Федеральное государственное
бюджетное учреждение науки
Институт математики и механики им. Н.Н. Красовского
Уральского отделения Российской академии наук

на правах рукописи

Авербух Юрий Владимирович

Вероятностные методы анализа игровых задач управления

Резюме диссертации на соискание ученой степени
доктора математических наук

Москва
2020

Введение

Диссертация в форме совокупности статей включает результаты, касающиеся приложения стохастических методов к изучению игровых задач управления. Рассматриваются как игры с конечным, так и игры с бесконечным числом игроков. Отметим, что задачи теории управления возникают в различных областях науки, в том числе, в робототехнике, экономике, финансах и биологии. Кроме того, задачи теории управления и вопросы теории уравнений в частных производных тесно связаны благодаря т.н. принципу динамического программирования.

Исследования, составившие настоящую диссертацию стимулированы позиционным подходом, разработанным в уральской школе по теории управления Н.Н. Красовского. Отличительными чертами разработанной в рамках этой школы методологии является использование разрывных стратегий, многозначного и негладкого анализа, а также теории выживаемости.

Резюме содержит следующие разделы. В разделе 1 дается короткое введение в теорию дифференциальных игр и игр среднего поля. Этот раздел предваряется рассказом об основных результатах теории управления для конечномерных объектов на конечном промежутке времени. Затем мы рассматриваем как антагонистические, так и неантагонистические дифференциальные игры. Также дается необходимые в дальнейшем сведения из теории управляемых систем с динамикой среднего поля и теории игр среднего поля. Отметим, что игры среднего поля представляют собой идеализированную модель игровой задачи управления с бесконечным числом однотипных игроков. В разделе 2 описываются основные результаты диссертации и приводится список публикаций. Раздел 3 содержит подробное описание основных результатов диссертации.

1 Теория управления и дифференциальные игры. Основные сведения

1.1 Теория управления в конечномерном пространстве

Классическим объектом изучения в теории управления является динамическая система на конечном промежутке времени с динамикой, задаваемой обыкновенным дифференциальным уравнением:

$$\frac{d}{dt}x(t) = f(t, x(t), u(t)). \tag{1}$$

Здесь t – время, $x(t)$ – текущее состояние системы, $u(t)$ – мгновенное управление. Мы будем предполагать, что $x(t) \in \mathbb{R}^d$, в то время как значения функции $u(\cdot)$ выбираются из множества U. В рамках теории управления существует множество востребованных в приложениях постановок задач. Мы ограничиваемся лишь двумя:

- минимизация (максимизация) некоторого целевого функционала;
- выживаемость в некотором множестве.

Отметим, что большое число постановок предполагает сочетание задач минимизации/максимизации функционала с теми или иными ограничениями. Также в научной литературе существует большой интерес к задачам быстродействия, которые заключаются в том, что необходимо привести управляемую систему на целевое множество за минимальное время.

Первая задача предполагает, что лицо, принимающее решения, выбирает управление $[0,T] \ni t \mapsto u(t) \in U$ так, чтобы минимизировать величину

$$\gamma(x(T)) + \int_{t_0}^{T} g(t, x(t), u(t))dt. \tag{2}$$

Здесь $x(\cdot)$ удовлетворяет уравнению (1) и начальному условию $x(t_0) = x_0$. Задача оптимального управление может рассматриваться как обобщение теории вариационного исчисления. Однако в фокусе теории оптимального управления находится вопрос о нахождении строгих экстремумов. Необходимое условие оптимальности дается принципом максимума Понтрягина, который сводит исходную (бесконечномерную) задачу оптимизации к краевой задаче для системы обыкновенных дифференциальных уравнений, дополненной поточечным условием максимизации предгамильтониана вдоль оптимальной траектории.

Второй метод нахождения оптимального управления основан на принципе динамического программирования, которой состоит в том, что часть оптимальной траектории сама по себе оптимальна. Чтобы строго описать метод динамического программирования, предположим, что лицо, принимающее решения наблюдает текущее положение системы и использует позиционные стратегии, которые суть функции от времени t и текущего положения $x(t)$. Стратегии, зависящии только от времени принято называть программными. Отметим, что программный и позиционный подходы для задач оптимального управления эквивалентны.

Если t_0 – начальный момент времени, а x_0 – начальное положение системы, то положим

$$\begin{aligned}\mathrm{Val}(t_0, x_0) \triangleq \sup\Big\{&\gamma(x(T)) + \int_{t_0}^{T} g(t, x(t), u(t))dt : \\ &(x(\cdot), u(\cdot)) \text{ удовлетворяющих (1)}, \quad x(t_0) = x_0\Big\}.\end{aligned} \tag{3}$$

Метод динамического программирования влечет, что функция φ удовлетворяет уравнению Беллмана

$$\frac{\partial\varphi}{\partial t} + H(t, x, \nabla\varphi) = 0, \quad \varphi(T, x) = \gamma(x), \tag{4}$$

в минимаксном/вязкостном смысле тогда и только тогда, когда она является функцией цены [Sub95]. Здесь $H : [0,T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ – гамильтониан, задаваемый по правилу:

$$H(t, x, p) = \min_{u \in U}[\langle p, f(t, x, u)\rangle + g(t, x, u)].$$

Отметим, что для рассматриваемой задачи теории управления уравнение Беллмана является уравнением в частных производных первого порядка типа Гамильтона-Якоби.

Кроме того, если $\varphi : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ удовлетворяет уравнению (4), то $\varphi = \mathrm{Val}$ и оптимальная стратегия может быть найдена с использованием градиента функции φ. В частности, в гладком случае оптимальная стратегия определяется равенством

$$u^*(t, x) = \underset{u \in U}{\operatorname{argmin}}[\langle\nabla\varphi(t, x), f(t, x, u)\rangle + g(t, x, u)].$$

Вторая рассматриваемая нами задача – задача о выживаемости. Она состоит в следующем [Aub09], [CLSW98]. Говорят, что множество $K \subset \mathbb{R}^d$ выживает относительно динамики (1) если для всякого начального момента времени t_0 и начальной позиции $x_0 \in K$ можно найти управление $u(\cdot)$ такое, что соответствующая траектория, выходящая из (t_0, x_0), лежит в K. Необходимые и достаточные условия выживаемости формулируются в терминах касательных и нормальных конусов [Aub09], [CLSW98]. Отметим, что задача оптимального управления может быть переформулирована в терминах теории выживаемости. В самом деле, функция цены удовлетворяет следующим условиям:

- ее подграфик выживает относительно расширенной динамической системы:
$$\frac{d}{dt}x(t) = f(t, x(t), u(t)), \quad \frac{d}{dt}z(t) = -g(t, x(t), u(t));$$
- ее надграфик выживает относительно каждой из динамических систем с постоянным управлением $u \in U$:
$$\frac{d}{dt}x(t) = f(t, x(t), u), \quad \frac{d}{dt}z(t) = -g(t, x(t), u).$$

Выше мы рассматривали лишь детерминированные задачи управления. Стохастические задачи управления являются естественным расширением этих постановок. Наиболее общий подход к теории стохастического управления предполагает, что динамика задается генератором типа Леви-Хинчина. Пусть $\mathcal{D}$ подпространство $C(\mathbb{R}^d)$, содержащее $C_c^2(\mathbb{R}^d)$, и пусть оператор $L_t[u] : \mathcal{D} \to C(\mathbb{R}^d)$ задан следующим способом:

$$\begin{aligned} L_t[u]\phi \triangleq \langle f(t, x, u), \nabla\rangle\phi(x) + \frac{1}{2}\langle G(t, x, u)\nabla, \nabla\rangle\phi(x) \\ + \int_{\mathbb{R}^d} [\phi(x+y) - \phi(x) - \langle y, \nabla\phi(x)\rangle \mathbf{1}_{B_1}(y)]\nu(t, x, u, dy) \end{aligned}$$

где G – симметрическая неотрицательная матрица, B_a обозначает шар с центром в начале координат и радиусом $a > 0$, а ν является мерой Леви. В рамках задач стохастического управления предполагается, что функционал платы равен

$$\mathbb{E}\left[\gamma(x(T)) + \int_{t_0}^{T} g(t, x(t), u(t))dt\right]. \tag{5}$$

Определение 1.1. Мы говорим (следуя [FS06]), что набор $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[s,T]}, P, u, X)$ является управляемым процессом на $[s, T]$ допустимым для генератора L, если

- $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[s,T]}, P)$ является вероятностным пространством с фильтрацией;
- u – $\{\mathcal{F}_t\}_{t\in[s,T]}$-прогрессивно измеримый процесс со значениями в U; X – càdlàg процесс, принимающий значения в $\mathbb{R}^d$, согласованный с фильтрацией $\{\mathcal{F}_t\}_{t\in[s,T]}$;
- для всех $\phi \in \mathcal{D}$ процесс
$$\phi(X(t)) - \int_s^t L_\tau^n[u(\tau)]\phi(X(\tau))d\tau$$
является $\{\mathcal{F}_t\}_{t\in[s,T]}$-мартингалом.

Уравнение Беллмана для систем со стохастической динамикой принимает следующий вид:

$$\frac{\partial\varphi}{\partial t} + \max_{u\in U}\Big[(L_t[u]\varphi)(x) + g(t,x,u)\Big] = 0, \quad \varphi(T,x) = \gamma(x). \tag{6}$$

Отметим, что этот максимально общий подход до сих пор изучен достаточно слабо. Насколько мне известно, метод динамического программирования развит для общих стохастических управляемых систем лишь при дополнительном предположении, что уравнение Беллмана (6) имеет гладкое решение изучен [FS06].

Исчерпывающие результаты получены в двух принципиально важных частных случаях: для управляемых систем с динамикой, задаваемой стохастическим дифференциальным уравнением [YZ99] и в случае динамики, задаваемой марковской цепью с непрерывным временем [GHL03]. В первом случае предполагается, управляемый процесс подчиняется стохастическому дифференциальному уравнению

$$dX = f(t,X(t),u(t))dt + \sigma(t,X(t),u(t))dW_t.$$

Это предположение соответствует генератору

$$L_t[u]\phi \triangleq \langle f(t,x,u),\nabla\rangle\phi(x) + \frac{1}{2}\langle G(t,x,u)\nabla,\nabla\rangle\phi(x)$$

с $G(t,x,u) = \sigma(t,x,u)\sigma^T(t,x,u)$.

Случай марковской динамики предполагает, что задана управляемая марковская цепь с матрицей Колмогорова $(Q_{x,y}(t,u))_{x,y\in\mathcal{S}}$, где $\mathcal{S}$ – некоторое конечное множество. Марковские задачи управления являются задачами теории управления для стохастических систем с генератором типа Леви-Хинчина равным:

$$L_t[u]\phi(x) \triangleq \sum_{y\in\mathcal{S},y\neq x}[\phi(y)-\phi(x)]Q_{x,y}(t,u).$$

Отметим, что для задач стохастического управления, задаваемая СДУ, приводит к уравнению Беллмана, являющемуся уравнением в частных производных второго порядка параболического типа. В то время как, уравнение Беллмана для марковской управляемой системы является системой обыкновенных дифференциальных уравнений. Таким образом, упомянутые стохастические задачи управления оказываются с точки зрения метода динамического программирования проще детерминированного случая.

1.2 Антагонистические дифференциальные игры в конечномерном пространстве

Антагонистические дифференциальные игры можно рассматривать как задачи теории управления с двумя лицами, принимающими решения (называемым игроками) в предположении о том, что интересы игроков противоположны. Мы рассматриваем дифференциальные игры на конечном промежутке времени, предлагая, что игроки стремятся минимизировать/максимизировать некоторый функционал[1]. Таким образом, в фокусе нашего внимания находится управляемая система с динамикой

$$\frac{d}{dt}x(t) = f(t,x(t),u(t),v(t)), \quad t\in[0,T], \quad x(t)\in\mathbb{R}^d, \quad u(t)\in U, \quad v(t)\in V. \tag{7}$$

[1]Другие постановки теории дифференциальных игр могут быть найдены, к примеру в [KS88].

Мы предполагаем, что первый игрок, управляя переменой $u(t)$, стремится минимизировать значение функционала

$$\gamma(x(T)) + \int_{t_0}^{T} g(t, x(t), u(t), v(t))dt.$$

В то время, как второй игрок управляет переменой $v(t)$ и имеет противоположную цель.

Основополагающую роль в построении теории дифференциальных игр играет выбор класса допустимых стратегий.

Прежде всего можно предположить, что управления зависят лишь от времени т.е. мы рассматриваем программные стратегии. По сути мы предполагаем, что игрок слеп и не обладает информацией о текущей позиции. Этот класс стратегий является весьма ограничительным. Предполагая, что игроки могут наблюдать текущее положение мы приходим к концепции позиционные стратегии $u(t, x)$ и $v(t, x)$.

Заданную позиционную стратегию $u(t, x)$ можно напрямую подставить в динамику (7). Это требует либо предположения о непрерывности стратегий u по x, либо перехода к дифференциальным включениям. Отметим, что Н.Н. Барабанова (Субботина) и А.И. Субботин показали, что как бы мы не подставляли позиционную стратегию напрямую, этот подход не является оптимальным по сравнению с пошаговыми схемами [BS71]. Пошаговая схема, впервые предложенная Н.Н. Красовским и А.И. Субботиным [KS70] (см. также [KS88]), предполагает, что задано разбиение отрезка управления $\Delta = \{t_i\}$ и на каждом интервале $[t_i, t_{i+1}]$ используется постоянное управление $u(t_i, x(t_i))$.

