ однозначно восстанавливается правилом

$$\varphi^{opt}[w] = \begin{cases} \varphi_-, & \check{\mathcal{V}}(w + \pi_{d+1}(1+\pi_0 w)) > \check{\mathcal{V}}(w), \\ \varphi_+, & \check{\mathcal{V}}(w + \pi_{d+1}(1+\pi_0 w)) < \check{\mathcal{V}}(w) \end{cases}$$

при $|\pi_{d+1}w| < 1$, $\pi_0 w \in [0;\Gamma)$, $\check{\mathcal{V}}(w+\pi_{d+1}(1+\pi_0 w)) \neq \check{\mathcal{V}}(w)$ и может быть произвольной, если хотя бы одно из этих условий не выполнено. Аналогичное свойство имеет место и для $\bar\nu^{opt} = (v^{opt}, \psi^{opt})$.

4.5. Оценки для нулевой координаты, не зависящей от действий игроков

Зафиксируем некоторое $z_0 \in \mathcal{Z}_<$ с $\pi_0 z_0 = 0$ и дираковскую меру ϱ, сосредоточенную в этой точке. Тогда для произвольных, возможно, зависящих от времени рандомизированных стратегий игроков $\bar\mu, \bar\nu$ найдется линейно зависящее от этой дираковской меры распределение $\check{\mathbb{P}}^{all}_{\delta_{z_0}}$ процесса $(\check Y(t))_{t\geq 0}$. Для простоты обозначений примем до конца этого раздела $\check{\mathbb{P}} \stackrel{\triangle}{=} \check{\mathbb{P}}^{all}_{\delta_{z_0}}$.

Отметим, что вне зависимости от действий игроков процесс $\pi_0\check Y(t)$ имеет независимые приращения, а случайные величины $\pi_0(\check Y(t_0+t) - \check Y(t))/h$ распределены по Пуассону с параметром t/h; в частности, их математическое ожидание и дисперсия равны t/h. Тогда $\check{\mathbb{E}}|\pi_0\check Y(t) - t|^2 = ht$; в частности, для $t = T+2\gamma$

$$\check{\mathbb{E}}|\pi_0\check Y(T+2\gamma) - (T+2\gamma)|^2 \leq (T+2\gamma)h \overset{(3.1)}{\leq} 2hT.$$

Наконец, поскольку мартингалом является $\pi_0\check Y(t) - t$, то по неравенству Дуба имеем

$$\check{\mathbb{E}} \max_{t\in[0;T+2\gamma]} |\pi_0\check Y(t) - t|^2 \leq 4\check{\mathbb{E}}|\pi_0\check Y(T+2\gamma) - (T+2\gamma)|^2 \leq 8hT \overset{(3.1)}{\leq} 2\gamma^2. \tag{4.4}$$

Теперь отдельно рассмотрим распределенную по Пуассону с параметром Γ/h случайную величину $\xi = \pi_0\check Y(\Gamma)/h$. Согласно границе Чернова (вариант неравенства Маркова) для всякого отрицательного δ выполнено

$$\check{\mathbb{P}}(\xi \leq \tfrac{T}{h}) \leq \frac{\check{\mathbb{E}}e^{\delta\xi}}{e^{\delta\frac{T}{h}}} = e^{\frac{\Gamma}{h}(e^\delta - 1)}e^{-\delta\frac{T}{h}}.$$

Последний знак равенства здесь — подстановка значения производящей функции (в точке δ) для распределения Пуассона с параметром Γ/h. Теперь, воспользовавшись $e^\delta - 1 \leq \delta + \delta^2/2$ в силу $\delta < 0$, подставляя $\delta = -\gamma/\Gamma$, $T = \Gamma - \gamma$ и $\gamma = \sqrt{-2hT\ln(hT)}$, наконец, получаем

$$\check{\mathbb{P}}(\pi_0\check Y(\Gamma) \leq T) \leq e^{\frac{\Gamma}{h}(\frac{\gamma^2}{2\Gamma^2} - \frac{\gamma}{\Gamma}) + \frac{\gamma T}{h\Gamma}} = e^{\frac{\gamma}{h}(\frac{\gamma}{2\Gamma} - 1 + \frac{\Gamma-\gamma}{\Gamma})} = e^{-\frac{\gamma^2}{2h\Gamma}} = e^{\frac{T\ln(hT)}{\Gamma}} = e^{\frac{\ln(hT)}{1+\gamma/T}}.$$

Аналогично, для $\check{\mathbb{P}}(\pi_0\check Y(\Gamma+\gamma) \leq \Gamma)$ при $\delta = -\gamma/(\Gamma+\gamma)$ имеем

$$\check{\mathbb{P}}(\pi_0\check Y(\Gamma+\gamma) \leq \Gamma) \leq e^{\frac{\Gamma+\gamma}{h}(\frac{\gamma^2}{2(\Gamma+\gamma)^2} - \frac{\gamma}{\Gamma+\gamma}) + \frac{\gamma\Gamma}{h(\Gamma+\gamma)}} = e^{\frac{\gamma}{h}(\frac{\gamma}{2(\Gamma+\gamma)} - 1 + \frac{\Gamma}{\Gamma+\gamma})} = e^{\frac{\ln(hT)}{1+2\gamma/T}}.$$

Снова воспользовавшись $\gamma = \sqrt{-2hT\ln(hT)} < T/2$ и $2\sqrt{hT} \leq \gamma$ (см. (3.1)), получаем

$$\check{\mathbb{P}}(\pi_0\check Y(\Gamma) \leq T) + \check{\mathbb{P}}(\pi_0\check Y(\Gamma+\gamma) \leq \Gamma) \leq e^{\frac{\ln(hT)}{1+\gamma/T}} + e^{\frac{\ln(hT)}{1+2\gamma/T}} < 2\sqrt{hT} \leq \gamma. \tag{4.5}$$

4.6. Оценки на траектории в марковской игре

Задача этого раздела — получить аналог оценки [2, Lemma 12] на нужных классах стратегий.

Зафиксируем некоторую пару рандомизированных стратегий $\bar\mu = (\mu, \varphi_\mu)$, $\bar\nu = (\nu, \psi_\nu)$, а значит, распределение $\check{\mathbb{P}} \stackrel{\triangle}{=} \check{\mathbb{P}}^{all}_{\varrho}$ и меру Ито $\eta(z;\cdot) \stackrel{\triangle}{=} \check\eta(z,\bar\mu,\bar\nu;\cdot)$. Такой мере Ито соответствует генератор Леви—Хинчина (см. [16, (5.1); 27, (2.14)]), сопоставляющий каждой функции $g \in C_c^2(\mathcal{Z}_<)$ отображение $x \mapsto \check{\Lambda}g(x)$ правилом

$$\check{\Lambda}g(x) = \int\limits_{\mathbb{X}} [g(x+y) - g(x)]\eta(x; dy) \quad \forall x \in \mathcal{Z}_<. \tag{4.6}$$

Поскольку все числа $\eta(x;\mathbb{X})$ равномерно по x ограничены, носители мер $\eta(x;\cdot)$ ограничены в совокупности, и $\eta(x;G) \equiv 0$ для некоторой окрестности G нуля, то выполнены все предположения [27, Assumptions 1, 2] и имеет место формула Дынкина [27, Proposition 2.3]: для всех $g \in C_c^2(\mathcal{Z}_<)$ и всех неотрицательных t, r $(t \geq r)$ имеет место

$$\check{\mathbb{E}}\left(g(\hat{Y}(t)) - \int\limits_r^t \check{\Lambda}g(\check{Y}(s))\,ds \,\middle|\, \check{Y}(r)\right) = \check{\mathbb{E}}g(\check{Y}(r)), \tag{4.7}$$

т. е. процесс

$$g(\check{Y}(t)) - \int\limits_r^t \check{\Lambda}g(\check{Y}(s))\,ds \tag{4.8}$$

является мартингалом. Заметим, что в силу равномерной ограниченности носителей меры Ито формула Дынкина (4.7) имеет место и для всех $g \in C^2(\mathcal{X})$, поскольку эти функции всегда можно считать равными нулю вне компактной окрестности объединения всех носителей.

Пусть теперь стратегии $\bar{\mu} = (\mu, \varphi_\mu)$, $\bar{\nu} = (\nu, \psi_\nu)$ пошаговые с некоторым разбиением $(t_k)_{k\in\mathbb{N}}$. В силу допустимости такой стратегии мы снова имеем распределение $\check{\mathbb{P}}_{\varrho}^{all}$. Зафиксируем некоторый момент $t' \stackrel{\triangle}{=} t_{k-1}$ из разбиения и элемент w' в $\mathcal{X}$, возможно, зависящий от $\check{Y}(t')$. Теперь на промежутке $[t'; t'') \stackrel{\triangle}{=} [t_{k-1}; t_k)$ стратегия $\bar{\mu} = (\mu, \varphi_\mu)$, $\bar{\nu} = (\nu, \psi_\nu)$ как программная стратегия, возможно, зависящая от $\check{Y}(\cdot)|_{[0;t')}$, также имеет свой генератор Леви—Хинчина (4.6), для которого справедлива формула Дынкина (4.7).

Обозначим через $\mathbb{E}_{y'}^{all}$ условное математическое ожидание относительно $y' = \check{Y}(t')$. Применим построенный генератор (4.6) к функции $g(w) \stackrel{\triangle}{=} ||w - w'||_d^2$ на промежутке $[t'; t'')$. Тогда согласно (4.8) мартингалом является

$$g(\check{Y}(t)) + \int\limits_{t_k}^t \left[\sum_{i=0}^d 2\pi_i \check{f}(\check{Y}(s), \bar{\mu}, \bar{\nu}) \cdot \pi_i(\check{Y}(s) - w') - \int\limits_{\mathcal{Z}_<} ||y||_d^2\, \eta(\check{Y}(t); dy)\right] ds, \tag{4.9}$$

откуда следует, что для всех $t \in [t'; t'')$ выполнено

$$\check{\mathbb{E}}_{y'}^{all} g(\check{Y}(t)) - g(y') = \check{\mathbb{E}}_{y'}^{all} \int\limits_{t'}^t \left[\int\limits_{\mathcal{Z}_<} ||y||_d^2\, \eta(\check{Y}(s); dy) - \sum_{i=0}^d 2\pi_i \check{f}(\check{Y}(s), \bar{\mu}, \bar{\nu}) \cdot \pi_i(\check{Y}(s) - w')\right] ds.$$

Таким образом, имеет место неравенство

$$\check{\mathbb{E}}_{y'}^{all} g(\check{Y}(t)) - g(y') \leq \check{\mathbb{E}}_{y'}^{all} \int\limits_{t'}^t \left[2||\check{f}(\check{Y}(s), \bar{\mu}, \bar{\nu})||_d \sqrt{g(\check{Y}(s))} + \int\limits_{\mathcal{Z}_<} ||y||_d^2\, \eta(\check{Y}(s); dy)\right] ds. \tag{4.10}$$

Примем $M \stackrel{\triangle}{=} \sqrt{hL\sqrt{d+1} + L^2}$. Оценим удвоенное произведение в (4.10) суммой квадратов. В силу ограниченности (4.1) и (4.2) по норме числами L и $hL\sqrt{d+1}$ имеем

$$\check{\mathbb{E}}_{y'}^{all}\big(g(\check{Y}(t)) - g(y')\big) \leq M^2(t - t') + \int\limits_{t'}^t \check{\mathbb{E}}_{y'}^{all} g(\check{Y}(s))\,ds,$$

откуда по неравенству Гронуолла для всех $t \in [t'; t'')$ следует

$$\check{\mathbb{E}}^{all}_{y'} g(\check{Y}(t)) \leq (M^2 + g(y'))e^{t-t'} - M^2, \tag{4.11}$$

т. е. $e^{-t}\check{\mathbb{E}}^{all}_{y'}(g(\check{Y}(t)) + M^2) \leq e^{-t'}(g(y') + M^2) = e^{-t'}(g(\check{Y}(t')) + M^2)$. Таким образом, на всяком промежутке $[t'; t'') \stackrel{\triangle}{=} [t_{k-1}; t_k)$ разбиения супермартингалом становится процесс

$$e^{-t}(||\check{Y}(t) - w'||^2 + M^2). \tag{4.12}$$

Поскольку переключение пошаговой стратегии происходит лишь в заранее заданные моменты времени, складывая по всем промежуткам разбиения, получаем супермартингал (4.12) уже для всех $t \geq t'$, равно как и неравенство (4.11). При этом из (4.11) при $w' = y' = \check{Y}(t')$ следует

$$\check{\mathbb{E}}^{all}_{y'} ||\check{Y}(t) - y'||_d \leq M\sqrt{e^{t-t'} - 1}. \tag{4.13}$$

Теперь в силу $||\check{f}||_d \leq L$ и оценки $hL\sqrt{d+1}$ сверху для нормы у (4.2) из (4.10) следует

$$\check{\mathbb{E}}^{all}_{y'} g(\check{Y}(t)) - g(y') \stackrel{(4.10)}{\leq} \check{\mathbb{E}}^{all}_{y'} \int\limits_{t'}^{t} \left[2L\sqrt{g(\check{Y}(s))} + hL\sqrt{d+1}\right] ds$$

для всех t из промежутка разбиения $[t'; t'')$. Подставляя сюда при $w' = y' = \check{Y}(t')$ оценку (4.13), получаем

$$\begin{aligned}
\check{\mathbb{E}}^{all}_{y'} ||\check{Y}(t) - y'||^2_d &\stackrel{(4.10)}{\leq} \check{\mathbb{E}}^{all}_{y'} \int\limits_{t'}^{t} \left[2L\sqrt{g(\check{Y}(s))} + hL\sqrt{d+1}\right] ds \\
&\stackrel{(4.13)}{\leq} hL\sqrt{d+1}|t - t'| + 2L\int\limits_{t'}^{t} M\sqrt{e^{s-t'} - 1}\, ds \\
&= hL\sqrt{d+1}|t - t'| + 4LM(\sqrt{e^{t-t'} - 1} - \operatorname{arctg}\sqrt{e^{t-t'} - 1}) \\
&\leq hL\sqrt{d+1}|t - t'| + 4LM(e^{t-t'} - 1)^{3/2}/3,
\end{aligned} \tag{4.14}$$

где в последнем неравенстве применено $\operatorname{arctg} x \geq x - x^3/3$. Итак, в силу (4.14) на всяком промежутке $[t'; t'')$ из выбранного для пошаговой стратегии разбиения при любом $t \in [t'; t'')$

$$\check{\mathbb{E}}^{all}_{\check{Y}(t')} ||\check{Y}(t) - \check{Y}(t')||^2_d \leq hL\sqrt{d+1}|t - t'| + 4L\sqrt{hL\sqrt{d+1} + L^2}(e^{t-t'} - 1)^{3/2}/3. \tag{4.15}$$

Отметим, наконец, что поскольку исследуемые нами функции g не зависят от последней координаты, то ни выкладки, ни полученные неравенства не зависят от выбора φ_μ, φ_ν.

5. Схема поводыря и двойная игра

В качестве фазового пространства новой игры возьмем произведение $\mathbb{X} \stackrel{\triangle}{=} \mathcal{X}_< \times \mathcal{Z}_<$, теперь пространством путей станет пространство Скорохода $D(\mathbb{R}_+, \mathcal{X}_< \times \mathcal{Z}_<)$. Компонента $\mathcal{X}_<$ будет содержать исходную игру, игра на компоненте $\mathcal{Z}_<$ (кроме последней координаты) будет рассчитываться первым игроком, а ее положение и станет тем поводырем, что будет помогать построить стратегию и в $\mathcal{X}_<$. При этом множествами управляющих параметров игроков станут $\Upsilon_1 \stackrel{\triangle}{=} \mathbb{U} \times \check{\mathbb{U}}_\varsigma \times \mathbb{V}_\varsigma$, $\Upsilon_2 \stackrel{\triangle}{=} \bar{\mathbb{V}}$. Тогда всевозможные их совместные мгновенные управления составят множество $\Upsilon \stackrel{\triangle}{=} \mathbb{U} \times \check{\mathbb{U}}_\varsigma \times \mathbb{V}_\varsigma \times \bar{\mathbb{V}}$. Для того чтобы описать допустимую динамику двойной игры, построим каждой программной стратегии $\upsilon(\cdot) \stackrel{\triangle}{=} (u,\mu,\varphi,\nu,v,\psi)(\cdot) \in D(\mathbb{R}_+,\Upsilon)$ линейно зависящее от начального условия $\varrho \in \mathcal{P}(\mathbb{X})$ распределение $\widetilde{\mathbb{P}}^G_\upsilon[\varrho]$. Постулируя линейность от ϱ, можно далее ограничиться случаем, когда ϱ — мера Дирака, сосредоточенная в некоторой точке $(x,z) \in \mathcal{X}_< \times \mathcal{Z}_<$.