При этом результат первого игрока оценивается следующим образом. Для позиционно стратегии u, разбиения отрезка управления $[t_0, T]$ $\Delta = \{t_i\}_{i=0}^{n}$, и некоторого управления второго игрока $v(\cdot)$, вычислим величину

$$J(t_0, x_0, u, \Delta, v) = \gamma(x(T)) + \int_{t_0}^{T} g(t, x(t), u(t), v(t))dt,$$

где $u(t)$ – пошаговая реализация стратегии $u(t, x)$,

$$u(t) = u(t_i, x(t_i)) \text{ для } t \in [t_i, t_{i+1}),$$

$x(\cdot)$ обозначает порожденную траекторию. Напомним, что первый игрок, корректируя свое управление как можно чаще, стремится минимизировать величину J для всех действий второго игрока. Таким образом точная верхняя оценка результата первого игрока есть

$$\mathrm{Val}^{+}(t_0, x_0) = \inf_{u} \limsup_{\delta \downarrow 0} \sup\{J(t_0, x_0, u, \Delta, v) : d(\Delta) \leq \delta\}.$$

Здесь супремум берется по всем возможным реализациям управления второго игрока v.

Переставляя игроков мы получаем оценку результата второго игрока функцией $\mathrm{Val}^{-}(t_0, x_0)$.

Очевидно, что

$$\mathrm{Val}^{+}(t_0, x_0) \geq \mathrm{Val}^{-}(t_0, x_0).$$

Известно [Sub95], что функции цены удовлетворяют уравнениям Беллмана

$$\frac{\partial\varphi}{\partial t} + H(t,x,\nabla\varphi) = 0, \quad \varphi(T,x) = \gamma(x)$$

с гамильтонианами

$$H^+(t,x,p) \triangleq \min_{u\in U}\max_{v\in V}[\langle p, f(t,x,u,v)\rangle + g(t,x,u,v)]$$

для случая верхней функции цены, и

$$H^-(t,x,p) \triangleq \max_{v\in V}\min_{u\in U}[\langle p, f(t,x,u,v)\rangle + g(t,x,u,v)]$$

для случая нижней функции цены.

Если $H^+(t,x,p) = H^-(t,x,p)$ (это условие называется условием Айзекса), мы имеем, что (см. [BD96], [KS88], [Sub95]) игра имеет цену, которую мы обозначим через Val,

$$\mathrm{Val}^+ = \mathrm{Val}^- = \mathrm{Val}.$$

Отметим, что верхняя и нижняя функции цены могут быть охарактеризованы с использование подхода теории выживаемости [KS88], [Sub95].

Функция $\varphi : [0,T]\times\mathbb{R}^d \to \mathbb{R}$ называется u-стабильной, если, для всех $s,r \in [0,T]$, $s<r$, $y\in\mathbb{R}^d$ и постоянного управления второго игрока $v\in V$, можно найти $(x(\cdot), z(\cdot))$, удовлетворяющие условиям

$$(\dot{x}(t), \dot{z}(t)) \in \mathrm{co}\{(f(t,x,u,v), g(t,x,u,v)) : u\in U\}, \quad x(s) = y, \quad z(s) = 0$$

и

$$\varphi(s,y) \geq \varphi(r, x(r)) + z(r).$$

Верхняя функция цены является u-стабильной. Также, каждая u-стабильная функция задает оценку верхней функции цены сверху.

Условие v-стабильности определяется аналогично. Функция $\varphi : [0,T]\times\mathbb{R}^d \to \mathbb{R}$ называется v-стабильной, если для всех $s,r\in[0,T]$, $s<r$, $y\in\mathbb{R}^d$ и всякого постоянного управления первого игрока $u\in U$ существует решение задачи Коши для дифференциального включения $(x(\cdot), z(\cdot))$

$$(\dot{x}(t), \dot{z}(t)) \in \mathrm{co}\{(f(t,x,u,v), g(t,x,u,v)) : v\in V\}, \quad x(s) = y, \quad z(s) = 0$$

такое, что

$$\varphi(s,y) \leq \varphi(r, x(r)) + z(r).$$

Как и выше, всякая v-стабильная функция является нижней оценкой нижней функции цены, которая в свою очередь сама v-стабильна.

По заданной u-стабильной функции φ, компакту начальных позиций $K\subset\mathbb{R}^d$ и положительному числу ε можно построить стратегию первого игрока $u^{\varepsilon}_{\varphi,K}(t,x)$, гарантирующую результат не более, чем $\varphi(s,y)+\varepsilon$, в каждой начальной позиции (s,y) из компакта K. Если мы выберем φ равной функции цены, то получим ε-оптимальную стратегию для компакта K. Такие же построения могут быть проделаны для второго игрока с использованием v-стабильных функций.

Существуют и другие подходы к формализации дифференциальных игр. Среди них необходимо отметить квазистратегии и стратегии управления с поводырем (моделью). Квазистратегией первого игрока (случай второго игрока рассматривается аналогично) называется отображение α, ставящие в соответствие каждому измеримому управлению второго игрока управление первого игрока такое, что равенство $v_1(t) = v_2(t)$ на $[0, \tau]$ влечет равенство $\alpha[v_1](t) = \alpha[v_2](t)$ для всех $t \in [0, \tau]$. Отметим, что позиционная формализация дифференциальных игр и формализация в классе квазистратегий эквивалентны [SC81].

Стратегии с поводырем представляют собой частный случай стратегий с памятью. Они были предложены Н.Н. Красовским и А.И. Субботиным для того, чтобы преодолеть неустойчивость позиционных стратегий к информационным помехам. Основная идея этого подхода состоит в том, что управление формируется пошагово с использованием информации о положении вспомогательной системы (поводыря). Отметим, что функции цены в классе позиционных стратегий и стратегий с поводырем совпадают.

Выше было отмечено, что, используя метод динамического программирования, мы можем свести игровую задачу управления к уравнению в частных производных первого порядка. При этом функция цены является минимаксным/вязкостным решением этого уравнения. Отметим, что в случае детерминированной динамики понятие минимаксного/вязкостного решения вводится достаточно непросто [Sub95]. Более того, обыкновенно это решение является негладким. Поэтому для построения приближенно оптимальных стратегий используется такой аппарат негладкого анализа как квазиградиенты и проксимальные градиенты.

Отметим, что эти конструкции упрощаются, если рассмотреть вместо детерминированной динамики (7) либо динамику, задаваемую стохастическим дифференциальным уравнением

$$dX(t) = f(t, X(t), u(t), v(t))dt + \sigma(t, X(t))dW_t,$$

либо динамику, задаваемую управляемой марковской цепью с матрицей Колмогорова

$$(Q_{x,y}(t, u, v))_{x,y \in \mathcal{S}},$$

где $\mathcal{S}$ – пространство возможных состояний. В первом случае уравнение Беллмана – это уравнение в частных производных второго порядка параболического типа, при этом оптимальные стратегии могут быть построены с использованием производных этого решения [HL95]. Во втором случае уравнение Беллмана сводится к системе обыкновенных дифференциальных уравнений, а оптимальные стратегии определяются напрямую решением этой системы [Zac64].

Н.Н. Красовский и А.Н. Котельникова предложили использовать решения стохастической дифференциальной игры для построения приближенно оптимальных стратегий в исходной дифференциальной игре [KK10]. Этот подход может быть обобщен на случай игры, задаваемой генератором более общего вида. Отметим, что диссертация включает ряд подобных результатов.

1.3 Неантагонистические дифференциальные игры

Неантагонистическая дифференциальная игра конечного числа лиц в конечномерном пространстве на ограниченном промежутке времени задается следующим образом. Предполагается, что динамика система определяется дифференциальным уравнением

$$\begin{gathered}\frac{d}{dt}x(t) = f(t, x(t), u_1(t), \ldots, u_N(t)), \\ t \in [0, T], \quad x(t) \in \mathbb{R}^d, \quad u_i(t) \in U_i, \quad i = 1, \ldots, N.\end{gathered} \tag{8}$$

Где i-й игрок управляет переменой $u_i(t)$ и стремится максимизировать выигрыш, равный

$$\gamma_i(x(T)) + \int_{t_0}^{T} g_i(t, x(t), u_1(t), \ldots, u_N(t))dt.$$

Отметим, что антагонистическая дифференциальная игра возникает, если мы положим $N = 2$, $\gamma_1 = -\gamma_2$, $g_1 = -g_2$.

Существует несколько различных подходов к определению решения неантагонистической дифференциальной игры. Мы ограничимся лишь концепцией равновесия по Нэшу, которое состоит в том, что ищется набор стратегий такой, что любое индивидуальное отклонение не выгодно.

Равновесие по Нэшу в некоторых случаях может быть изучено с помощью метода динамического программирования, который сводит нашу игровую задачу к системе уравнений Беллмана. Для того, чтобы описать этот метод предположим, что корректно определены функции:

$$\begin{aligned}H_i(t, x,&p_1, \ldots, p_N) \triangleq \\ &\langle p_i, f(t, x, u_1^*(t, x, p_1, \ldots, p_N), \ldots, u_N^*(t, x, p_1, \ldots, p_N))\rangle \\ &\quad + g_i(t, x, u_1^*(t, x, p_1, \ldots, p_N), \ldots, u_N^*(t, x, p_1, \ldots, p_N)),\end{aligned} \tag{9}$$

где набор стратегий $(u_1^*(t, x, p_1, \ldots, p_N), \ldots, u_N^*(t, x, p_1, \ldots, p_N))$ удовлетворяет следующему свойству:

$$\begin{aligned}&\langle p_i, f(t, x, u_1^*(t, x, p_1, \ldots, p_N), \ldots, u_N^*(t, x, p_1, \ldots, p_N))\rangle \\ &\quad + g_i(t, x, u_1^*(t,x, p_1, \ldots, p_N), \ldots, u_N^*(t, x, p_1, \ldots, p_N)) = \\ &\qquad \max_{u_i \in U_i}[\langle p_i, f(t, x, u_1^*(t, x, p_1, \ldots, p_N), \ldots, u_i, \ldots, u_N^*(t, x, p_1, \ldots, p_N))\rangle \\ &\qquad\quad + g_i(t, x, u_1^*(t, x, p_1, \ldots, p_N), \ldots, u_i, \ldots, u_N^*(t, x, p_1, \ldots, p_N))]\end{aligned} \tag{10}$$

В этом случае показано [Fri70], что, если существует гладкое решение системы уравнений Беллмана

$$\frac{\partial \varphi_i}{\partial t} + H_i(t, x, \nabla\varphi_1, \ldots, \varphi_N) = 0, \quad \varphi_i(T, x) = \gamma_i(x), \tag{11}$$

то набор стратегий $(u_i^*(t, x, \nabla\varphi_1(t, x), \ldots, \nabla\varphi_N(t, x)))_{i=1,\ldots,N}$ задает равновесие по Нэшу в классе позиционных стратегий в каждой начальной позиции.

Поскольку нам известно, что даже в антагонистическом случае у уравнения Беллмана есть только негладкое решение, то разумным является попытаться построить теорию негладких решений для системы уравнений Беллмана. Это было сделано для ряда частных случаев в работах [BS04a], [BS04b], [CP03].

Однако общей теории обобщенных решений системы (9), (10) до настоящего времени не построено. Более того есть ряд примеров, показывающих, что в общем случае эта система допускает лишь разрывные решения [Ave15].

Второй подход к исследованию неантагонистических игр состоит в использовании стратегий наказания. Этот подход предполагает, что игроки выбирают некоторую траекторию, реализуемую в силу динамики (8), и наказывают за все индивидуальные отклонения от этой траектории. Для того, чтобы траектория обеспечивала равновесие по Нэшу необходимо и достаточно, чтобы ожидаемый выигрыш каждого игрока вдоль траектории был не меньше значения функции цены антагонистической игры, в которой этот игрок стремится максимизировать свой выигрыш, а остальные игроки ему противодействуют. Стратегии наказания позволяют построить равновесие по Нэшу в классе позиционных стратегий для случая двух игроков [Kle93]. В играх N игроков с $N > 2$ позиционное равновесие может быть построено лишь для случая, когда i-й игрок воздействует напрямую лишь на свою переменую x_i. В общем случае требуется некоторый внешний центр, распространяющий среди игроков информацию об игроке-уклонисте [Kle93].

Отметим основные недостатки стратегий наказаний.

1. Игроки должны договариваться о выборе траектории.

2. Имеется множественность равновесий.

3. Конструкция равновесия по Нэшу опирается на угрозы, которые зачастую сложно назать адекватными.

4. Построенное в рамках стратегий наказания позиционное равновесие не является универсальным т.е. нельзя построить один общий набор стратегий, подходящий во всех начальных позициях из заданного множества.

Метод динамического программирования проще реализуется в стохастическом случае: для неантагонистической игры с динамкой задаваемой стохастическим дифференциальным уравнением, система уравнений Беллмана (11) становится системой параболических уравнений второго порядка, а для неантагонистической марковской игры – системой обыкновенных дифференциальных уравнений [Lev13, Theorems 2 and 5]. Для этих случаев теоремы существования решения и конструкции стратегий могут быть получены при достаточно широких предположениях [Man04], [Man14], [HM19], [Lev13, Theorem 6].

1.4 Управляемые системы с динамикой среднего поля и игры среднего поля

Игры среднего поля (предложенные в работах Lasry и Lions [LL06a], [LL06b] а также, независимо Caines, Malhamé, Huang [HMC05]) и задачи управления с динамикой среднего поля [AD01] исследуют системы, состоящие из бесконечно большого числа однотипных элементов таких, что динамика каждого элемента определяется его положением, его действиями и распределением всех прочих частиц. Это приводит к динамической системе в пространстве вероятностных мер, которое оснащается метрикой Канторовича. Кратко опишем этот объект.

Прежде всего введем некоторые дополнительные обозначения. Если (X, ρ_X) – польское пространство, то множество всех борелевских вероятностей на X через $\mathcal{P}(X)$. Если $(\Omega', \mathcal{F}')$ и $(\Omega'', \mathcal{F}'')$ – некоторые измеримые пространства, m – вероятность на $\mathcal{F}'$, $h : \Omega' \to \Omega''$ – измеримое отображение, то $h_{\#}m$ обозначает образ меры

m в силу отображения h, определяемый следующим образом: для $\Upsilon \in \mathcal{F}''$

$$(h_{\#}m)(\Upsilon) \triangleq m(h^{-1}(\Upsilon)).$$

Также нам потребуются операторы проектирования. Если X_1, X_2 – некоторые множества, то обозначим естественную проекцию $X_1 \times X_2$ на X_i через p^i.

При $p \geq 1$, $\mathcal{P}^p(X)$ обозначает множество вероятностей m на X таких, что для какого-то (а значит, и для всех) $x_0 \in X$

$$\int_X \rho_X^p(x, x_0) m(dx) < \infty.$$

Метрика Канторовича[2] с показателем p определяется следующим образом: если $m', m'' \in \mathcal{P}(X)$, то

$$W_p(m', m'') \triangleq \left[\inf\left\{\int_{X\times X} \rho_X^p(x', x'')\pi(d(x', x'')) : \pi \in \Pi(m', m'')\right\}\right]^{1/p},$$

где $\Pi(m', m'')$ – множество планов между m' и m'' т.е. $\pi \in \Pi(m', m'')$, если $\pi \in \mathcal{P}^1(X \times X)$ и $\mathrm{p}^1{}_{\#}\pi = m'$, $\mathrm{p}^2{}_{\#}\pi = m''$.

Пространство $\mathcal{P}^p(X)$, снабженное метрикой W_p, является польским, если пространство X само является польским [AGS05]. Если X – компакт, то $\mathcal{P}^p(X)$ также компакт [AGS05]. Однако свойство σ-компактности $\mathcal{P}^p(X)$ не следует из того, что X – σ-компакт.

Управляемый процесс с динамикой типа среднего поля может быть определен с использованием генераторной техники. В этом случае мы предполагаем, что динамика каждого элемента системы задается генератором типа Леви-Хинчина

$$\begin{aligned} L_t[m,u]\phi \triangleq \langle f(t,x,m,u), \nabla\rangle\phi(x) + \frac{1}{2}\langle G(t,x,m,u)\nabla, \nabla\rangle\phi(x) \\ + \int_{\mathbb{R}^d}[\phi(x+y) - \phi(x) - \langle y, \nabla\phi(x)\rangle \mathbf{1}_{B_1}(y)]\nu(t,x,m,u,dy). \end{aligned}$$

Здесь m обозначает вероятность, соответствующую распределению всех элементов системы. Управление u может быть выбрано либо программным (т.е. как случайный процесс со значениями в U), либо позиционным $u(t,x)$.

В первом случае мы приходим к следующему определению.

Определение 1.2. Будем говорить, что поток вероятностей $[s,T] \ni t \mapsto m(t) \in \mathcal{P}^p(\mathbb{R}^d)$ порожден генератором L, если существует управляемый процесс $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t\in[s,T]}, P, u, X)$, допустимый для генератора $L_t[m(t), u]$, такой, что

$$m(t) = X(t,\cdot)_{\#}P.$$

В этом случае движение каждого элемента системы определяется генератором $L_t[m(t), u]$.

[2]В англоязычной литературе распространено название метрика Васерштейна (Wasserstein distance). Вопрос о точном наименовании разъяснен в [BK12, §1.1].

Отметим, что, если мы рассматриваем позиционные стратегии $u(t,x)$ или $u(t,x,m)$, то можно положить $u(t,\omega) = u(t, X(t,\omega))$ или $u(t,\omega) = u(t, X(t,\omega), m(t))$.

Предложенная конструкция включает в себя управляемую динамику для уравнений типа МакКина-Власова, когда динамика каждого элемента определяется уравнением

$$dX(t) = f(t, X(t), m(t), u)dt + \sigma(t, X(t), m(t), u)dW_t, \tag{12}$$

где $m(t)$ – распределение процесса $X(t)$ и более общие процессы, задаваемые скачкообразными процессами с диффузией. В диссертации преимущественно рассматриваются детерминированные управляемые процессы с динамикой типа среднего поля, которые получаются из уравнения (12) при $\sigma = 0$. Отметим, что в этом случае полезно выбрать и зафиксировать управляемое пространство.

Как было отмечено выше, и теория игр среднего поля, и теория управления управляемыми системами с динамикой типа среднего поля рассматривают управляемые динамические системы в пространстве вероятностей. Различие между ними следующее. В рамках теории игр среднего поля предполагается, что каждый элемент системы является самостоятельным игроком и стремится максимизировать выигрыш, равный

$$\mathbb{E}\left[\gamma(X(T), m(T)) + \int_s^T g(t, X(t), m(t), u(t))dt\right], \tag{13}$$

в то время как в рамках теории управления системами с динамикой типа среднего поля предполагается, что целью управления всеми частями системы является максимизация общего выигрыша равного

$$\gamma'(m(T)) + \mathbb{E}\int_s^T g(t, X(t), m(t), u(t))dt. \tag{14}$$

Теория игр среднего поля исследует равновесия в играх с бесконечным числом однотипных участников. Это приводит нас к следующему определению.

Определение 1.3. Пусть $m_0 \in \mathcal{P}^2(\mathbb{R}^d)$ – заданное начальное распределение игроков. Будем говорить, что набор $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t\in[t_0,T]}, P, u, X, m(\cdot))$ задает решение игры среднего поля для динамики, задаваемой генератором L, и функционалом выигрыша каждого игрока (13), если

- $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t\in[t_0,T]}, P, u, X)$ – управляемый процесс, допустимый для генератора $L_t[m(t), u]$;
- $m(t) = X(t,\cdot)_{\#}P$;
- $m(t_0) = m_0$;
- для всех $y \in \mathbb{R}^d$ и любого управляемого процесса $(\Omega', \mathcal{F}', \{\mathcal{F}'\}_{t\in[t_0,T]}, P', u', X')$, допустимого для $L_t[m(t), u]$, такого, что $X(t_0) = y$ P'-п.н. выполнено неравенство

$$\begin{aligned}\mathbb{E}\Big[\gamma(X(T), m(T)) + \int_{t_0}^T g(t, X(t), m(t), u(t))dt\Big| X(t_0) = y\Big] \\ \geq \mathbb{E}'\Big[\gamma(X'(T), m(T)) + \int_{t_0}^T g(t, X'(t), m(t), u'(t))dt\Big]\end{aligned}$$

Здесь $\mathbb{E}$ (соответственно, $\mathbb{E}'$) обозначает математическое ожидание в силу вероятности P (соответственно, P'.)

Определение 1.3 лежит в русле вероятностного подхода к играм среднего поля [CD18]. Другой подход (более популярный в существующей литературе) основан на системе дифференциальных уравнений в частных производных. Чтобы описать этот подход, предположим, что $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, P, u, X)$ задает равновесие в игре среднего поля. Далее положим

$$\varphi(s,y) \triangleq \mathbb{E}\Big[\gamma(X(T), m(T)) + \int_s^T g(t, X(t), m(t), u(t))dt \Big| X(s) = y\Big].$$

Функция φ определяет ожидаемый выигрыш пробного игрока, который стартует в момент t из позиции y. Метод динамического программирования позволяет заключить, что это функция удовлетворяет (возможно в обобщенном смысле) уравнению Беллмана

$$\frac{\partial\varphi}{\partial t} + \max_{u\in U}\Big(L_t[m(t), u]\varphi(x) + g(t, x, m(t), u)\Big) = 0, \quad \varphi(T, x) = \gamma(x, m(T)). \tag{15}$$

Одновременно, динамика потока вероятностей $m(\cdot)$ определяется равенством $m(t) = X(t,\cdot)_{\#}P$ со стохастическим процессом X, определяемым оптимальным управляемым процессом. Если мы выберем позиционное управление по правилу

$$u^*(t,x) = \underset{u\in U}{\operatorname{argmax}}\Big(L_t[m(t), u]\varphi(x) + g(t, x, m(t), u)\Big), \tag{16}$$

то поток вероятностей $m(\cdot)$ удовлетворяет следующему кинетическому уравнению

$$\frac{d}{dt}m(t) = L_t[m(t), u^*(t,\cdot)]m(t), \quad m(t_0) = m_0. \tag{17}$$

Систему уравнений (15), (17) вместе с условием (16) называют системой уравнений для игры среднего поля. Ее решение задает процесс, определяющий решение игры среднего поля в силу Определения 1.3. Теоремы существования решений системы уравнений для игры среднего поля доказываются при весьма широких предположениях на генераторы и интегральный показатель качества управления g [KLY11], [KY13].

Игры среднего поля рассматриваются как предельные игровые задачи управления для игровых задач управления с большим числом игроков. Это предположение становится строгим утверждением для случая равновесий в классе программных стратегий [Fis17], [Lac17]. Предельное поведение позиционных равновесий рассматривалось с помощью т.н. основного уравнения для игр среднего поля в [CDLL19]. Основное уравнение игр среднего поля – это дифференциальное уравнение в произведении конечномерного пространства и пространства вероятностных мер. Теорема о существовании и единственности решения этого уравнения была получена лишь в случае динамики, задаваемой стохастическим дифференциальным уравнением с невырожденным коэффициентом стохастичности, и коэрцитивного гамильтониана. Кроме того, существует ряд результатов о существании решения на малых промежутках времени [GS15]. Также на основе решения игры среднего поля могут быть построены приближенные равновесия по Нэшу для игр с конечным числом игроков [KLY11], [KTY14].

2 Результаты и публикации

В рамках диссертации получены следующие результаты, касающиеся применения вероятностных методов при анализе дифференциальных игр. Прежде всего были рассмотрены игры двух лиц (как антагонистические так и неантагонистические). В этой части исследования основным объектом являются конструкции приближенных решений и оценки функции цены, основанные на решении некоторой модельной игровой задачи управления. Вторая часть результатов посвящена исследованию систем большого числа однотипных частиц в пределе когда, число элементов стремится к бесконечности. Здесь, прежде всего, изучались игры среднего поля, предложенные в работах Lasry, Lions и (независимо) Huang, Malhamé, Caines. В этом направлении нами рассматривались вопросы существования решений, их устойчивости. Также для исследования систем, состоящих из большого числа частиц, были применены методы теории выживаемости.

Кратко опишем основные результаты диссертации.

1. Рассматривались приближенные решения антагонистической стохастической игры с непрерывным временем на основе решений модельной игры. Предполагалось, что динамики исходной и модельной игре, вообще говоря, различны. Построены стратегии игроков в исходной игре, приближенно реализующие значения функций, удовлетворяющих условиям стабильности в модельной игре.

2. На основе общего результата, была получена аппроксимация функции цены антагонистической дифференциальной игры на основе решения задачи Коши для параболического уравнения и на основе решения системы обыкновенных дифференциальных уравнений. Также с использованием этого подхода были построены приближенно оптимальные стратегии и оценки функции цены в игровой задаче управления марковской цепью, описывающей систему многих взаимодействующих частиц с конечным числом состояний.

3. Для неантагонистических дифференциальных игр двух лиц было построено приближенное равновесие по Нэшу в классе стохастических стратегий с памятью и общим сигналом. Предложенная конструкция использует пару функций, удовлетворяющую условию стабильности в модельной стохастической игре с непрерывным временем. Доказано, что если модельная игра сходится к исходной, то множество предельных выигрышей игроков лежит в выпуклой оболочке множества выигрышей игроков, соответствующих равновесию по Нэшу в классе позиционных стратегий.

4. Были найдены конкретные классы пар функций, удовлетворяющих условиям стабильности. В частности, показано, что приближенное равновесие можно построить по паре функций, являющейся строгим решением системы параболических уравнений и по решению системы дифференциальных включений, получающейся как система уравнений Беллмана для марковской игры.

5. Было исследовано условие выживаемости для системы бесконечного числа однотипных элементов. Была получена теорема о выживаемости типа Нагумо, которая опирается на аналог касательного конуса для подмножества пространства вероятностных мер.

6. Был предложен минимаксный подход к детерминированным играм среднего поля. Этот подход является вариантом вероятностного подхода, в котором решение задается распределениями траекторий в расширенном фазовом пространстве. Были исследованы вопросы существования решения, его устойчивости по отношению к стохастическим возмущениям динамики. Кроме того, по заданному решению игры среднего поля были построены приближенные равновесия по Нэшу в игре конечного числа игроков.

7. Исследовалась зависимость решения игры среднего поля от начального распределения игроков. Для этого было введено понятие мультифункции цены – отображения, которое ставит в соответствие начальному моменту времени и начальному распределению игроков множество ожидаемых выигрышей пробного игрока. Было показано, что если некоторое многозначное отображение является выживающим относительно динамики, задаваемой игрой среднего поля, то это отображение – мультифункция цены. Были найдены инфинитезимальные варианты этого условия выживаемости. Это условие может рассматриваться как обобщенный вариант основного уравнения для игр среднего поля.

Список публикаций, выносимых на защиту

[1] Yu. Averboukh. Approximate solutions of continuous-time stochastic games. *SIAM J. Control Optim.*, 54(5):2629–2649, 2016.

[2] Yu. Averboukh. Extremal shift rule for continuous-time zero-sum Markov games. *Dyn. Games Appl.*, 7(1):1–20, 2017.

[3] Yu. Averboukh. Approximate public-signal correlated equilibria for nonzero-sum differential games. *SIAM J. Control Optim.*, 57(1):743–772, 2019.

[4] Yu. Averboukh. Markov approximations of nonzero-sum differential games. *Vestnik Udmurtskogo Universiteta. Matematika. Mekhanika. Komp'yuternye Nauki*, 30:3–17, 2020.

[5] Yu. Averboukh. Viability theorem for deterministic mean field type control systems. *Set-Valued Var. Anal.*, 26:993–1008, 2018.

[6] Yu. Averboukh. A minimax approach to mean field games. *Sb. Math.*, 206(7):893–920, 2015.

[7] Yu. Averboukh. Deterministic limit of mean field games associated with nonlinear Markov processes. *Appl. Math. Opt.*, 81:711–738, 2020.

[8] Yu. Averboukh. Viability analysis of the first-order mean field games. *ESAIM Contr. Optim. Ca.*, 26:33, 35 pages, 2020.