Построим сначала вспомогательные процессы $(\hat{Y}_t)_{t\geq 0}$, $(\check{Y}_t)_{t\geq 0}$, для чего рассмотрим на $\mathcal{X}_<$ динамику $\hat{\mathbb{P}}^G_{(u,0,v,\psi)}[\delta_x](\cdot)$ дифференциальной игры с нулевой интенсивностью первого игрока на последней координате, а для $\mathcal{Z}_<$ введем динамику $\check{\mathbb{P}}^G_{(\mu,\varphi,\nu,0)}[\delta_z](\cdot)$ марковской игры с нулевой интенсивностью второго игрока на последней координате. Обе эти динамики допустимые, найдутся соответствующие им процессы $(\hat{Y}_t)_{t\geq 0}$, $(\check{Y}_t)_{t\geq 0}$. При этом, поскольку $\pi_{d+1}\hat{Y}$ и $\pi_{d+1}\check{Y}$ управляются интенсивностями ψ,φ соответственно, то можно определить $\hat{\theta}_2,\check{\theta}_1$ как моменты остановки игры в силу каждого из игроков, а значит, и момент остановки $\theta_{\min} \stackrel{\triangle}{=} \min(\check{\theta}_1,\hat{\theta}_2)$. Теперь каждой паре траекторий $(\hat{Y}_t,\check{Y}_t)_{t\geq 0}$ вместе с моментами остановки $\hat{\theta}_2,\check{\theta}_1$ можно сопоставить траекторию $(\hat{w},\check{w})(\cdot) \in D([0;\infty),\mathbb{R}^2)$ правилом: $(\hat{w},\check{w}) \equiv (0,0)$ при $\theta_{\min} \geq T$,

$$(\hat{w},\check{w}) \equiv (\pi_{d+1}\hat{Y}_t)\Big(1,\frac{1+\pi_0\check{Y}_{\hat{\theta}_2}}{1+\hat{\theta}_2}\Big) \quad \text{и} \quad (\hat{w},\check{w}) \equiv (\pi_{d+1}\check{Y}_t)\Big(\frac{1+\check{\theta}_1}{1+\pi_0\check{Y}_{\check{\theta}_1}},1\Big)$$

при $\theta_{\min} = \hat{\theta}_2 < T$ и $\theta_{\min} = \check{\theta}_1 < T$ соответственно.

Определим, наконец, процесс $(Y'(t),Y''(t))_{t\geq 0}$ со значениями в $\mathcal{X}_< \times \mathcal{Z}_<$ как образ процесса $(\hat{Y}_t,\check{Y}_t,(\hat{w},\check{w})(t))_{t\geq 0}$ в силу следующего неупреждающего отображения:

$$(\hat{y}_t,\check{y}_t,(\hat{w},\check{w})(t)) \mapsto \Big(y'(t) \stackrel{\triangle}{=} (\pi_0\hat{y}_t,\ldots,\pi_d\hat{y}_t,\hat{w}(t)), y''(t) \stackrel{\triangle}{=} (\pi_0\check{y}_t,\ldots,\pi_d\check{y}_t,\check{w}(t))\Big).$$

Полученный нами процесс задает динамику для каждой совместной программной стратегии из $D(\mathbb{R}_+,\Upsilon)$. Таким образом, допустимая динамика построена.

Отметим нужные далее свойства процесса. Во-первых, динамика в каждой компоненте по всем координатам, кроме последней, полностью совпадает с описанными ранее для достроенной исходной игры и марковской игры. Во-вторых, в каждой компоненте длины прыжков и их интенсивности вплоть до момента $\min\{t\,|\,\pi_0 Y''(t) \geq \Gamma\} \cup \{T\}$ подчиняются тем же мерам Леви, что и в указанных ранее играх. В-третьих, до момента времени T последние координаты в каждой компоненте, будучи нулевыми в начальный момент, становятся ненулевыми только одновременно. Таким образом, недетерминированность компоненты $Y'(t)$ полностью описывается недетерминированностью последней координаты компоненты $Y''(t)$, а значит, для расчета последней координаты компоненты $Y''(t)$ не требуется в ходе игры знать интенсивность второго игрока, поскольку эта координата может измениться лишь одновременно с завершением игры. Наконец, это означает, что управление второго игрока можно подать в формате управления исходной игры, что и было сделано выбором $\Upsilon_1 = \mathbb{U} \times \check{\mathbb{U}}_\varsigma \times \mathbb{V}_\varsigma$, $\Upsilon_2 = \bar{\mathbb{V}}$.

5.1. Нацеливание

При построении стратегии первого игрока и воображаемого им ответа второго игрока нам потребуется нацеливаться и отклоняться от позиций на каждой компоненте. Введем необходимые для этого функции.

Прежде всего заметим, что в силу (1.2) для любых векторов $x,w \in \mathcal{X}$ найдутся управления $u^{to\ w}(x) \in \mathbb{U}$, $v^{from\ w}(x) \in \mathbb{V}$ такие, что для всех $(\varphi,\psi) \in [\varphi_-;\varphi_+] \times [\psi_-;\psi_+]$ выполнено

$$\begin{aligned}\min_{u\in\mathbb{U}}\max_{v\in\mathbb{V}}\langle x-w,\hat{f}(x,u,\varphi,v,\psi)\rangle &= \max_{v\in\mathbb{V}}\min_{u\in\mathbb{U}}\langle x-w,\hat{f}(x,u,\varphi,v,\psi)\rangle \\ = \min_{u\in\mathbb{U}}\langle x-w,\hat{f}(x,u,\varphi,v^{from\ w}(x),\psi)\rangle &= \max_{v\in\mathbb{V}}\langle x-w,\hat{f}(x,u^{to\ w}(x),\varphi,v,\psi)\rangle.\end{aligned}$$

Введем также для любого вектора $w \in \mathcal{X}$ управление первого игрока правилом

$$\mathcal{X} \ni x \mapsto \bar{u}^{to\ w}(x) \stackrel{\triangle}{=} (u^{to\ w}(x),\varphi^{opt}[w]) \in \bar{\mathbb{U}}.$$

При этом для всех $x, w \in [0;T) \times \mathbb{R}^{d+1}$ в силу липшицевости $\hat{f}$ имеем

$$\begin{gathered}\langle x-w, \hat{f}(x, u^{to\ w}(x), \varphi, v^{from\ w}(x), \psi) - \hat{f}(w, u^{to\ x}(w), \varphi, v^{from\ x}(w), \psi)\rangle \\ = \min_{u\in\mathbb{U}, v'\in\mathbb{V}} \max_{v\in\mathbb{V}, u'\in\mathbb{U}} \langle x-w, \hat{f}(x,u,\varphi,v,\psi) - \hat{f}(w,u',\varphi,v',\psi)\rangle \le 3L||x-w||_d^2,\end{gathered}$$

где неравенство выполнено в силу подстановки $u = u', v' = v$.

Для того чтобы определить поведение второго игрока на компоненте $\mathcal{Z}_<$, зададим для каждого $x \in \mathcal{X}$ стационарную рандомизированную стратегию второго игрока в виде

$$\mathcal{Z}_< \ni w \mapsto \bar{\nu}^{from\ x}[w] = (\delta_{v^{from\ x}(w)}, \psi^{opt}).$$

Тогда для любой стационарной рандомизированной стратегии $\bar{\mu}$ первого игрока выполнено

$$\begin{gathered}\langle x-w, \check{f}(w, \bar{\mu}, \bar{\nu}^{from\ x})\rangle \ge \min_{\bar{\mu}\in\check{\mathbb{U}}_\varsigma} \langle x-w, \check{f}(w,\bar{\mu},\bar{\nu}^{from\ x})\rangle \\ \overset{(1.2)}{=} \min_{\bar{u}\in\bar{\mathbb{U}}} \max_{\bar{v}\in\bar{\mathbb{V}}} \langle x-w, \hat{f}(w,\bar{u},\bar{v})\rangle = \langle x-w, \check{f}(w, \bar{u}^{to\ x}, \bar{v}^{from\ x})\rangle,\end{gathered}$$

откуда для всех $\bar{v}, \bar{\mu}, \varphi, \psi$ в случае $x, w \in [0;T) \times \mathbb{R}^{d+1}$ имеем

$$\begin{gathered}\langle x-w, \hat{f}(x, u^{to\ x}(w), \varphi, \bar{v}) - \check{f}(w, \bar{\mu}, \bar{\nu}^{from\ w}(x))\rangle \\ \le \langle x-w, \hat{f}(x, u^{to\ x}(w), \varphi, v^{from\ x}(w), \psi) - \check{f}(w, \bar{\mu}^{to\ w}, \bar{\nu}^{from\ w})\rangle \le 3L||x-w||_d^2.\end{gathered}$$

Зафиксируем некоторую пару $(x', w') \in \mathcal{X} \times \mathcal{Z}_<$. Рассмотрим не зависящее от $(d+1)$-х координат своих аргументов отображение $R_{x',w'} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+$, заданное правилом

$$\hat{R}_{x',w'}(x) = \sum_{i=0}^{d} (\pi_i x' - \pi_i w')(\pi_i x - \pi_i x') \quad \forall x \in \mathcal{X}. \tag{5.1}$$

Для получения аналога (4.9) каждой стационарной стратегии $\bar{\mu}$ сопоставим порожденный совместной стационарной стратегией $(\bar{\mu}, \bar{\nu}^{from\ x'}(w'))$ для функции $g(w) \overset{\triangle}{=} R_{w',x'}(w)$ генератор (4.6):

$$\check{\Lambda}[\bar{\mu}, \bar{\nu}^{from\ x'}(w')] R_{w',x'}(w) = \sum_{i=0}^{d} (\pi_i w' - \pi_i x') \pi_i \hat{f}(w, \bar{\mu}, \bar{\nu}^{from\ x'}) = -\langle x'-w', \hat{f}(w, \bar{\mu}, \bar{\nu}^{from\ x'})\rangle.$$

Теперь для $x, w, x', w' \in [0;T] \times \mathbb{R}^{d+1}$ в силу липшицевости $\hat{f}$ имеем

$$\check{\Lambda}[\bar{\mu}, \bar{\nu}^{from\ x'}(w')] R_{w',x'}(w) \le -\langle x'-w', \hat{f}(w', \bar{\mu}, \bar{v}^{from\ x'})\rangle + 3L||w-w'||_d \cdot ||x'-w'||_d. \tag{5.2}$$

Поскольку для всякого порожденного совместным управлением $(\bar{u}, \bar{v})$ решения $t \mapsto y(t)$ системы (3.3) имеет место

$$\frac{dg(y(t))}{dt} - \Big\langle \frac{\partial g(x)}{\partial x}\Big|_{x=y(t)}, \hat{f}(y(t), \bar{u}(t), \bar{v}(t))\Big\rangle = 0$$

для всякой гладкой функции g, то, подобно (4.9), для $\hat{\Lambda}[\bar{u}, \bar{v}]g(x) \overset{\triangle}{=} \Big\langle \frac{\partial g(x)}{\partial x}|_{x=y(t)}, \hat{f}(x, \bar{u}, \bar{v})\Big\rangle$ мартингалом становится процесс

$$g(Y'(t)) - \int_{t_0}^{t} \hat{\Lambda}[\bar{u}, \bar{v}] g(Y'(s))\, ds \tag{5.3}$$

для всех гладких функций $g : \mathcal{X} \to \mathbb{R}$, не зависящих от последней координаты; в частности, выполнена формула Дынкина (4.7). Кроме того, при этом выполнено

$$\hat{\Lambda}[\bar{u}^{to\ w'}(x'), \bar{v}]R_{x',w'}(x) = \sum_{i=0}^{d}(\pi_i x' - \pi_i w')\pi_i \hat{f}(x, \bar{u}^{to\ w'}(x'), \bar{v})$$
$$= \langle x' - w', f(x, \bar{u}^{to\ w'}(x'), \bar{v})\rangle \leq \langle x' - w', f(x', \bar{u}^{to\ w'}(x'), \bar{v})\rangle + 3L||x - x'||_d \cdot ||x' - w'||_d \qquad (5.4)$$

в случае $x, w, x', w' \in [0;T] \times \mathbb{R}^{d+1}$. Складывая (5.4) с (5.2), получаем для всех таких x, w, x', w' при любых $\bar{\mu} \in \check{\mathbb{U}}_\varsigma$, $\bar{v} \in \bar{\mathbb{V}}$ оценку

$$\hat{\Lambda}[\bar{u}^{to\ w'}(x'), \bar{v}]R_{x',w'}(x) + \check{\Lambda}[\bar{\mu}, \bar{\nu}^{from\ x'}]R_{w',x'}(w) \leq \frac{3L}{2}\big(4||x' - w'||_d^2 + ||x - x'||_d^2 + ||w - w'||_d^2\big). \qquad (5.5)$$

5.2. Конструкция поводыря для первого игрока

Пусть в некоторые заранее заданные моменты времени $(t_k)_{k\in\mathbb{N}}$ первый игрок считывает позицию траектории $x(\cdot)$ реальной игры. Будем считать эту последовательность моментов времени возрастающей, полагая также, что некоторое t_k равно T, а $t_0 = 0$. Определим по этой последовательности моментов времени мелкость разбиения $r \stackrel{\triangle}{=} \max_{k\in\mathbb{N}, t_k \leq T} |t_k - t_{k-1}|$ и отображение $\mathbb{R}_+ \ni t \mapsto \tau_\Delta(t) \stackrel{\triangle}{=} \max\{t_i \,|\, t_i \leq t\}$.

Опишем пошаговую стратегию первого игрока при выбранном разбиении $(t_k)_{k\in\mathbb{N}}$.

Пусть в каждый момент времени первый игрок применяет на компоненте $\mathcal{Z}_<$ оптимальную стационарную стратегию μ^{opt} , в частности, он ставит $\varphi^{opt}[z(t)]$ в качестве своей интенсивности завершения игры. Тогда $\varphi^{opt}[z(t)]$ он ставит и на первой компоненте, т. е. в реальной игре.

Пусть на первой, реальной, компоненте он старается минимизировать расстояние от $z(\cdot)$ до $x(\cdot)$ следующим образом: на промежутке $[t_{k-1}; t_k)$ он ставит управление, наилучшее (в момент разбиения t_{k-1}) для смещения x к $z(t_{k-1})$. Таким образом он играет по правилу: $\mathbb{R}_+ \ni t \mapsto (u^{to\ z(\tau_\Delta(t))}(x(\tau_\Delta(t)), \mu^{opt}[z], \varphi^{opt}[z]) \in \Upsilon_1 \stackrel{\triangle}{=} \mathbb{U} \times \check{\mathbb{U}}_\varsigma$. Далее, в рамках рассчитываемой первым игроком на $\mathcal{Z}_<$ подыгры пусть воображаемый им второй игрок старается максимизировать расстояние от $z(\cdot)$ до $x(\cdot)$, применяя стационарную стратегию, наихудшую для смещения z в сторону $x(\tau_\Delta(t))$. Но тогда в качестве ν подается зависящая неупреждающе от x и стационарно от z стратегия $\mathbb{R}_+ \ni t \mapsto \nu^{from\ x(\tau_\Delta(t))}$. Итак, задана следующая стратегия первого игрока: в каждый момент времени

$$\widetilde{v}_t[x, z] = (u^{to\ z(\tau_\Delta(t))}(x(\tau_\Delta(t))), \mu^{opt}[z], \varphi^{opt}[z], \nu^{from\ x(\tau_\Delta(t))}) \in \Upsilon_1. \qquad (5.6)$$

Поскольку эта стратегия является пошаговой, она допустима.

Все необходимые конструкции построены, и можно приступить к доказательству теоремы.

6. Доказательство теоремы 1

Предположим, что второй игрок обладает всей информацией о самой стратегии первого игрока, в частности, он знает как о правиле (5.6), так и о позиции x и z в каждый момент t, а значит, и о реализующейся по ходу игры траектории процесса двойной игры первого игрока. Зафиксируем произвольную контрстратегию второго игрока, неупреждающе сопоставляющую каждой траектории $(x, z)(\cdot)$ двойной игры распределение $\widetilde{\mathbb{Q}}^{II}[\widetilde{\mathbb{P}}^I]_{(x,z)(\cdot)}$ на пространстве Скорохода $D(\mathbb{R}_+, \mathbb{U} \times \check{\mathbb{U}}_\varsigma \times \mathbb{V}_\varsigma \times \bar{\mathbb{V}})$, проекция которого на $D(\mathbb{R}_+, \mathbb{U} \times \check{\mathbb{U}}_\varsigma \times \mathbb{V}_\varsigma)$ совпадет с распределением $\widetilde{\mathbb{P}}^I$, выбранным первым игроком вместе с правилом (5.6). В частности, в качестве контрстратегии $\widetilde{\mathbb{Q}}^{II}$ мы могли взять произвольный элемент $\mathbb{Q}^{II}$ в $\mathfrak{Q}^{II}$, т. е. контрстратегию исходной игры; в этом случае отображение выше не зависело бы от z.