3 Краткое описание основных результатов

3.1 Приближенные решения антагонистических игр

Результаты этого раздела опубликованы в статьях [1], [2]. В этих статьях предполагалось, что заданы две системы с динамикой, задаваемой генератором типа Леви-Хинчина $L_t^i[u,v]$, $i = 1, 2$:

$$L_t^i[u,v]\phi(x) = \frac{1}{2}\langle G^i(t,x,u,v)\nabla,\nabla\rangle\phi(x) + \langle f^i(t,x,u,v),\nabla\rangle\phi(x) \\ + \int_{\mathbb{R}^d}[\phi(x+y) - \phi(x) - \langle y, \nabla\phi(x)\rangle \mathbf{1}_{B_1}(y)]\nu^i(t,x,u,v,dy).$$

Здесь u (соответственно, v) является управлением первого (соответственно, второго) игрока. Далее мы предполагаем, что каждый генератор L^i определен на пространстве $\mathcal{D}^i \subset C(\mathbb{R}^d)$, содержащем $C_b(\mathbb{R}^d)$, а также линейные и квадратичные функции. Первый игрок стремится минимизировать величину

$$\mathbb{E}\gamma(X(T)),$$

в то время как цель второго игрока противоположна. Мы предполагаем, что генератор L^1 задает динамику исходной системы, а генератор L^2 задает модельную систему, решение которой (в форме стабильной функции) предполагается известным. Целью упомянутых статей является построения приближенного решения для исходной задачи управления.

Определим используемый класс стратегий. Мы предполагаем, что игроки используют память и дополнительную информацию, которая поступает из модельной системы и имеет вероятностную природу. Это приводит к следующем определению.

Определение 3.1 [1, Definition 1]**.** Пусть t_0 – начальное время. Стратегией первого игрока на $[t_0, T]$ назовем набор $\mathfrak{u} = (\Omega^U, \mathcal{F}^U, \{\mathcal{F}_s^U\}_{s\in[t_0,T]}, u_{x(\cdot)}, P_{x(\cdot)}^U)$, удовлетворяющий следующим условиям:

1. $(\Omega^U, \mathcal{F}^U, \{\mathcal{F}_s^U\}_{s\in[t_0,T]})$ измеримое пространство с фильтрацией;
2. для каждой функции $x(\cdot) \in \mathbb{D}_{t_0}$ $u_{x(\cdot)}$ является $\{\mathcal{F}_s^U\}_{s\in[t_0,T]}$-прогрессивно измеримым процессом со значениями в U, а $P_{x(\cdot)}^U$ – вероятность на $(\Omega^U, \mathcal{F}^U, \{\mathcal{F}_s^U\}_{s\in[t_0,T]})$;
3. если для всех $s \in [t_0, t]$, $y(s) = x(s)$ то
   - для всех $A \in \mathcal{F}_t^U$ $P_{x(\cdot)}^U(A) = P_{y(\cdot)}^U(A)$;
   - для всех $s \in [t_0, t]$ $u_{x(\cdot)}(s) = u_{y(\cdot)}(s)$ $P_{x(\cdot)}^U$-п.н.;
4. для всякого $t \in [t_0, T]$ функция $(x(\cdot), s, \omega) \mapsto u_{x(\cdot)}(s, \omega)$ измерима относительно $\mathbb{F}_{t_0,t} \otimes \mathcal{B}([t_0, t]) \otimes \mathcal{F}_{t_0,t}^U$.

Здесь $x(\cdot)$ – траектория, выбираемая из пространства Скорохода $\mathbb{D}_{t_0} \triangleq D([t_0, T], \mathbb{R}^d)$. $\mathbb{F}_{s,t}$ обозначает σ-алгебру на $\mathbb{D}_{t_0}$, порожденную событиями на интервале $[s, t]$.

Стратегия $\mathfrak{u} = (\Omega^U, \mathcal{F}^U, \{\mathcal{F}^U_s\}_{s\in[t_0,T]}, u_{x(\cdot)}, P^U_{x(\cdot)})$ называется *пошаговой*, если существует разбиение $\Delta = \{t_l\}_{l=1}^r$ промежутка управления $[t_0, T]$ такое, что равенство $x(t_k) = y(t_k)$, $k = 0, \dots, l-1$, влечет равенства $P_{x(\cdot)}(A) = P_{y(\cdot)}(A)$ для всех $A \in \mathcal{F}^U_{t_l-0}$ и $u_{x(\cdot)}(s) = u_{y(\cdot)}(s)$ для всех $s \in [0, t_l)$.

Отметим, что введено определение стратегий включает в себя как пошаговые стратегии, так и рандомизированные пошаговые стратегии.

Стратегией второго игрока назовем набор $\mathfrak{v} = (\Omega^V, \mathcal{F}^V, \{\mathcal{F}^V_s\}_{s\in[t_0,T]}, v_{x(\cdot)}, P^V_{x(\cdot)})$, удовлетворяющий условиям, аналогичным условиям Определения 3.1 для $v_{x(\cdot)}$, принимающим значения во множестве V.

Определение 3.2 [1, Definition 2]**.** Пусть (t_0, x_0) – начальная позиция, и пусть $\mathfrak{u} = (\Omega^U, \mathcal{F}^U, \{\mathcal{F}^U_s\}_{s\in[0,T]}, u_{x(\cdot)}, P^U_{x(\cdot)})$, $\mathfrak{v} = (\Omega^V, \mathcal{F}^V, \{\mathcal{F}^V_s\}_{s\in[t_0,T]}, v_{x(\cdot)}, P^V_{x(\cdot)})$ являются стратегиями первого и второго игроков соответственно. Набор $(\Omega^X, \mathcal{F}^X, \{\mathcal{F}^X_s\}_{s\in[t_0,T]}, X(\cdot), P)$ является *реализацией движения, порожденного стратегиями* $\mathfrak{u}$, $\mathfrak{v}$ *и начальной позицией* (t_0, x_0), если выполнены следующие условия:

1. $(\Omega^X, \mathcal{F}^X, \{\mathcal{F}^X_s\}_{s\in[t_0,T]})$ – измеримое пространство с фильтрацией;

2. P – вероятность на $(\Omega, \mathcal{F}, \{\mathcal{F}_s\}_{s\in[t_0,T]})$, где $\Omega \triangleq \Omega^X \times \Omega^U \times \Omega^V$, $\mathcal{F} \triangleq \mathcal{F}^X \otimes \mathcal{F}^U \otimes \mathcal{F}^V$, $\mathcal{F}_s \triangleq \mathcal{F}^X_s \otimes \mathcal{F}^U_s \otimes \mathcal{F}^V_s$;

3. $X(\cdot)$ – случайный процесс на $(\Omega, \mathcal{F}, \{\mathcal{F}_s\}_{s\in[t_0,T]})$, согласованный с фильтрацией $\{\mathcal{F}_s\}_{s\in[t_0,T]}$, и значениями в $\mathbb{R}^d$;

4. $X(t_0) = x_0$ P-п.н.;

5. для всех $\phi \in \mathcal{D}^1$ процесс
$$\varphi(X(t)) - \int_{t_0}^t L^1_\tau[u(\tau), v(\tau)]\phi(X(\tau))d\tau \tag{18}$$
является $\{\mathcal{F}_s\}_{s\in[t_0,T]}$-мартингалом; здесь u и v – $\{\mathcal{F}_s\}_{s\in[t_0,T]}$-прогрессивно измеримые процессы, определяемые правилами
$$u(\tau, \omega^X, \omega^U, \omega^V) \triangleq u_{X(\cdot,\omega^X,\omega^U,\omega^V)}(\tau, \omega^U),$$
$$v(\tau, \omega^X, \omega^U, \omega^V) \triangleq v_{X(\cdot,\omega^X,\omega^U,\omega^V)}(\tau, \omega^V),$$
где $(\omega^X, \omega^U, \omega^V) \in \Omega$;

6. для всякого $x(\cdot) \in \mathbb{D}_{t_0}$ и произвольной случайной величины ϕ' на $(\Omega^U, \mathcal{F}^U)$
$$\mathbb{E}^U_{x(\cdot)}\phi' = \mathbb{E}(\phi'|X(\cdot) = x(\cdot)),$$
где $\mathbb{E}^U_{x(\cdot)}$ обозначает математическое ожидание в силу вероятности $P^U_{x(\cdot)}$;

7. для всякого $x(\cdot) \in \mathbb{D}_{t_0}$ и произвольной случайной величины ϕ'' на $(\Omega^V, \mathcal{F}^V)$
$$\mathbb{E}^V_{x(\cdot)}\phi'' = \mathbb{E}(\phi''|X(\cdot) = x(\cdot)),$$
где $\mathbb{E}^V_{x(\cdot)}$ обозначает математическое ожидание в силу вероятности $P^V_{x(\cdot)}$.

Отметим, что если обе стратегии являются пошаговыми, то можно показать существование хотя бы одной реализации. Однако для произвольных стратегий $\mathfrak{u}$, $\mathfrak{v}$ выигрыш определен, вообще говоря, неединственным способом. Величины

$$J^*(t_0, x_0, \mathfrak{u}, \mathfrak{v}) \triangleq \sup\{\mathbb{E}g(X(T)) : (\Omega^X, \mathcal{F}^X, \{\mathcal{F}^X_s\}_{s\in[0,T]}, X(\cdot), P) \text{ реализуют движение, порожденное стратегиями } \mathfrak{u} \text{ и } \mathfrak{v} \text{ и начальной позицией } (t_0, x_0)\},$$

$$J_*(t_0, x_0, \mathfrak{u}, \mathfrak{v}) \triangleq \inf\{\mathbb{E}g(X(T)) : (\Omega^X, \mathcal{F}^X, \{\mathcal{F}^X_s\}_{s\in[0,T]}, X(\cdot), P) \text{ реализуют движение, порожденное стратегиями } \mathfrak{u} \text{ и } \mathfrak{v} \text{ и начальной позицией } (t_0, x_0)\}.$$

являются верхней и нижней оценками выигрыша для стратегий $\mathfrak{u}$ и $\mathfrak{v}$. В этом случае, верхняя функция цены равна

$$\mathrm{Val}_+(t_0, x_0) = \inf_{\mathfrak{u}} \sup_{\mathfrak{v}} J^*(t_0, x_0, \mathfrak{u}, \mathfrak{v}).$$

А нижняя функция цены определяется как

$$\mathrm{Val}_-(t_0, x_0) = \sup_{\mathfrak{v}} \inf_{\mathfrak{u}} J_*(t_0, x_0, \mathfrak{u}, \mathfrak{v}).$$

Очевидно, что

$$\mathrm{Val}_-(t_0, x_0) \le \mathrm{Val}_+(t_0, x_0).$$

Основной результат статьи [1] состоит в оценках этих значений на основе решений модельной игры.

Мы предполагаем, что задано решение модельной игры в форме стабильной функции.

Определение 3.3 [1, Definition 4]**.** Функция $c_+ : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ называется u-стабильной для генератора L^2, если

1. $c_+(T, x) = g(x)$;

2. для всех $t, \theta \in [0, T]$, $t < \theta$, существует вероятностное пространство с фильтрацией $(\widetilde{\Omega}^{t,\theta}, \widetilde{\mathcal{F}}^{t,\theta}, \{\widetilde{\mathcal{F}}^{t,\theta}_s\}_{s\in[t,\theta]})$ такое, что для всех $y \in \mathbb{R}^d$, $v \in V$ можно найти $\{\widetilde{\mathcal{F}}^{t,\theta}_s\}_{s\in[t,\theta]}$-прогрессивно измеримое обобщенное управление первого игрока на $[t, \theta]$ $\mu^{t,\theta}_{y,v}$, согласованный с фильтрацией $\{\widetilde{\mathcal{F}}^{t,\theta}_s\}_{s\in[t,\theta]}$ процесс $Y^{t,\theta}_{y,v}$, принимающий значения в $\mathbb{R}^d$, и вероятность $\widetilde{P}^{t,\theta}_{y,v}$ на $\widetilde{\Omega}^{t,\theta}$ такие, что $Y^{t,\theta}_{y,v}(t) = y$ $\widetilde{P}^{t,\theta}_{y,v}$-п.н. для всех $\phi \in \mathcal{D}^2$,
$$\phi(Y^{t,\theta}_{y,v}(s)) - \int_t^s \int_U L^2_\tau[w, v]\phi(Y^{t,\theta}_{y,v}(\tau))\mu^{t,\theta}_{y,v}(\tau, dw)d\tau \tag{19}$$
является $\{\widetilde{\mathcal{F}}^{t,\theta}_s\}_{s\in[t,\theta]}$-мартингалом и
$$c_+(t, y) \ge \widetilde{\mathbb{E}}^{t,\theta}_{y,v} c_+(\theta, Y^{t,\theta}_{y,v}(\theta)); \tag{20}$$

3. для всех случайных величин ϕ на $\widetilde{\Omega}^{t,\theta}$ зависимость $\widetilde{\mathbb{E}}^{t,\theta}_{y,v}\phi$ от y и v измерима;

4. для всех $\phi \in \mathcal{D}^2$ функция $(y, v, s) \mapsto \widetilde{\mathbb{E}}^{t,\theta}_{y,v}\phi(Y^{t,\theta}_{y,v}(s))$ измерима.

Здесь $\widetilde{\mathbb{E}}^{t,\theta}_{y,v}$ обозначает математическое ожидание в силу вероятности $\widetilde{P}^{t,\theta}_{y,v}$.

Далее, положим

$$\Sigma^i(t,x,u,v) \triangleq \sum_{j=1}^{d} G^i_{jj}(t,x,u,v) + \int_{\mathbb{R}^d} \|y\|^2 \nu^i(t,x,u,v,dy);$$

$$b^i(t,x,u,v) \triangleq f^i(t,x,u,v) + \int_{\mathbb{R}^d \setminus B_1} y\nu^i(t,x,u,v,dy).$$

Функция Σ^i оценивает стохастичность системы, задаваемой генератором L^i, в то время как b^i играет роль эффективного сдвига. Мы накладываем условия непрерывности динамики (включая липшицевость b^i) и ограниченности функций Σ^i, b^i. Наиболее важным является условие Айзекса, которое состоит в том, что хотя бы для одного $i = 1, 2$ и всех $t \in [0,T]$, $x, p \in \mathbb{R}^d$, $u \in U$, $v \in V$

$$\min_{u \in U} \max_{v \in V} \langle p, b^i(t,x,u,v)\rangle = \max_{v \in V} \min_{u \in U} \langle p, b^i(t,x,u,v)\rangle.$$

Пусть M^i_0 – константа такая, что

$$|\Sigma^i(t,x,u,v)| \leq M^i_0 \quad t \in [0,T], \quad x \in \mathbb{R}^d, \quad u \in U, \quad v \in V,$$

K_i – константа Липшица для функции $x \mapsto b^i(t,x,u,v)$,

$$\varkappa \triangleq \frac{1}{2} \sup_{t \in [0,T], x \in \mathbb{R}^d, u \in U, v \in V} \|b^1(t,x,u,v) - b^2(t,x,u,v)\|^2,$$

$$\epsilon = \varkappa + M^1_0 + M^2_0. \tag{21}$$

Далее выберем i так чтобы для b^i выполнялось условие Айзекса и положим

$$\beta \triangleq 3 + 2K^i. \tag{22}$$

Наконец, пусть $C \triangleq \sqrt{2Te^{\beta T}}$.

Теорема 3.4 [1, Theorem 8]**.** *Если c_+ – u-стабильная функция для генератора L^2, то можно конструктивно построить стратегию $\widehat{\mathfrak{u}}_\Delta$ такую, что для всех $(t_0,x_0) \in [0,T] \times \mathbb{R}^d$*

$$\lim_{\delta \downarrow 0} \sup\{J(t_0,x_0,\widehat{\mathfrak{u}}_\Delta,\mathfrak{v}) : d(\Delta) \leq \delta\} \leq c_+(t_0,x_0) + R \cdot C\sqrt{\epsilon}.$$

Здесь R – константа Липшица для терминальной функции платы γ. Доказательство основано на аналоге метода экстремального сдвига для стохастических систем.

Общая теорема влечет следующие утверждения.

Следствие 3.5 [1, Corollary 9]**.** *Если c_+ – u-стабильная функция для генератора L^2, то, для всех $(t_0,x_0) \in [0,T] \times \mathbb{R}^d$*

$$\mathrm{Val}_+(t_0,x_0) \leq c_+(t_0,x_0) + R \cdot C\sqrt{\epsilon}.$$

Понятие v-стабильной функции может быть введено тем же самым способом, что и понятие u-стабильности. Тогда мы получаем следующую оценку.

Следствие 3.6 [1, Corollary 10]**.** *Если функция c_- – v-стабильна для генератора L^2, то для всех $(t_0,x_0) \in [0,T] \times \mathbb{R}^d$*

$$c_-(t_0,x_0) - R \cdot C\sqrt{\epsilon} \leq \mathrm{Val}_-(t_0,x_0).$$

3.2 Частные случаи

Рассмотрим ряд конкретных пар и на основе этих пар построим оценки решений игровых задач управления.

3.2.1 Стохастические модели дифференциальных игр

Мы предполагаем, что динамика исходной игры задается обыкновенным дифференциальным уравнением

$$\frac{d}{dt}x(t) = f(t, x(t), u, v), \tag{23}$$

а модельная система описывается стохастическим дифференциальным уравнением

$$dX(t) = f(t, X(t), u, t)dt + \sigma dW_t.$$

Таким образом, генераторы принимают вид

$$L_t^1[u, v]\phi(x) = \langle f(t, x, u, v), \nabla\phi(x)\rangle,$$

$$L_\tau^2[u, v]\varphi(x) = \langle f^1(t, x, u, v), \nabla\varphi(x)\rangle + \frac{\sigma^2}{2} \cdot \triangle\varphi(x).$$

Мы предполагаем, что условие Айзекса выполнено для исходной системы. В этом случае она имеет цену, которую мы обозначим через Val. Применяя общую теорию, можно получить следующее утверждение.

Предложение 3.7 [1, Proposition 16 и Corollary 18]**.** *Если ψ_σ является решением краевой задачи для дифференциального уравнения в частных производных*

$$\frac{\partial\psi}{\partial t} + \min_{u\in U}\max_{v\in V}\langle\nabla\psi, f^1(t, x, u, v)\rangle + \frac{\sigma^2}{2}\triangle\psi = 0, \quad \psi(T, x) = \gamma(x),$$

то ψ_σ – одновременно u- и v-стабильная функция относительно генератора

$$L_\tau^2[u, v]\phi(x) = \langle\nabla\phi(x), f^1(t, x, u, v)\rangle + \frac{\sigma^2}{2} \cdot \triangle\phi(x). \tag{24}$$

Таким образом, существует константа C_1 такая, что

$$|\mathrm{Val}(t_0, x_0) - \psi_\sigma(t_0, x_0)| \le RC_1\sigma.$$

3.2.2 Марковская модель дифференциальной игры

Как и выше мы предполагаем, что динамика исходной системы задается обыкновенным дифференциальным уравнением (23). Определим модельную систему с марковской динамикой следующим образом. Пусть h некоторое положительное число, $f^1(t, x, u, v) = (f_1^1(t, x, u, v), \ldots, f_d^1(t, x, u, v))$ и пусть e^i обозначает i-й координатный вектор. Обозначим

$$\chi_i(t, x, u, v) = \begin{cases} e^i, & f_i^1(t, x, u, v) > 0, \\ -e^i, & f_i^1(t, x, u, v) < 0, \\ 0, & f_i^1(t, x, u, v) = 0. \end{cases}$$

Функция $\chi_i(t,x,u,v)$ показывает направление движения в силу динамики $f^1(t,x,u,v)$ по i-й оси.

Для $A\subset\mathbb{R}^d$ положим

$$\nu^2(t,x,u,v,A)\triangleq\frac{1}{h}\sum_{i=1}^n|f_i(t,x,u,v)|\delta_{h\chi_i(t,x,u,v)}(A).$$

Напомним, что δ_z обозначает меру Дирака, сосредоточенную z.

Далее, определим генератор L^2 по правилу

$$\begin{aligned}L^2_t[u,v]\phi(x)&\triangleq\int_{\mathbb{R}^d}[\phi(x+y)-\phi(x)]\nu^2(t,x,u,v,dy)\\&=\sum_{i=1}^n|f_i(t,x,u,v)|\frac{\phi(x+h\chi_i(t,x,u,v))-\phi(x)}{h}.\end{aligned}\tag{25}$$

Этот генератор соответствует марковской цепи на $h\mathbb{Z}^d$ с матрицей Колмогорова

$$Q^h_{xy}(t,u,v)=\begin{cases}\frac{1}{h}|f_i(t,x,u,v)|, & y=x+h\chi_i(t,x,u,v),\\ -\frac{1}{h}\sum_{i=1}^d|f_i(t,x,u,v)|, & x=y,\\ 0, & y\neq x,\quad y\neq x+h\chi_i(t,x,u,v),\end{cases}\tag{26}$$

Следующая система является аналогом уравнения Айзекса-Беллмана для марковской цепи с матрицей Колмогорова, заданной (26):

$$\frac{d}{dt}\zeta^+_h(t,x)+\min_{u\in U}\max_{v\in V}\sum_{i=1}^d|f_i(t,x,u,v)|\frac{\zeta^+_h(t,x+h\chi_i(t,x,u,v))-\zeta^+_h(t,x)}{h}=0,$$
$$\zeta^+_h(T,x)=\gamma(x).\tag{27}$$

Здесь $x\in h\mathbb{Z}^d$ – параметр.

Теорема 3.8 [1, Proposition 20 и Theorem 21]**.** *Уравнение (27) имеет единственное решение. Оно задает следующую оценку для функции цены дифференциальной игры*

$$|\mathrm{Val}(t_0,x_0)-\zeta^+_h(t_0,x_0)|\leq RC_2\sqrt{h}.$$

Здесь C_2 – константа, определяемая функцией f.

3.2.3 Детерминированная модель для мультиагентной системы

Мы рассматриваем конструкции управления системой, состоящей из N однотипных взаимодействующих частиц, каждая из которых принимает не более чем конечное число значений. Мы предполагаем, что

- множество состояний есть $\{1,\ldots,d\}$, где d – натуральное число;
- динамика каждой частицы задается марковской цепью с матрицей Колмогорова $Q(t,x,u,v)=(Q_{i,j}(t,x,u,v))_{i,j=\overline{1,d}}$; здесь $x=(x_1,\ldots,x_d)$ – вектор, описывающий населенность уровней.

Отметим, что $x \in \mathbb{R}^d$, $x_i \geq 0$, $\sum_{i=1}^d x_i = 1$. Полагая, что $h = 1/N$, мы получаем, что исследуемая система описывается генератором

$$L^1_t[u,v]\phi(x) = \sum_{i,j=1}^d \frac{1}{h} x_i Q_{ij}(t,x,u(t),v(t))[\phi(x - he^i + he^j) - \phi(x)].$$

Здесь e^i – i-й орт.

В пределе $h \to 0$, мы получаем следующую модельную систему:

$$L^2_t[u,v]\phi(x) = \sum_{k=1}^d \sum_{i \neq k} [x_i Q_{ik}(t,x,u(t),v(t)) - x_k Q_{ki}(t,x,u(t),v(t))] \frac{\partial \phi}{\partial x_k}(x).$$

Характеристики этой системы удовлетворяют обыкновенному дифференциальному уравнению

$$\frac{d}{dt} x(t) = x(t) Q(t, x(t), u(t), v(t)). \tag{28}$$

В [2] (независимо от [1]) был получен следующий результат.

Теорема 3.9 [2, Theorem 1]. *Пусть c_+ – u-стабильная функция для дифференциальной игры с динамикой (28), тогда можно построить стратегию первого игрока $\widehat{\mathfrak{u}}_\Delta$ такую, что для всех $(t_0,x_0) \in [0,T] \times \mathbb{R}^d$*

$$\lim_{\delta \downarrow 0} \sup\{J(t_0, x_0, \widehat{\mathfrak{u}}_\Delta, \mathfrak{v}) : d(\Delta) \leq \delta\} \leq c_+(t_0,x_0) + R \cdot C_3 \sqrt{h}.$$

Здесь C_3 – константа, определяемая матрицей Q и числом d.

Аналогичный результат справедлив и для v-стабильных функций. Наконец, если игровая задача управления для системы, состоящей из N частиц ($h = 1/N$), имеет цену, то имеет место следующая оценка

$$|\mathrm{Val}^h(t_0,x_0) - \mathrm{Val}(t_0,x_0)| \leq RC_3\sqrt{h}.$$

Здесь мы обозначили через Val^h – функцию цены в игровой задаче управления N частицами ($N = 1/h$), а через Val функцию цены в предельной игре с динамикой (28).

3.3 Приближенные равновесия в неантагонистических играх

В настоящем разделе мы используем методологию статьи [1] для неантагонистических игр. Напомним, что во многих ситуациях неантагонистические стохастические дифференциальные игры и неантагонистические марковские игры оказываются проще дифференциальных игр. Таким образом, заманчивой является идея построить приближенное равновесие в дифференциальной игре на основе решения неантагонистической дифференциальной игры со стохастической динамикой.

Мы рассматриваем дифференциальную игру с динамикой

$$\dot{x} = f_1(t,x,u) + f_2(t,x,v), \quad t \in [0,T], x \in \mathbb{R}^d, \quad u \in U, \quad v \in V. \tag{29}$$

Здесь u (соответственно, v) обозначает управление первого (соответственно, второго) игрока. Мы предполагаем, что i-й игрок стремится максимизировать терминальный показатель качества $\gamma_i(x(T))$. Кроме того, полагаем, что множества U и V – метрические компакты. Для краткости будем также использовать обозначение $f^1(t,x,u,v) \triangleq f_1(t,x,u) + f_2(t,x,v)$.

Приближенное равновесие по Нэшу будет построено в классе стохастических стратегий с общим сигналом и памятью. Неформально, этот класс стратегий можно описать следующим образом. Мы предполагаем, что игроки наблюдают как всю предысторию, так и некоторый внешний сигнал. В предложенной конструкции этот сигнал продуцируется модельной системой.

Эта идея формализуется следующим образом.

Определение 3.10 [3, Definition 2.1]**.** Набор $\mathfrak{w} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[t_0,T]}, u_{x(\cdot)}, v_{x(\cdot)}, P_{x(\cdot)})$ называется набором стратегий с общим сигналом и памятью на $[t_0,T]$, если

(i) $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[t_0,T]})$ – измеримое пространство с фильтрацией;

(ii) для каждого $x(\cdot) \in C([t_0,T];\mathbb{R}^d)$ $P_{x(\cdot)}$ – вероятность на $\mathcal{F}$;

(iii) для каждого $x(\cdot) \in C([t_0,T];\mathbb{R}^d)$ $u_{x(\cdot)}$ (соответственно, $v_{x(\cdot)}$) является $\{\mathcal{F}_t\}_{t\in[t_0,T]}$-прогрессивно измеримым процессом со значениями в U (соответственно, в V);

(iv) если $x(t) = y(t)$ для всех $t \in [t_0, r]$, то

- для каждого $A \in \mathcal{F}_r$ $P_{x(\cdot)}(A) = P_{y(\cdot)}(A)$,
- для каждого $t \in [t_0, r]$ $u_{x(\cdot)}(t) = u_{y(\cdot)}(t)$, $v_{x(\cdot)}(t) = v_{y(\cdot)}(t)$ $P_{x(\cdot)}$-п.н.

(v) для любых r сужение функций $(x(\cdot), t, \omega) \mapsto u_{x(\cdot)}(t,\omega)$, $(x(\cdot), t, \omega) \mapsto v_{x(\cdot)}(t,\omega)$ на $C([t_0,T];\mathbb{R}^d) \times [t_0,r] \times \Omega$ измеримы относительно $\mathbb{F}_{t_0,r} \otimes \mathcal{B}([t_0,r]) \otimes \mathcal{F}_r$;

(vi) для всех $A \in \mathcal{F}$ функции $x(\cdot) \mapsto P_{x(\cdot)}(A)$ измеримы относительно $\mathbb{F}_{t_0,T}$.