Зафиксируем начальные условия ($x_0 \stackrel{\triangle}{=} (0, x_*, 0), z_0 = (0, y_0, 0)$) со свойствами $||x_* - y_0||^2 \leq 7L^2\gamma^2 e^{-12LT}$, $y_0 \in h\mathbb{R}^d$. Для этого достаточно (см. (3.1)) в качестве y_0 выбрать ближайший к x_* элемент из $h\mathbb{Z}^d$. В качестве стратегии $\mathbb{P}^I$ первого игрока возьмем описанную выше пошаговую стратегию, зафиксируем произвольную реализацию $\widetilde{\mathbb{P}}^{all}_{\delta_{(x_0,z_0)}}$ получившейся совместной стратегии $\widetilde{\mathbb{P}}^{I+II}$. Тем самым нами заданы и отслеживающая лишь положение системы вероятность $\widetilde{\mathbb{P}}^{\chi}_{\delta_{(x_0,z_0)}}$, и не содержащая поводыря реализация $\hat{\mathbb{P}}^{all}_{\delta_{x_0}}$. В частности, при сделанном вторым игроком выборе контрстратегии $\mathbb{Q}^{II} = \widetilde{\mathbb{Q}}^{II}$ среди контрстратегий исходной игры распределение $\hat{\mathbb{P}}^{all}_{\delta_{x_0}}$ становится реализацией некоторой совместной стратегии из $\mathbb{Q}^{II}[\mathbb{P}^I]$.

Поскольку начальное положение зафиксировано, мы далее будем опускать соответствующие нижние индексы.

6.1. Расхождение траекторий

Определим момент остановки $\tau \stackrel{\triangle}{=} \min\{t \,|\, \pi_0 Y''(t) \geq T\} \cup \{T\}$. Рассмотрим два последовательных момента t_{k-1} и t_k из разбиения, а с ними и моменты остановки $t' \stackrel{\triangle}{=} \min(t_{k-1}, \tau)$, $t'' \stackrel{\triangle}{=} \min(t_k, \tau)$. На промежутке $[t'; t'')$ первым игроком осуществляется константное управление $(\bar{u}^{to\ w'}(x'), \bar{\mu}^{opt}, \nu^{from\ x'})$, где $x' = Y'(t')$, $\quad w' = Y''(t')$. Дезинтегрируя $\widetilde{\mathbb{P}}^{\chi}$ по $x' = Y'(t'), w' = Y''(t')$, рассмотрим для всех $x' \in \mathcal{X}, w' \in \mathcal{Z}_<$ соответствующее ему условное математическое ожидание $\widetilde{\mathbb{E}}^{\chi}_{t',x',w'}$ разности $||Y''(t'') - Y'(t'')||^2_d - ||Y''(t') - Y'(t')||^2_d$.

Из определений $R_{x',w'}, R_{w',x'}$ (см. (5.1)) и цепочки равенств

$$||w - x||^2_d - ||w - x + x' - w'||^2_d = -2\sum_{i=0}^{d} \pi_i(x' - w') \cdot \pi_i(w - x) - ||x' - w'||^2_d$$

$$= 2\sum_{i=0}^{d} \pi_i(x' - w') \cdot \pi_i(x - x' + w' - w) + ||x' - w'||^2_d = 2R_{x',w'}(x) + 2R_{w',x'}(w) + ||x' - w'||^2_d$$

следует неравенство $\big(||x - w||^2_d - ||x' - w'||^2_d\big)/2 \leq ||x - x'||^2_d + ||w - w'||^2_d + R_{x',w'}(x) + R_{w',x'}(w)$. Теперь, сокращая правую часть неравенства до $S(x, w, x', w') \stackrel{\triangle}{=} ||x - x'||^2_d + R_{x',w'}(x) + R_{w',x'}(w) + ||w - w'||^2_d$, получаем

$$\widetilde{\mathbb{E}}^{\chi}_{t',x',w'} ||Y'(t'') - Y''(t'')||^2_d \leq ||x' - w'||^2_d + 2\widetilde{\mathbb{E}}^{\chi}_{t',x',w'} S(Y'(t''), Y''(t''), x', w'). \tag{6.1}$$

Оценим математические ожидания от R и S. Подставляя момент остановки t'', поскольку (4.8) и (5.3) являются мартингалами, мы имеем

$$\widetilde{\mathbb{E}}^{\chi}_{t',x',w'} R_{x',w'}(Y'(t'')) = \widetilde{\mathbb{E}}^{\chi}_{t',x',w'} \int_{t'}^{t''} \hat{\Lambda}[\bar{u}^{to\,w'}(x'), \bar{v}] R_{x',w'}(Y'(s))\, ds,$$

$$\widetilde{\mathbb{E}}^{\chi}_{t',x',w'} R_{w',x'}(Y''(t'')) = \widetilde{\mathbb{E}}^{\chi}_{t',x',w'} \int_{t'}^{t''} \check{\Lambda}[\bar{\mu}^{opt}, \bar{v}^{from\,x'}(w')] R_{w',x'}(Y''(s))\, ds.$$

Складывая, в силу неравенства (5.5) получаем

$$\widetilde{\mathbb{E}}^{\chi}_{t',x',w'} \big(R_{x',w'}(Y'(t'')) + R_{w',x'}(Y''(t''))\big)$$

$$= \widetilde{\mathbb{E}}^{\chi}_{t',x',w'} \int_{t'}^{t''} \Big[\hat{\Lambda}[\bar{u}^{to\,w'}(x'), \bar{v}] R_{x',w'}(Y'(s)) + \check{\Lambda}[\bar{\mu}^{opt}, \bar{v}^{from\,x'}(w')] R_{w',x'}(Y''(s))\Big]\, ds$$

$$\leq \frac{3L}{2}\int\limits_{t'}^{t''}\widetilde{\mathbb{E}}^{\chi}_{t',x',w'}\Big(4||x'-w'||_d^2+||Y'(s)-x'||_d^2+||Y''(s)-w'||_d^2\Big)\,ds.$$

Таким образом, для S выполнено неравенство

$$\widetilde{\mathbb{E}}^{\chi}_{t',x',w'}\Big[S(Y'(t''),Y''(t''),x',w')-||Y'(t'')-x'||_d^2-||Y''(t'')-w'||_d^2\Big]$$
$$\leq 3L\widetilde{\mathbb{E}}^{\chi}_{t',x',w'}\int\limits_{t'}^{t''}\Big(2||x'-w'||_d^2+\frac{||Y'(s)-x'||_d^2+||Y''(s)-w'||_d^2}{2}\Big)\,ds.$$

В силу оценок (3.4) и (4.15) для $M^2=hL\sqrt{d+1}+L^2$ показано

$$\widetilde{\mathbb{E}}^{\chi}_{t',x',w'}||Y'(t)-x'||_d^2\overset{(3.4)}{\leq}L^2(t_{k-1}-t_k)^2,$$
$$\widetilde{\mathbb{E}}^{\chi}_{t',x',w'}||Y''(t)-w'||_d^2\overset{(4.15)}{\leq}L(t-t_{k-1})\Big(h\sqrt{d+1}+\frac{4M(e^{t-t_{k-1}}-1)^{3/2}}{3(t-t_{k-1})}\Big)$$

для всех $t\in[t_{k-1};t_k]$. Подставляя для $\widetilde{\mathbb{E}}^{\chi}_{t',x',w'}S(Y'(t''),Y''(t''),x',w')$ в качестве t момент остановки $t''\leq t'+t_{k-1}-t_k$, установив $g(r)\overset{\triangle}{=}(h\sqrt{d+1}+4M(e^r-1)^{3/2}/3r+Lr)(1+Lr)/6$, где $r=\max_{t_i\leq T}(t_i-t_{i-1})$, имеем

$$\widetilde{\mathbb{E}}^{\chi}_{t',x',w'}S(Y'(t''),Y''(t''),x',w')\leq 6L(t''-t')||x'-w'||_d^2+6L(t_k-t_{k-1})g(r).$$

Теперь, возвращаясь к (6.1), для $\widetilde{\mathbb{E}}^{\chi}_{t',x',w'}||Y'(t'')-Y''(t'')||_d^2$ получаем

$$\widetilde{\mathbb{E}}^{\chi}_{t',x',w'}||Y'(t'')-Y''(t'')||_d^2\leq(1+12L(t_k-t_{k-1}))\widetilde{\mathbb{E}}^{\chi}_{t',x',w'}||Y'(t')-Y''(t')||_d^2+12L(t_k-t_{k-1})g(r).$$

Интегрируя полученную оценку по всем $t'=\min(t_{k-1},\tau)$ и $(x',w')=(Y'(t'),Y''(t'))$, имеем

$$\widetilde{\mathbb{E}}^{all}\big[||Y'(t'')-Y''(t'')||_d^2-e^{12L(t_k-t_{k-1})}||Y'(t')-Y''(t')||_d^2\big]\leq 12L(t_k-t_{k-1})g(r);$$