Ключевой концепцией в определении понятии равновесия по Нэшу является одностороннее уклонение. В рамках рассматриваемого подхода предполагается, что игрок-уклонист может использовать дополнительно информацию от некоторого другого устройства. Это приводит к следующему определению.

Определение 3.11 [3, Definition 2.2]**.** Пусть задан набор стратегий с памятью и общим сигналом $\mathfrak{w} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[t_0,T]}, P_{x(\cdot)}, u_{x(\cdot)}, v_{x(\cdot)})$. Будем говорить, что набор стратегий $\mathfrak{w}^c = (\Omega^c, \mathcal{F}^c, \{\mathcal{F}^c_t\}_{t\in[t_0,T]}, P^c_{x(\cdot)}, u^c_{x(\cdot)}, v^c_{x(\cdot)})$ является односторонним отклонением первого (соответственно, второго) игрока, если можно найти вероятностное пространство $(\Omega', \mathcal{F}', \{\mathcal{F}'\}_{t\in[t_0,T]})$ такое, что

(i) $\Omega^c = \Omega \times \Omega'$;

(ii) $\mathcal{F}^c = \mathcal{F} \otimes \mathcal{F}'$;

(iii) $\mathcal{F}^c_t = \mathcal{F}_t \otimes \mathcal{F}'_t$ для $t \in [t_0,T]$;

(iv) для всех $x(\cdot) \in C([t_0,T];\mathbb{R}^d)$ и всех $A \in \mathcal{F}$ $P^c_{x(\cdot)}(A \times \Omega') = P_{x(\cdot)}(A)$;

(v) для всех $x(\cdot)$, $t \in [t_0, T]$, $\omega \in \Omega$, $\omega' \in \Omega'$ $v_{x(\cdot)}(t, \omega, \omega') = v_{x(\cdot)}(t, \omega)$ (соответственно, $u_{x(\cdot)}(t, \omega, \omega') = u_{x(\cdot)}(t, \omega)$).

Введем теперь понятие реализации движения, порожденного набором стратегий с памятью и общим сигналом.

Определение 3.12 [3, Definition 2.3]**.** Пусть $t_0 \in [0, T]$, $x_0 \in \mathbb{R}^d$, $\mathfrak{w} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [t_0, T]}, P_{x(\cdot)}, u_{x(\cdot)}, v_{x(\cdot)})$ – набор стратегий с памятью и общим сигналом $[t_0, T]$. Будем говорить, что пара $(X(\cdot), P)$ является реализацией движения, порожденного $\mathfrak{w}$ и начальной позицией (t_0, x_0), если

(i) P – вероятность на $\mathcal{F}$;

(ii) $X(\cdot)$ – процесс, согласованный с фильтрацией $\{\mathcal{F}_t\}_{t \in [t_0, T]}$, принимающий значения в $\mathbb{R}^d$;

(iii) $X(t_0) = x_0$ P-п.н.;

(iv) для P-п.в. $\omega \in \Omega$

$$\frac{d}{dt} X(t, \omega) = f_1(t, X(t, \omega), u_{X(\cdot, \omega)}(t, \omega)) + f_2(t, X(t, \omega), v_{X(\cdot, \omega)}(t, \omega)).$$

(v) $P_{x(\cdot)} = P(\cdot | X(\cdot) = x(\cdot))$ т.е. для заданного $A \in \mathcal{F}$

$$P(A) = \int_{C([t_0, T]; \mathbb{R}^d)} P_{x(\cdot)}(A) \chi(d(x(\cdot))),$$

где χ – вероятность $C([t_0, T]; \mathbb{R}^d)$ определенная по правилу: для $\Upsilon \in \mathbb{F}_{t_0, T}$ $\chi(\Upsilon) \triangleq P\{\omega : X(\cdot, \omega) \in \Upsilon\}$.

Отметим, что если набор стратегий с памятью и общим сигналом является пошаговым (т.е. стратегии определяются значениями предыстории лишь в конечном числе точек), то можно доказать существование хотя бы одной реализации.

Отметим, что динамика (8) соответствует генератору

$$L_t^1[u, v]\phi(x) \triangleq \langle f_1(t, x, u) + f_2(t, x, v), \nabla \phi(x) \rangle.$$

Чтобы построить приближенное равновесие мы используем вспомогательную стохастическую игру с непрерывным временем, у которой динамика задается генератором

$$(L_t^2[u, v]\phi)(x) \triangleq \frac{1}{2} \langle G^2(t, x, u, v) \nabla, \nabla \rangle \phi(x) + \langle f^2(t, x, u, v), \nabla \rangle \phi(x)$$
$$+ \int_{\mathbb{R}^d} [\phi(x + y) - \phi(x) - \langle y, \nabla \phi(x) \rangle \mathbf{1}_{B_1}(y)] \nu^2(t, x, u, v, dy). \quad (30)$$

Мы предлагаем, что целевой функционал игрока i во вспомогательной стохастической игре равен

$$\mathbb{E} \left[\gamma_i(X(T)) + \int_{t_0}^{T} g_i(t, X(t), u(t), v(t)) dt \right]. \quad (31)$$

Следующее условие стабильности играет ключевую роль в конструкции приближенно оптимально равновесия в классе стратегий с памятью и общим сигналом.

Определение 3.13 [3, Definition 3.2]**.** Пусть $c_1, c_2 : [0,T]\times\mathbb{R}^d \to \mathbb{R}$ – непрерывные функции. Будем говорить, что пара функций (c_1, c_2) удовлетворяет *условию* $(\mathcal{C})$, если для всех $s, r \in [0,T]$, $s < r$, существует измеримое пространство с фильтрацией $(\widehat{\Omega}^{s,r}, \widehat{\mathcal{F}}^{s,r}, \{\widehat{\mathcal{F}}^{s,r}_t\}_{t\in[s,r]})$, удовлетворяющее следующим свойствам:

(i) для заданного $y \in \mathbb{R}^d$ можно найти процессы $\eta^{s,r}_y$, $\widehat{Y}^{s,r}_y$ и вероятность $\widehat{P}^{s,r}_y$ так, чтобы набор $(\widehat{\Omega}^{s,r}, \widehat{\mathcal{F}}^{s,r}, \{\widehat{\mathcal{F}}^{s,r}_t\}_{t\in[s,r]}, \widehat{P}^{s,r}_y, \eta^{s,r}_y, \widehat{Y}^{s,r}_y)$ был бы управляемой системой, допустимой для генератора $L^2_t[u,v]$, и для $i = 1, 2$

$$\widehat{\mathbb{E}}^{s,r}_y \left[c_i(r, \widehat{Y}^{s,r}_y(r)) + \int_s^r \int_{U\times V} g_i(t, \widehat{Y}^{s,r}_y(t), u, v)\eta^{s,r}_y(t, d(u,v))dt \right] = c_i(s, y);$$

(ii) для произвольных $y \in \mathbb{R}^d$ и $v \in V$ можно найти обобщенное управление первого игрока $\mu^{s,r}_{y,v}$, процесс $\overline{Y}^{1,s,r}_{y,v}$, принимающий значение в $\mathbb{R}^d$, и вероятность $\overline{P}^{1,s,r}_{y,v}$ такие, что набор $(\widehat{\Omega}^{s,r}, \widehat{\mathcal{F}}^{s,r}, \{\widehat{\mathcal{F}}^{s,r}_t\}_{t\in[s,r]}, \overline{P}^{1,s,r}_{y,v}, \mu^{s,r}_{y,v}\otimes\delta_v, \overline{Y}^{1,s,r}_{y,v})$ является управляемым процессом, допустимым для генератора $L^2_t[u,v]$, и

$$\overline{\mathbb{E}}^{1,s,r}_{y,v} \left[c_2(r, \overline{Y}^{1,s,r}_{y,v}(r)) + \int_s^r \int_U g_2(t, \overline{Y}^{1,s,r}_{y,v}(t), u, v)\mu^{s,r}_{y,v}(t, du)dt \right] \le c_2(s, y);$$

(iii) для произвольных $y \in \mathbb{R}^d$ и $u \in U$ можно найти обобщенное управление второго игрока $\nu^{s,r}_{y,u}$, процесс $\overline{Y}^{2,s,r}_{y,u}$ и вероятность $\overline{P}^{2,s,r}_{y,u}$ так, чтобы набор $(\widehat{\Omega}^{s,r}, \widehat{\mathcal{F}}^{s,r}, \{\widehat{\mathcal{F}}^{s,r}_t\}_{t\in[s,r]}, \overline{P}^{2,s,r}_{y,u}, \delta_u\otimes\nu^{s,r}_{y,u}, \overline{Y}^{2,s,r}_{y,u})$ был бы управляемым процессом, допустимым для генератора $L^2_t[u,v]$, и

$$\overline{\mathbb{E}}^{2,s,r}_{y,u} \left[c_1(r, \overline{Y}^{2,s,r}_{y,u}(r)) + \int_s^r \int_V g_1(t, \overline{Y}^{2,s,r}_{y,u}(t), u, v)\nu^{s,r}_{y,u}(t, dv)dt \right] \le c_1(s, y).$$

Здесь $\widehat{\mathbb{E}}^{s,r}_y$ (соответственно, $\overline{\mathbb{E}}^{1,s,r}_{y,u}$, $\overline{\mathbb{E}}^{2,s,r}_{y,u}$) обозначает математическое ожидание в силу вероятности $\widehat{P}^{s,r}_y$ (соответственно, $\overline{P}^{1,s,r}_{y,u}$, $\overline{P}^{2,s,r}_{y,u}$).

Неформально смысл условия $(\mathcal{C})$ следующий. Первая часть говорит, что оба игрока могут совместными действиями поддерживать выигрыш на уровне $(c_1(s,y), c_2(s,y))$ на любом интервале $[s,r]$. Условия (ii), (iii) говорят о том, что если игрок i выбирает постоянное управление на промежутке управления $[s,r]$, то другой игрок может своими действиями добиться того, что выигрыш игрока i на промежутке $[s,r]$ будет не больше $c_i(s,y)$. Здесь предполагается, что терминальный показатель платы i-го игрока на промежутке $[s,r]$ равен $c_i(r,\cdot)$. Кроме того, чтобы избежать технических сложностей, мы предполагаем, что во всех трех случаях можно использовать общее вероятностное пространство.

Далее наложим условие близости исходной и модельной систем, которое формулируется следующим образом: для всех $t \in [0,T]$, $x \in \mathbb{R}^d$, $u \in U$, $v \in V$

$$|\Sigma^2(t,x,u,v)| \le \delta^2, \quad \|f^1(t,x,u,v) - b^2(t,x,u,v)\|^2 \le 2\delta^2, \quad |g_i(t,x,u,v)| \le \delta.$$

Здесь, как и ранее,

$$\Sigma^2(t,x,u,v) \triangleq \sum_{j=1}^d G^2_{jj}(t,x,u,v) + \int_{\mathbb{R}^d} \|y\|^2\nu^2(t,x,u,v,dy);$$

$$b^2(t,x,u,v) \triangleq f^2(t,x,u,v) + \int_{\mathbb{R}^d \setminus B_1} y\nu^2(t,x,u,v,dy).$$

Предполагая дополнительно некоторые условия регулярности, мы получаем следующее утверждение.

Теорема 3.14 [3, Theorem 3.3]. *Пусть функции $c_1, c_2 : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ таковы, что*

- $c_i(T,x) = \gamma_i(x)$*;*
- (c_1, c_2) *удовлетворяет условию* $(\mathcal{C})$*.*

Тогда для всех $(t_0, x_0) \in [0,T] \times \mathbb{R}^d$, и $\varepsilon > (RC+T)\delta$ существует набор стратегий с памятью и общим сигналом $\mathfrak{w}^*$, который задает ε-равновесие в (t_0,x_0). Более того, если X^* и P^* порождены $\mathfrak{w}^*$ и (t_0,x_0), а $\mathbb{E}^*$ обозначает математическое ожидание в силу вероятности P^*, то

$$|\mathbb{E}^* \gamma_i(X^*(T)) - c_i(t_0,x_0)| \leq \varepsilon.$$

В Теореме 3.14 C есть константа, определяемая лишь динамикой исходной игры.

Естественным является вопрос о предельном поведеним системы в случае, когда δ стремится к 0. В этой связи получен следующий результата [3, Theorem 5.2]: предельные значения ожидаемых выигрышей игроков, соответствующих приближенным равновесиям в классе стохастических стратегий с памятью и общим сигналом, лежат в выпуклой оболочке множества выигрышей в классе детерминированных стратегий наказания.

3.4 Равновесия, построенные по системам уравнений Беллмана

3.4.1 Случай гладкий решений

Основной результат в этом направлении следующий.

Теорема 3.15 [3, Theorem 4.1]. *Пусть $c_1, c_2 : [0,T] \times \mathbb{R}^d \to \mathbb{R}$, $u^0 : [0,T] \times \mathbb{R}^d \to U$, $v^0 : [0,T] \times \mathbb{R}^d \to V$ класса C^2 и удовлетворяют следующим условиям:*

$$\begin{aligned} &\frac{\partial c_i}{\partial t} + L_t^2[u^0(t,x), v^0(t,x)]c_i(t,x) + g_i(t,x,u^0(t,x),v^0(t,x)) = 0, \\ &c_i(T,x) = \gamma_i(x), \end{aligned} \tag{32}$$

где $u^0(t,x)$ и $v^0(t,x)$ задают равновесие по Нэшу в статической игре с платежными функциями

$$L_t^2[u,v]c_i(t,x) + g_i(t,x,u,v), \quad i = 1,2.$$

Предположим, дополнительно, что для заданного промежутка $[s,r] \subset [0,T]$ существует решение задачи о мартингалах на $[s,r]$ с генераторами $L_t^2[u^0(t,\cdot), v^0(t,\cdot)]$, $L_t^2[u, v^0(t,\cdot)]$, $L_t^2[u^0(t,\cdot), v]$, $u \in U$, $v \in V$. Более того, можно выбрать общее вероятностное пространство с фильтрацией, подходящее для всех упомянутых задач.