теперь, подставляя $t'=\min(t_{k-1},\tau)$, $t''=\min(t_k,\tau)$, получаем

$$\widetilde{\mathbb{E}}^{all}||Y'(\min(t_k,\tau))-Y''(\min(t_k,\tau))||_d^2$$
$$\leq e^{12L(t_k-t_{k-1})}\widetilde{\mathbb{E}}^{all}||Y'(\min(t_{k-1},\tau))-Y''(\min(t_{k-1},\tau))||_d^2+12L(t_k-t_{k-1})g(r)e^{12L(t_k-t_{k-1})}.$$

Поскольку это верно для всех промежутков $[t_{i-1};t_i)$, дополнительно домножая на $e^{12L(t_k-t_i)}$, получаем также

$$e^{12L(t_k-t_i)}\widetilde{\mathbb{E}}^{all}||Y'(\min(t_i,\tau))-Y''(\min(t_i,\tau))||_d^2$$
$$\leq e^{12L(t_k-t_{i-1})}\widetilde{\mathbb{E}}^{all}||Y'(\min(t_{i-1},\tau))-Y''(\min(t_{i-1},\tau))||_d^2+12L(t_i-t_{i-1})g(r)e^{12L(t_k-t_{i-1})}.$$

Последовательно подставляя эти неравенства друг в друга, имеем

$$\widetilde{\mathbb{E}}^{all}||Y'(\min(t_k,\tau))-Y''(\min(t_k,\tau))||_d^2$$
$$\leq e^{12Lt_k}\widetilde{\mathbb{E}}^{all}||Y'(0)-Y''(0)||_d^2+12Lg(r)\sum_{i=1}^{k}(t_i-t_{i-1})e^{12L(t_k-t_{i-1})};$$

заменяя здесь сумму по нижним прямоугольникам на интеграл, получаем

$$\widetilde{\mathbb{E}}^{all}||Y'(\min(t_k,\tau))-Y''(\min(t_k,\tau))||_d^2\leq e^{12Lt_k}\widetilde{\mathbb{E}}^{all}||Y'(0)-Y''(0)||_d^2+g(r)(e^{12Lt_k}-1)$$

для всех натуральных k. Теперь из $||Y'(0)-Y''(0)||_d^2 \le 7L^2e^{-12LT}\gamma^2$, в силу того, что T принадлежит разбиению, наконец, следует

$$\widetilde{\mathbb{E}}^{all}_{t'}||Y'(\min(T,\tau))-Y''(\min(T,\tau))||_d^2 \le 7L^2\gamma^2+e^{12LT}g(r). \tag{6.2}$$

Напомним, что, как показано в (4.4), имеет место

$$\widetilde{\mathbb{E}}^{all}|T-\tau|^2 \le \widetilde{\mathbb{E}}^{all}\max_{t\in[0;T+2h]}|\pi_0Y'(t)-\pi_0Y''(t)|^2 \le 2\gamma^2. \tag{6.3}$$

Тогда из оценки (3.4) следует

$$\widetilde{\mathbb{E}}^{all}||Y'(T)-Y'(\min(T,\tau))||_d \le 3\sqrt{2}L\gamma. \tag{6.4}$$

Далее, напомним, что при $M^2=hL\sqrt{d+1}+L^2$ супермартингалом является $e^{-t}(||\check{Y}(t)-w'||_d^2+M^2)$ (см. (4.12)). Тогда для моментов остановки $\min(T,\tau)$ и $T+2\gamma$ имеет место неравенство $e^{-T-2\gamma}(1+\widetilde{\mathbb{E}}^{all}||Y''(T+2\gamma)-Y''(\min(T,\tau))||_d^2/M^2) \le \widetilde{\mathbb{E}}^{all}e^{-\min(T,\tau)}$, оценив правую часть которого при помощи

$$\widetilde{\mathbb{E}}^{all}e^{-\min(T,\tau)}=e^{-T}\widetilde{\mathbb{E}}^{all}[1+e^{\max(0,T-\tau)}-1] \le e^{-T}+\widetilde{\mathbb{E}}^{all}\max(0,T-\tau) \overset{(6.3)}{\le} e^{-T}+\sqrt{2}\gamma$$

и установив $H(\gamma)\overset{\triangle}{=}(hL\sqrt{d+1}+L^2)\gamma e^{2\gamma}(2+\sqrt{2}e^T)$, мы получаем

$$e^{-T-2\gamma}(1+\widetilde{\mathbb{E}}^{all}||Y''(T+2\gamma)-Y''(\min(T,\tau))||_d^2/M^2) \le e^{-T}+\sqrt{2}\gamma$$

и

$$\widetilde{\mathbb{E}}^{all}||Y''(T+2\gamma)-Y''(\min(T,\tau))||_d^2 \le M^2(e^{2\gamma}-1+\sqrt{2}e^{T+2\gamma}\gamma) \le H(\gamma).$$

Вместе с (6.2) и (6.4), не без помощи $\sqrt{7}+3\sqrt{2}<7$, это дает

$$\widetilde{\mathbb{E}}^{all}||Y'(T)-Y''(T+2\gamma)||_d \le 7L\gamma+\sqrt{H(\gamma)}+e^{6LT}\sqrt{g(r)}. \tag{6.5}$$

6.2. Расхождение платежей

Обозначим через A событие $(\pi_0Y''(T+2\gamma)>T+\gamma)\,\&\,(\pi_0Y''(T+\gamma)>T)\,\&\,(\bar\sigma(\hat Y)\equiv\bar\sigma(Y'))$; здесь $\hat Y(\cdot)$ — траектория исходной игры. Отметим, что $\hat Y(\cdot)$ может не совпасть с $Y'(\cdot)$ только на $[T;\infty)$ и только из-за прыжка Y' по последней координате. Прыжки невозможны после момента остановки $\tau'=\min\{t\,|\,\pi_0Y''(t)>T+\gamma\}$. Поскольку интенсивности игроков не превосходят L, получаем, что вероятность такого прыжка не превосходит

$$2L(\gamma+\widetilde{\mathbb{E}}^{all}|\tau'-T|)<2L\gamma+2L\widetilde{\mathbb{E}}^{all}\max_{t\in[0;T+2\gamma]}|\pi_0Y''(t)-t| \overset{(6.3)}{\le} 2L(1+\sqrt{2})\gamma,$$

а следовательно, не больше и вероятность события $(\bar\sigma(\hat Y)\ne\bar\sigma(Y'))$. С другой стороны, из (4.5) следует $\widetilde{\mathbb{P}}^{all}((\pi_0Y''(T+2\gamma)>T+\gamma)\,\&\,(\pi_0Y''(T+\gamma)>T))\ge 1-\gamma$. Теперь в силу ограниченности модулей функций W и $\bar\sigma$ единицей из $L\ge 2$ следует

$$\widetilde{\mathbb{P}}^{all}(\text{ не } A)\,\widetilde{\mathbb{E}}^{all}\big(|\bar\sigma(\hat Y)-\bar\sigma(Y'')|\,\big|\text{ не } A\big) \le 2L(3+2\sqrt{2})\gamma \le 6L^2\gamma. \tag{6.6}$$

Рассмотрим такую траекторию, для которой событие A выполнено. Напомним, что $\hat Y(t)=Y'(t)=Y'(t_T)$, $Y''(t)=Y''(t_T)$ при всех $t\ge t_T\overset{\triangle}{=}\sup\{t\,|\,\pi_0Y''(t)<T+\gamma\}\cup\{T\}$. Теперь для траектории процесса в случае выполнения события A имеет место

$$\bar\sigma(Y'')=\int_0^\infty he^{h(T+2\gamma-t)}1_{[0;T+2\gamma)}(t)W(Y''(t))\,dt=W(Y''(T+2\gamma)),$$

$$\bar\sigma(\hat Y)=\int_0^\infty he^{h(T+2\gamma-t)}1_{[0;T+2\gamma)}(t)W(Y'(t))\,dt=W(Y'(T+2\gamma)).$$

Кроме того, поскольку при выполнении события A прыжок по последней координате если происходит, то ровно один раз и одновременно в обоих компонентах на расстояния $1+\pi_0 Y'$ и $1+\pi_0 Y''$ соответственно, то из (6.3) вытекает

$$\widetilde{\mathbb{E}}^{all} \max_{t\in[0;T+2\gamma]} |\pi_{d+1}(Y'(t)-Y''(t))|^2 \le 2\gamma^2,$$

откуда, в частности, получаем $\widetilde{\mathbb{P}}^{all}(A)\widetilde{\mathbb{E}}^{all}\big(|\pi_{d+1}(Y'(t_T)-Y''(t_T))|\,\big|\,A\big) \le \sqrt{2}\gamma \le L\gamma$. Тогда с учетом $3L$-липшицевости отображения $y\mapsto W(y)$ из (6.5) следует

$$\widetilde{\mathbb{P}}^{all}(A)\widetilde{\mathbb{E}}^{all}\big(|\bar\sigma(\hat Y)-\bar\sigma(Y'')|\,\big|\,A\big) \le 3L(8L\gamma+\sqrt{H(\gamma)}+e^{6LT}\sqrt{g(r)}).$$

Это вместе с (6.6) дает

$$\widetilde{\mathbb{E}}^{all}|\bar\sigma(\hat Y)-\bar\sigma(Y'')| \le 3L(10L\gamma+\sqrt{H(\gamma)}+e^{6LT}\sqrt{g(r)}). \tag{6.7}$$

Напомним, что по построению двойной игры Y'' — это процесс, порожденный в марковской игре, в которой первый игрок управляет на каждом промежутке $[t_{k-1};t_k)$ стратегией $\bar\mu^{opt}\equiv \bar\mu^{\downarrow\check{\mathcal{V}}}$. Тогда на каждом таком промежутке $\check{\mathcal{V}}(Y''(t))$ становится супермартингалом, т. е.

$$\check{\mathbb{E}}^{all}(\mathcal{V}(Y''(t_k))|Y''(t_{k-1})) \le \mathcal{V}(Y''(t_{k-1})).$$

Объединяя по всем промежуткам, получаем, что $\check{\mathcal{V}}(z_0)\ge \widetilde{\mathbb{E}}^{all}\bar\sigma(Y'')$, т. е. в силу (6.7)

$$\check{\mathcal{V}}(z_0)\ge \widetilde{\mathbb{E}}^{all}\bar\sigma(\hat Y)-30L^2\gamma-3L\sqrt{H(\gamma)}-3Le^{6LT}\sqrt{g(r)}$$

при любых действиях противника $\bar v$. Поскольку $\widetilde{\mathbb{E}}^{all}\bar\sigma(\hat Y)$ — в точности целевая функция в исходной игре, а игра с поводырем в рамках двойной игры — один из способов играть в исходной игре первому игроку в случае его дискриминации вторым, то показано, что

$$V^+=\hat{\mathcal{V}}^+(x_0)\le \check{\mathcal{V}}(z_0)+30L^2\gamma+3L\sqrt{H(\gamma)}+3Le^{6LT}\sqrt{g(r)}.$$

Аналогично, конструируем при помощи поводыря стратегию для второго игрока с симметричной оценкой. Так как $\gamma=\sqrt{2hT\ln(hT)}$, а шаг решетки h и мелкость разбиения r можно выбрать произвольно малыми, равенство $V^+=V^-$ также показано.

СПИСОК ЛИТЕРАТУРЫ

1. **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. Vol. 9, no. 2. P. 121–144. doi: 10.1007/s10436-012-0210-5 .
2. **Averboukh, Y.** Approximate solutions of continuous-time stochastic games // SIAM J. Control Optim. 2016. Vol. 54, no. 5. P. 2629–2649. doi:10.1137/16M1062247 .
3. **Averboukh, Y.** Approximate Public-Signal Correlated Equilibria For Nonzero-Sum Differential Games // SIAM J. Control Optim. 2019. Vol. 57, no. 1. P. 743–772. doi:10.1137/17M1161403 .
4. **Basu A., Stettner L.** Zero-sum Markov games with impulse controls // SIAM J. Control Optim. 2020. Vol. 58, no. 1. P. 580–604. doi:10.1137/18M1229365 .
5. **Bensoussan, A., Friedman, A.** Nonlinear variational inequalities and differential games with stopping times // J. Functional Analysis. 1974. Vol. 16, no. 3. P. 305–352. doi:10.1016/0022-1236(74)90076-7 .
6. **Bensoussan, A., Friedman, A.** Nonzero-sum stochastic differential games with stopping times and free boundary problems // Trans. Amer. Math. Soc. 1977. Vol. 231, no. 2. P. 275–327. doi:10.1090/S0002-9947-1977-0453082-7 .
7. **Bielecki T. R., Crépey, S., Jeanblanc M., Rutkowski M.** Arbitrage pricing of defaultable game options with applications to convertible bonds // Quantitative Finance. 2008. Vol. 8, no. 8. P. 795–810. doi:10.1080/14697680701401083 .
8. **Биллингсли П.** Сходимость вероятностных мер. М.: Наука, 1977. 352 с.

9. **Боровков А.А.** Теория вероятностей. М.: УРСС, 1999. 470 с.
10. **Дынкин Е.Б.** Игровой вариант задачи об оптимальной остановке // Докл. АН СССР. 1969. Т. 185, вып. 1. С. 16–19.
11. **Ekström E., Peskir G.** Optimal stopping games for Markov processes // SIAM J. Control Optim. 2008. Vol. 47, no. 2. P. 684–702. doi:10.1137/060673916 .
12. **Gensbittel F., Grün C.** Zero-sum stopping games with asymmetric information // Math. Oper. Research. 2019. Vol. 44, no. 1. P. 277–302. doi: 10.1287/moor.2017.0924 .
13. **Gromova E., Malakhova A., Palestini A.** Payoff Distribution in a Multi-Company Extraction Game with Uncertain Duration // Mathematics. 2018. Vol. 6, no. 9, art. no. 165. doi:10.3390/math6090165 .
14. **Guo X, Hernández-Lerma O.** Zero-sum continuous-time Markov games with unbounded transition and discounted payoff rates // Bernoulli. 2005. Vol. 11, no. 6. P. 1009–1029. doi:10.3150/bj/1137421638 .
15. **Hamadéne S.** Mixed zero-sum stochastic differential game and American game options // SIAM J. Control Optim. 2006. Vol. 45, no. 2. P. 496–518. doi:10.1137/S036301290444280X .
16. **Kolokoltsov V.N.** Markov processes, semigroups and generators. Berlin: De Gryuter, 2011. 430 p. (Ser. De Gruyter Studies in Mathematics; vol. 38.) ISBN 978-3-11-025010-7 .
17. **Красовский Н.Н.** Игра сближения-уклонения со стохастическим поводырем // Докл. АН СССР. 1977. Т. 237, вып. 5. С. 1020–1023.
18. **Красовский Н.Н., Котельникова А.Н.** О дифференциальной игре на перехват // Тр. МИАН. 2010. Т. 268. С. 168–214. doi: 10.1134/S008154381001013X .
19. **Красовский Н.Н., Котельникова А.Н.** Стохастический поводырь для объекта с последействием в позиционной дифференциальной игре // Тр. Ин-та математики и механики УрО РАН. 2011. Т. 17, № 2. С. 97–104.
20. **Красовский Н.Н., Субботин А.И.** Позиционные дифференциальные игры. М.: Наука, 1974. 456 с.
21. **Krasovskii N.N., Subbotin A.I.** Game-theoretical control problems. NY: Springer, 1988. 517 p. ISBN: 978-1-4612-8318-8 .
22. **Laraki R., Solan E.** The value of zero-sum stopping games in continuous time // SIAM J. Control Optim. 2005. Vol. 43, no. 5. P. 1913–1922. doi:10.1137/S0363012903429025 .
23. **Marin-Solano J., Shevkoplyas E.** Non-constant discounting and differential games with random time horizon // Automatica. 2011. Vol. 47, no. 12. P. 2626–2638. doi:10.1016/j.automatica.2011.09.010 .
24. **Mazalov V. V.** Dynamic games with optimal stopping. In: Game theory and Applications, vol. 2, L.A. Petrosjan and V.V. Mazalov (eds.), NY: Nova Science Publ., 1996. 223 p. P. 37–46. ISBN 1-56072-390-4 .
25. **Мейер П.-А.** Вероятность и потенциалы. М.: Мир, 1973, 324 с.
26. **Neyman, A.** Continuous-time stochastic games // Games and Economic Behavior. 2017. Vol. 104. P. 92–130. doi: 10.1016/j.geb.2017.02.004 .
27. **Prieto-Rumeau T., Hernández-Lerma O.** Selected topics on continuous-time controlled Markov chains and Markov games. London: Imperial College Press, 2012. 279 p. (ICP Advanced Texts in Math.; vol 5). ISBN-13 978-1-84816-848-0 .
28. **Rockafellar R. T., Wets R. J. B.** Variational analysis. Berlin: Springer-Verlag, 2009. 734 p. (A Series of Comprehensive Studies in Math.; vol. 317). ISBN 978-3-540-62772-2 .
29. **Sorin S., Vigeral G.** Reversibility and oscillations in zero-sum discounted stochastic games // J. Dyn. Games. 2015. Vol. 2, no. 1. P. 103–115. doi:10.3934/jdg.2015.2.103 .
30. **Субботин А.И.** Обобщенные решения уравнений в частных производных первого порядка. Перспективы динамической оптимизации. М.; Ижевск: Институт компьютерных исследований, 2003. 336 с. ISBN 5-93972-206-7 .