Тогда пара (c_1,c_2) удовлетворяет условию $(\mathcal{C})$. В частности, для любых $(t_0,x_0) \in [0,T] \times \mathbb{R}^d$, $\varepsilon > (RC+T)\delta$ существует ε-равновесие в классе стохастических стратегий с памятью и общим сигналом в (t_0,x_0).

3.4.2 Приближенные равновесия на основе решений систем уравнений в частных производных второго порядка

Предположение о том, что система уравнений Беллмана имеет C^2-гладкое решение вообще говоря весьма ограничительно. Мы его ослабим в предположении о том, что вспомогательная игра имеют динамику, определяемую стохастическим дифференциальным уравнением следующего вида:

$$dX(t) = f^2(t, X(t), u(t), v(t))dt + \sigma(t, X(t))dW(t). \tag{33}$$

Эта система соответствует генератору

$$L_t^2[u,v]\phi(x) = \langle f^2(t,x,u,v), \nabla\phi(x)\rangle + \frac{1}{2}\langle G^2(t,x)\nabla, \nabla\phi(x)\rangle, \tag{34}$$

где $G^2(t,x) = \sigma(t,x)\sigma^T(t,x)$. Мы предполагаем, что $W(t)$ – d-мерный винеровский процесс, σ – невырожденная и ограниченная $d \times d$-матрица. Принципиальным является предположение о том, что корректно определены функции $u^N(t,x,p_1,p_2)$, $v^N(t,x,p_1,p_2)$ со значениями в U и V соответственно такие, что для любых $t \in [0,T]$, $x, p_1, p_2 \in \mathbb{R}^d$, $u \in U$, $v \in V$

$$\mathcal{H}_1(t,x,p_1,u^N(t,x,p_1,p_2), v^N(t,x,p_1,p_2)) \geq \mathcal{H}_1(t,x,p_1,u,v^N(t,x,p_1,p_2)),$$

$$\mathcal{H}_2(t,x,p_2,u^N(t,x,p_1,p_2), v^N(t,x,p_1,p_2)) \geq \mathcal{H}_2(t,x,p_2,u^N(t,x,p_1,p_2),v).$$

Здесь I обозначает единичную матрицу, а

$$\mathcal{H}_i(t,x,p,u,v) \triangleq \langle p, f^2(t,x,u,v)\rangle + g_i(t,x,u,v).$$

Теорема 3.16 [3, Theorem 4.5]. *Предположим, что (c_1, c_2) – сильное решение уравнения*

$$\frac{\partial c_i}{\partial t} + \mathcal{H}_i(t,x,\nabla c_i, u^N(t,x,\nabla c_1,\nabla c_2), v^N(t,x,\nabla c_1,\nabla c_2)) \\ + \langle G(t,x)\nabla,\nabla\rangle c_i(t,x) = 0, \quad c_i(T,x) = \gamma_i(x).$$

Тогда (c_1, c_2) удовлетворяет условию $(\mathcal{C})$ для генератора L^2, задаваемого формулой (34).

3.4.3 Приближенные равновесия на основе решений систем дифференциальных включений

Перепишем систему (8) в координатной форме

$$\frac{d}{dt}x_j(t) = f_{1,j}(t, x_1(t), \ldots, x_d(t), u) + f_{2,j}(t, x_1(t), \ldots, x_d(t), v), \quad j = 1, \ldots d.$$

Здесь $x_j(t)$ обозначает j-ю координату вектора $x(t)$. Пусть h – некоторое положительное числа. Как и выше, через e^j мы обозначаем j-й координатный вектор.

Положим

$$Q^1_{x,y}(t,u) \triangleq \begin{cases} \frac{1}{h}|f_{1,j}(t,x,u)|, & y = x + h\,\mathrm{sgn}(f_{1,j}(t,x,u))\cdot e^j, \\ -\frac{1}{h}\sum_{j=1}^d |f_{1,j}(t,x,u)|, & y = x, \\ 0, & \text{иначе.} \end{cases}$$

Также, пусть

$$Q^2_{x,y}(t,v) \triangleq \begin{cases} \frac{1}{h}|f_{2,j}(t,x,v)|, & y = x + h\,\mathrm{sgn}(f_{2,j}(t,x,v))\cdot e^j, \\ -\frac{1}{h}\sum_{j=1}^d |f_{2,j}(t,x,v)|, & y = x, \\ 0, & \text{иначе.} \end{cases}$$

Мы рассмотрим марковскую игру с пространством состояний $h\mathbb{Z}^d$ и матрицей Колмогорова $Q(t,u,v) = Q^1(t,u) + Q^2(t,v)$, в которой первый игрок стремится максимизировать показатель качества

$$\mathbb{E}\left[\gamma_1(X(T)) + \int_{t_0}^T g_1(t,X(t),u(t))dt\right],$$

а второй максимизирует величину

$$\mathbb{E}\left[\gamma_2(X(T)) + \int_{t_0}^T g_2(t,X(t),v(t))dt\right].$$

В этом случае аналог системы уравнений Беллмана строится следующим образом. Для $t \in [0,T]$, $x \in h\mathbb{Z}^d$, $\zeta : h\mathbb{Z}^d \to \mathbb{R}$, $\mu \in \mathcal{P}(U)$, $\nu \in \mathcal{P}(V)$ положим

$$\widehat{H}_1(t,x,\zeta,\mu) \triangleq \int_U \left[\sum_{y\in h\mathbb{Z}^d} Q^1_{x,y}(t,x,u)\zeta(y) + g_1(t,x,u)\right]\mu(du),$$

$$\widehat{H}_2(t,x,\zeta,\nu) \triangleq \int_V \left[\sum_{y\in h\mathbb{Z}^d} Q^2_{x,y}(t,x,v)\zeta(y) + g_2(t,x,v)\right]\nu(dv).$$

Функции $\widehat{H}_1$, $\widehat{H}_2$ играют роль предгамильтонианов. Далее, положим

$$\mathcal{O}_1(t,x,\zeta) \triangleq \underset{\mu\in\mathcal{P}(U)}{\mathrm{Argmax}}\widehat{H}_1(t,x,\zeta,\mu), \quad \mathcal{O}_2(t,x,\zeta) \triangleq \underset{\nu\in\mathcal{P}(V)}{\mathrm{Argmax}}\widehat{H}_2(t,x,\zeta,\nu).$$

Наконец, если ζ_1,ζ_2 – вещественнозначные функции, определенные на $h\mathbb{Z}^d$, то обозначим

$$\begin{aligned}\mathcal{H}_1(t,x,\zeta_1,\zeta_2) \triangleq \max_{\mu\in\mathcal{P}(U)} &\widehat{H}_1(t,x,\zeta_1,\mu) \\ &+ \left\{\int_V \sum_{y\in h\mathbb{Z}^d} Q^2_{x,y}(t,x,v)\zeta_1(y)\nu(du) : \nu \in \mathcal{O}_2(t,x,\zeta_2)\right\},\end{aligned}$$

$$\begin{aligned}\mathcal{H}_2(t,x,\zeta_1,\zeta_2) \triangleq \max_{\nu\in\mathcal{P}(V)} &\widehat{H}_2(t,x,\zeta_2,\nu) \\ &+ \left\{\int_U \sum_{y\in h\mathbb{Z}^d} Q^1_{x,y}(t,x,u)\zeta_2(y)\mu(du) : \mu \in \mathcal{O}_1(t,x,\zeta_1)\right\}.\end{aligned}$$

Многозначные отображения $\mathcal{H}_1$ и $\mathcal{H}_2$ являются аналогами гамильтонианов.

В рассматриваемой неантагонистической марковской игре в качестве аналога системы уравнений Беллмана выступает следующая система дифференциальных включений:

$$\frac{d}{dt}\zeta_i(t,x) \in -\mathcal{H}_i(t,x,\zeta_1,\zeta_2), \quad \zeta_i(T,x) = \gamma_i(x), \quad i = 1,2. \tag{35}$$

Теорема 3.17 [4, Theorem 2]. *Пусть* $(\zeta_1(\cdot,\cdot),\zeta_2(\cdot,\cdot))$ *является решением (35). Тогда она удовлетворяет условию* $(\mathcal{C})$. *В частности, если* $|g_1(t,x,u)|, |g_2(t,x,v)| \leq \sqrt{h}$, *то можно построить* ε*-равновесие в классе стохастических стратегий с памятью и общим сигналом для всех* $\varepsilon > (RC+T)\sqrt{h}$.

3.5 Выживаемость для управляемых систем с динамикой среднего поля

Настоящий раздел основан на статье [5]. Рассмотрим задачу управления системой с динамикой среднего поля, при которой динамика каждого элемента системы задается обыкновенным дифференциальным уравнением

$$\frac{d}{dt}x(t) = f(t, x(t), m(t), u(t)). \tag{36}$$

Здесь $t \in [0,T]$ – время, $x(t)$ – обозначает положение элемента системы (агента), $m(t)$ – вероятность, описывающая распределение агентов, $u(t)$ – мгновенное управление агента. Для простоты в работы [5] предлагалось что фазовое пространство в настоящем случае есть d-мерный тор $\mathbb{T}^d \triangleq \mathbb{R}^d/\mathbb{Z}^d$. Формально интегрируя (36), мы приходим к следующем уравнению для потока вероятностей $m(\cdot)$

$$\frac{d}{dt}m(t) = \langle f(t,\cdot,m(t),u(t)), \nabla\rangle m(t). \tag{37}$$

Конечно можно задать поток вероятностей, используя Определение 1.2, положив генератор равным $L_t[m,u]\phi(x) = \langle f(t,x,m,u), \nabla\phi(x)\rangle$. Однако, в данном случае более удобно уточнить это определение, зафиксировав вероятностное пространство, которое определяет поток вероятностей.

Прежде всего нам требуются некоторые дополнительные обозначения. Если X и Y – польские пространства, m – мера X, то обозначим через $\Lambda(X,m,Y)$ множество мер на $X\times Y$ с маргинальным распределением на X равным m. Используя теорему о дезинтегрирование мер, мы отождествляем $\xi \in \Lambda(X,m,Y)$ с мерозначным отображением $X \ni x \mapsto \xi(\cdot|x) \in \mathcal{P}(Y)$ таким, что $\phi \in C_b(X\times Y)$,

$$\int_{X\times Y}\phi(x,y)\xi(d(x,y)) = \int_X\int_Y \phi(x,y)\xi(dy|x)m(dx).$$

Далее, обозначим

$$\mathcal{U} \triangleq \Lambda([0,T],\lambda,U),$$

где λ обозначает меру Лебега. Для заданной начальной позиции $(s,y)\in[0,T]\times\mathbb{T}^d$, $\xi\in\mathcal{U}$ и потока вероятностей $m(\cdot)$, траектория $x(\cdot)$ порождена этой начальной позицией и управлением ξ, если $x(\cdot)$ является решением задачи Коши для уравнения

$$\frac{d}{dt}x(t) = \int_U f(t,x(t),m(t),u)\xi(du|t), \quad x(s) = y.$$

Обозначим оператор, сопоставляющий y и ξ соответствующее движение, через $\mathrm{traj}^s_{m(\cdot)}$.

Далее, если $t\in[0,T]$, то e_t является оператором проектирования $C([0,T],\mathbb{T}^d)$ на $\mathbb{T}^d$, определенным по правилу: для $x(\cdot)\in C([0,T],\mathbb{T}^d)$

$$e_t(x(\cdot)) = x(t).$$

Для заданного $m \in \mathcal{P}^p(\mathbb{T}^d)$ положим

$$\mathcal{A}[m] \triangleq \Lambda(\mathbb{T}^d, m, \mathcal{U}).$$

Множество $\mathcal{A}[m]$ является множеством распределением пар (y, ξ) с маргинальным распределением на $\mathbb{T}^d$ равным m. Для заданного распределения $\alpha \in \mathcal{A}[m]$, мы предполагаем, что агенты, находящиеся в начальный момент в состоянии x выбирают свои управления в соответствии с распределением $\alpha(d\xi|x)$.

Определение 3.18. Предположим, что $s \in [0, T]$ – начальное время, m_* – начальное распределение агентов, $\alpha \in \mathcal{A}(m_*)$. Будем говорить, что поток вероятностей $[0, t] \ni t \mapsto m(t) \in \mathcal{P}^p(\mathbb{T}^d)$ порожден α, если можно найти вероятность $\chi \in \mathcal{P}^p(C([0, T], \mathbb{T}^d))$ такую, что

1. $\chi = \mathrm{traj}^s_{m(\cdot)\#}\alpha$;
2. $m(t) = e_{t\#}\chi$;
3. $m(s) = m_*$.

Также можно воспользоваться подходом на основе дифференциальных включений. Если мы положим,

$$F(t, x, m) \triangleq \mathrm{co}\{f(t, x, m, u) : u \in U\},$$

то управляемую систему (36) можно записать следующим образом:

$$\frac{d}{dt}x(t) \in F(t, x(t), m(t)), \tag{38}$$

в то время как уравнение (37) станет дифференциальным включение типа среднего поля (ДВСП)

$$\frac{d}{dt}m(t) \in \langle F(t, \cdot, m(t)), \nabla \rangle m(t). \tag{39}$$

Как и уравнение (37) выше, это включение пока лишь формальная запись. Будем говорить, что поток вероятностей $[0, T] \ni t \mapsto m(t) \in \mathcal{P}^p(\mathbb{T}^d)$ является решением дифефренциального включения типа среднего поля (39), если существует вероятность $\chi \in \mathcal{P}^p(C([0, T]; \mathbb{T}^d))$ такая, что

1. χ-п.н. $x(\cdot)$ удовлетворяет (38);
2. $m(t) = e_{t\#}\chi$.