Поступила 8.02.2021
После доработки 11.05.2021
Принята к публикации 17.05.2021

Хлопин Дмитрий Валерьевич
канд. физ.-мат. наук
зав. отделом
Институт математики и механики им. Н. Н. Красовского УрО РАН
г. Екатеринбург
e-mail: khlopin@imm.uran.ru

REFERENCES

1. Amir R., Evstigneev I.V., Schenk-Hoppé K.R. Asset market games of survival: a synthesis of evolutionary and dynamic games. *Ann. Finance*, 2013, vol. 9, no. 2, pp. 121–144. doi: 10.1007/s10436-012-0210-5 .
2. Averboukh Y. Approximate solutions of continuous-time stochastic games. *SIAM Journal on Control and Optimization*, 2016, vol. 54, no. 5, pp. 2629–2649. doi: 10.1137/16M1062247 .
3. Averboukh Y. Approximate public-signal correlated equilibria for nonzero-sum differential games. *SIAM J. Control Optim.*, 2019, vol. 57, no. 1, pp. 743–772. doi: 10.1137/17M1161403 .
4. Basu A., Stettner L. Zero-sum Markov games with impulse controls. *SIAM J. Control Optim.*, 2020, vol. 58, no. 1, pp. 580–604. doi: 10.1137/18M1229365 .
5. Bensoussan B., Friedman F. Nonlinear variational inequalities and differential games with stopping times. *J. Functional Analysis*, 1974, vol. 16, no. 3, pp. 305–352. doi: 10.1016/0022-1236(74)90076-7 .
6. Bensoussan A., Friedman A. Nonzero-sum stochastic differential games with stopping times and free boundary problems. *Trans. Amer. Math. Soc.*, 1977, vol. 231, no. 2, pp. 275–327. doi: 10.1090/S0002-9947-1977-0453082-7 .
7. Bielecki T.R., Crépey S., Jeanblanc M., Rutkowski M. Arbitrage pricing of defaultable game options with applications to convertible bonds. *Quantitative Finance*, 2008, vol. 8, no. 8, pp. 795–810. doi: 10.1080/14697680701401083 .
8. Billingsley P. *Convergence of probability measures.* NY: Wiley, 1968, 253 p. ISBN: 0471072427. Translated to Russian under the title *Skhodimost' veroyatnostnykh mer.* Moscow: Nauka Publ., 1977, 352 p.
9. Borovkov A.A. *Teoriya veroyatnostei* [Probability theory]. Moscow: URSS Publ., 1999, 470 p. ISBN: 5-901006-66-6 .
10. Dynkin E.B. Game variant of a problem on optimal stopping. *Soviet Math. Dokl.*, 1969, vol. 10, pp. 270–274.
11. Ekström E., Peskir G. Optimal stopping games for Markov processes. *SIAM J. Control Optim.*, 2008, vol. 47, no. 2, pp. 684–702. doi: 10.1137/060673916 .
12. Gensbittel F., Grün C. Zero-sum stopping games with asymmetric information. *Math. Oper. Research*, 2018, vol. 44, no. 1, pp. 277–302. doi: 10.1287/moor.2017.0924 .
13. Gromova E., Malakhova A., Palestini A. Payoff distribution in a multi-company extraction game with uncertain duration. *Mathematics*, 2018, vol. 6, no. 9, art. no. 165. doi: 10.3390/math6090165 .
14. Guo X., Hernández-Lerma O. Zero-sum continuous-time Markov games with unbounded transition and discounted payoff rates. *Bernoulli*, 2005, vol. 11, no. 6, pp. 1009–1029. doi: 10.3150/bj/1137421638 .
15. Hamadéne S. Mixed zero-sum stochastic differential game and American game options. *SIAM J. Control Optim.*, 2006, vol. 45, no. 2, pp. 496–518. doi: 10.1137/S036301290444280X .
16. Kolokoltsov V.N. *Markov processes, semigroups and generators.* Ser. De Gruyter Studies in Mathematics, vol. 38. Berlin; NY: De Gruyter, 2011, 430 p. doi: 10.1515/9783110250114 .
17. Krasovskii N.N. A convergence-evasion game with stochastic guide. *Dokl. Akad. Nauk SSSR*, 1977, vol. 237, no. 5. pp. 1020–1023 (in Russian).
18. Krasovskii N.N., Kotel'nikova A.N. On a differential interception game. *Proc. Steklov Inst. Math.*, 2010, vol. 268, no. 1, pp. 161–206. doi: 10.1134/S008154381001013X .
19. Krasovskii N.N., Kotel'nikova A.N. Stochastic guide for a time-delay object in a positional differential game. *Proc. Steklov Inst. Math.*, 2012, vol. 277, suppl. 1, pp. 145–151. doi: 10.1134/S0081543812050148 .
20. Krasovskii N.N., Subbotin A.I. *Pozicionnye differencial'nye igry* [Positional differential games]. Moscow: Nauka Publ., 1974, 458 p.
21. Krasovskii N.N., Subbotin A.I. *Game-theoretical control problems.* NY: Springer, 1988, 517 p. ISBN: 978-1-4612-8318-8 .
22. Laraki R., Solan E. The value of zero-sum stopping games in continuous time. *SIAM J. Control Optim.*, 2005, vol. 43, no. 5, pp. 1913–1922. doi: 10.1137/S0363012903429025 .
23. Marin-Solano J., Shevkoplyas E. Non-constant discounting and differential games with random time horizon. *Automatica*, 2011, vol. 47, no. 12, pp. 2626–2638. doi: 10.1016/j.automatica.2011.09.010 .
24. Mazalov V.V. Dynamic games with optimal stopping. In: *Game Theory and Applications*, vol. 2, L.A. Petrosjan and V.V. Mazalov (eds.), NY: Nova Science Publ., 1996, pp. 37–46. ISBN: 1-56072-390-4 .
25. Meyer P.-A. *Probability and potentials.* Massachusetts: Blaisdell Publ. Company, 1966, 266 p. Translated to Russian under the title *Veroyatnost' i potentsialy*, Moscow: Mir Publ., 1973, 324 p.
26. Neyman A. Continuous-time stochastic games. *Games and Economic Behavior*, 2017, vol. 104, pp. 92–130. doi: 10.1016/j.geb.2017.02.004 .

27. Prieto-Rumeau T., Hernández-Lerma O. *Selected topics on continuous-time controlled Markov chains and Markov games.* London: Imperial College Press, 2012, Ser. ICP Advanced Texts in Math., vol 5, 279 p. ISBN: 978-1-84816-848-0 .
28. Rockafellar R.T., Wets R.J.B. *Variational analysis.* Vol. 317. Berlin: Springer–Verlag, 2009, A Series of Comprehensive Studies in Mathematics, 734 p. ISBN: 978-3-540-62772-2 .
29. Sorin S., Vigeral G. Reversibility and oscillations in zero-sum discounted stochastic games. *J. Dyn. Games*, 2015, vol. 2, no. 1, pp. 103–115. doi: 10.3934/jdg.2015.2.103 .
30. Subbotin A.I. *Generalized solutions of first-order PDEs. The dynamical optimization perspective.* Basel: Birkhäuser, 1995, 314 p. doi: 10.1007/978-1-4612-0847-1 . Translated to Russian under the title *Obobshchennye resheniya uravnenii v chastnykh proizvodnykh pervogo poryadka: Perspektivy dinamicheskoi optimizatsii*, Moscow; Izhevsk: Inst. Komp'yuter. Issled. Publ., 2003, 336 p.

Received February 8, 2021
Revised May 11, 2021
Accepted May 17, 2021

Funding Agency: This work was supported by the Russian Science Foundation (project no. 17-11-01093).

Dmitry Valer'evich Khlopin, Cand. Sci. (Phys.-Math.), Krasovskii Institute of Mathematics and Mechanics of the Ural Branch of the Russian Academy of Sciences, Yekaterinburg, 620108 Russia, e-mail: khlopin@imm.uran.ru .

Cite this article as: D. V. Khlopin. Differential game with the possibility of early termination, *Trudy Instituta Matematiki i Mekhaniki UrO RAN*, 2021, vol. 27, no. 4, pp. 189–214 .