Отметим, что каждый поток вероятностей, порожденный s и α является решением (39). Обратно, если задан поток вероятностей $m(\cdot)$, который является решением (39) при $s \in [0, T]$ и $m_* = m(s)$, то можно найти управление $\alpha \in \mathcal{A}[m_*]$, порождающее поток $m(\cdot)$.

Рассмотрим теперь свойство выживаемости. Этот подход далее применяется при исследовании игр среднего поля первого порядка. Мы будем рассматривать вопросы выживаемости для случая $p = 1$ и однородной по времени динамики т.е. f не зависящего от t.

Определение 3.19 Definition 2 в статье [5]**.** Будем говорить, что множество $K \subset \mathcal{P}^p(\mathbb{T}^d)$ выживает относительно ДВСП (39) если для всех $m_0 \in K$ существуют $T > 0$ и решение ДВСП (39) на $[0, T]$ $m(\cdot)$ такие, что $m(0) = m_0$, и $m(t) \in K$ для всех $t \in [0, T]$.

Чтобы охарактеризовать свойство выживаемости, введем понятие касательного распределения.

Определение 3.20 [5, Definition 2]**.** Пусть $a > 0$, $m \in \mathcal{P}^1(\mathbb{T}^d)$. Будем говорить, что вероятность $\beta \in \mathcal{P}(\mathbb{T}^d \times \mathbb{R}^d)$ с маргинальным распределением на $\mathbb{T}^d$, равным m, является касательным распределением к множеству K в $m \in \mathcal{P}^1(\mathbb{T}^d)$ радиуса a, если существуют последовательности $\{\tau_n\}_{n=1}^{\infty} \subset (0, +\infty)$, $\{\beta_n\}_{n=1}^{\infty} \subset \mathcal{P}(\mathbb{T}^d \times \mathbb{R}^d)$ такие, что $\mathrm{p}^1{}_{\#}\beta = m$, $\mathrm{supp}(\beta_n) \subset \mathbb{T}^d \times B_a$ и

$$\frac{1}{\tau_n}\mathrm{dist}(\Theta^{\tau_n}{}_{\#}\beta_n, K) \to 0, \quad W_1(\beta_n, \beta) \to 0, \quad \tau_n \to 0 \text{ при } n \to \infty.$$

Обозначим множество касательных к K распределений в m радиуса a через $\mathcal{T}_K^a(m)$.

Далее пусть $\mathcal{F}(m)$ – множество вероятностей $\beta \in \mathcal{P}^1(\mathbb{T}^d \times \mathbb{R}^d)$ с маргинальным распределением на $\mathbb{T}^d$ равным m и таких, что

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} \mathrm{dist}(v, F(x, m))\beta(d(x, v)) = 0.$$

Теорема 3.21 [5, Theorem 1]**.** *Замкнутое множество $K \subset \mathcal{P}^1(\mathbb{T}^d)$ выживает относительно ДВСП (39) тогда и только тогда, когда существует константа $a > 0$ такая, что для всех $m \in K$*

$$\mathcal{T}_K^a(m) \cap \mathcal{F}(m) \neq \varnothing. \tag{40}$$

3.6 Игры среднего поля первого порядка

В рамках данного раздела предполагается, что фазовое пространство для каждого игрока есть $\mathbb{R}^d$. Мы предполагаем, что цель каждого игрока состоит в том, чтобы максимизировать величину

$$\gamma(x(T), m(T)) + \int_{t_0}^{T} g(t, x(T), m(t), u(t))dt \tag{41}$$

при том, что его динамика задается

$$\begin{gathered}\frac{d}{dt}x(t) = f(t, x(t), m(t), u(t)), \\ t \in [t_0, T], \quad x(t) \in \mathbb{R}^d, \quad m(t) \in \mathcal{P}^1(\mathbb{T}^d), \quad u(t) \in U,\end{gathered} \tag{42}$$

а $m(t)$ является распределением игроков в момент t. Удобно ввести дополнительную переменную и перейти к задаче с терминальным функционалом платы. Определим $z(\cdot)$ правилом:

$$\frac{d}{dt}z(t) = g(t, x(t), m(t), u(t)), \quad z(t_0) = 0. \tag{43}$$

Тогда можно считать, что каждый игрок стремится максимизирвать величину

$$\gamma(x(T), m(T)) + z(T). \quad (44)$$

Отметим, что $(x(t), z(t))$ лежит в расширенном фазовом пространстве $\mathbb{R}^d \times \mathbb{R}$.

Ниже мы уточним решение игры среднего поля применительно к случаю детерминированной динамики. Как и выше оператор проектирования $C([0,T], \mathbb{R}^d \times \mathbb{R})$ на $\mathbb{R}^d$ обозначим через e_t: для $w(\cdot) = (x(\cdot), z(\cdot))$ положим

$$e_t(w(\cdot)) = x(t).$$

Оператор проектирования принимающий значения в $\mathbb{R}^d \times \mathbb{R}$ обозначается через $\hat{e}_t$:

$$\hat{e}_t(w(\cdot)) = w(t).$$

Напомним, что обобщенные движения в силу (42), (43) удовлетворяют дифференциальному включению

$$(\dot{x}(t), \dot{z}(t)) \in F(t, x(t), m(t)), \quad (45)$$

где

$$F(t, x, m) \triangleq \mathrm{co}\{(f(t, x, m, u), g(t, x, m, u)) : u \in U\}. \quad (46)$$

Для заданного потока вероятностей $m(\cdot)$, $s, r \in [0, T]$, $s < r$, $y \in \mathbb{R}^d$, мы обозначим множество решений (45) на $[s, r]$ с начальным условием $x(s) = y$ через $\mathrm{Sol}(r, s, y, m(\cdot))$. Далее,

$$\mathrm{SOL}(r, s, m(\cdot)) \triangleq \bigcup_{y \in \mathbb{T}^d} \mathrm{Sol}(r, s, y, m(\cdot)).$$

Определение 3.22. Будем говорить, что пара $(\varphi, m(\cdot))$, где $\varphi : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}$ а $[t_0, T] \ni t \mapsto m(t) \in \mathcal{P}^p(\mathbb{R}^d)$ является (минимаксным) решением игры среднего поля (41), (42), если существует вероятность $\chi \in \mathcal{P}^p(C([t_0, T]; \times\mathbb{R}^d \times \mathbb{R}))$ такая, что

1. $m(t) = e_{t\#}\chi$;
2. $\varphi(s, y)$ – функция цены в задаче оптимального управления

$$\text{максимизировать } [\gamma(x(T), m(T)) + z(T) - z(s)]$$

$$\text{при } (x(\cdot), z(\cdot)) \in \mathrm{Sol}(T, s, y, m(\cdot));$$

3. $\mathrm{supp}(\chi) \subset \mathrm{SOL}(T, t_0, m(\cdot))$;
4. для любых $s, r \in [t_*, T]$, $s < r$ и $(x(\cdot), z(\cdot)) \in \mathrm{supp}(\chi)$

$$\varphi(s, x(s)) + z(s) = \varphi(r, x(r)) + z(r).$$

Прежде всего была получена теорема о существовании решения

Теорема 3.23 [6, Теорема 1]**.** *Для каждого $m_0 \in \mathcal{P}^1(\mathbb{T}^d)$, существует решение игры среднего поля (41), (42) такое, что $m(t_0) = m_0$.*

Замечание 3.24. В статье [6] эта теорема доказана при дополнительных предположениях о том, что $p = 1$ и вероятность m_0 сосредоточена на компактном множестве. Однако она легко может быть распространена на общий случай.

Предложенное определение решений игр среднего поля устойчиво по отношению к стохастическим возмущениям динамики. Для того чтобы это описать класс допустимых возмущений, мы рассматриваем последовательность стохастических игр среднего поля с динамикой, задаваемой генераторами типа Леви-Хинчина

$$\begin{aligned} L_t^n[m,u]\phi(x) = \frac{1}{2}\langle G^n(t,x,m,u)\nabla, \nabla\rangle\phi(x) + \langle f^n(t,x,m,u), \nabla\rangle\phi(x) \\ + \int_{\mathbb{R}^d} [\phi(x+y) - \varphi(x) - \langle y, \nabla\phi(x)\rangle \mathbf{1}_{B_1}(y)]\nu^n(t,x,m,u,dy). \end{aligned} \tag{47}$$

Предполагается, что каждый игрок стремится максимизировать выигрыш равный

$$\mathbb{E}\left[\gamma(X(T), m(T)) + \int_{t_0}^T g(t, X(T), m(t), u(t))dt\right],$$

где X – случайный процесс, порожденный генератором L^n. Введем следующие обозначения:

$$\Sigma^n(t,x,m,u) \triangleq \sum_{i=1}^d G_{ii}^n(t,x,m,u) + \int_{\mathbb{R}^d} \|y\|^2 \nu^n(t,x,m,u,dy),$$

$$b^n(t,x,m,u) \triangleq f^n(t,x,m,u) + \int_{\mathbb{R}^d \setminus B_1} y\nu^n(t,x,m,u,dy).$$

Мы предполагаем, что стохастические игры сходятся к исходной в следующем смысле:

$$\sup_{t\in[0,T], x\in\mathbb{R}^d, m\in\mathcal{P}^2(\mathbb{R}^d)} \frac{\Sigma^n(t,x,m,u)}{1 + \|x\|^2 + \varsigma^2(m)} \to 0 \text{ при } n \to \infty;$$

$$\sup_{t\in[0,T], x\in\mathbb{R}^d, m\in\mathcal{P}^2(\mathbb{R}^d)} \frac{\|b^n(t,x,m,u) - f(t,x,m,u)\|}{(1 + \|x\| + \varsigma(m))} \to 0 \text{ при } n \to \infty.$$

Здесь

$$\varsigma(m) \triangleq \left[\int_{\mathbb{R}^d} \|x\|^2 m(dx)\right]^{1/2}.$$

При некоторых дополнительных условиях регулярности имеет место следующее утверждение.

Теорема 3.25 [7, Theorem 1]**.** *Предположим, что для каждого n пара $(\varphi^n, m^n(\cdot))$ является решением игры среднего поля с динамикой (47) и функционалом платы (41).*

Тогда существует решение детерминированной игры среднего поля (41), (42) $(\varphi^*, m^*(\cdot))$ и последовательность $\{n_l\}_{l=1}^\infty$ такие, что

1. $\sup_{t\in[0,T]} W_2(m^{n_l}(t), \mu^*(t)) \to 0$ *при* $l \to \infty$;

2. $$\lim_{l\to\infty} \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\varphi^{n_l}(t,x) - \varphi^*(t,x)| = 0 \tag{48}$$

По заданному решению игры среднего поля для $p = 1$ было построено приближенное равновесие по Нэшу в игре конечного числа лиц, в которой игрок управляет движением группа агентов, двигающихся в силу динамики (42). Предполагается, что игрок стремится максимизировать общий выигрыш своих агентов. Если игрок i управляет группой агентов, стартующих из позиции $x^i_{N,0}$, $\mathbf{x} = (x^1_{N,0}, \ldots, x^N_{N,0})$, то обозначим

$$\delta^N_{\mathbf{x}} \triangleq \frac{1}{N}(\delta_{x^1_{N,0}} + \ldots + \delta_{x^N_{N,0}}).$$

По заданному решению игры среднего поля можно построить набор стратегий $\hat{\varrho}_N = (\hat{\varrho}^1_N, \ldots, \hat{\varrho}^N_N)$ где $\hat{\varrho}_N$ – вероятность на $\mathcal{U}$.

Теорема 3.26 [6, Theorem 3]**.** *Если*

$$W_1(m_0, \delta^N_{\mathbf{x}}) \to 0, \quad N \max_{i=\overline{1,N}} \int_{\mathbb{R}^d} \|x - x^i_{N,0}\| m^i_N(dx) \to 0 \text{ при } N \to \infty,$$

то для любого $\varepsilon > 0$ существует N_0 такое, что при всех $N > N_0$ набор стратегий $\hat{\varrho}_N$ является ε-равновесием по Нэшу.

Здесь $m^1_N, \ldots, m^N_N$ – меры на $\mathbb{R}^d$ такие, что $m_0 = m^1_N + \ldots + m^N_N$, $m^i_N(\mathbb{R}^d) = 1/N$ и

$$W_1(m_0, \delta^N_{\mathbf{x}}) = \sum_{i=1}^{N} \int_{\mathbb{R}^d} \|x - x^i_{N,0}\| m^i_N(dx).$$

3.7 Методы теории выживаемости для игр среднего поля первого порядка

Последняя часть диссертации посвящена исследованию зависимости решения игры среднего поля от начального распределения игроков. Отметим, что эта зависимость является объектом изучения в рамках подхода, основанного на основном уравнении игр среднего поля [CDLL19]. В рамках настоящего раздела предполагается, что фазовое пространство для каждого агента есть $\mathbb{T}^d \triangleq \mathbb{R}^d/\mathbb{Z}^d$. Также мы наделяем пространство вероятностей над $\mathbb{T}^d$ 1-метрикой Канторовича.

Основным объектом исследования в данном разделе является мультифункция цены.

Определение 3.27 [8, Definition 3.1]**.** Будем говорить, что многозначное отображение $\mathcal{V} : [0,T] \times \mathcal{P}^1(\mathbb{T}^d) \rightrightarrows C(\mathbb{T}^d)$ является мультифункцией цены для игры среднего поля (41), (42), если для всех $t_0 \in [0,T]$, $m_0 \in \mathcal{P}^1(\mathbb{T}^d)$ и $\phi \in \mathcal{V}(t_0, m_0)$ существует решение игры среднего поля (41), (42) $(\varphi, m(\cdot))$ такое, что

$$\varphi(t_0, \cdot) = \phi(\cdot), \quad m(t_0) = m_0. \tag{49}$$

Мультифункция цены может быть охарактеризована в терминах теории выживаемости. Для этого введем динамику на $C(\mathbb{T}^d) \times \mathcal{P}^1(\mathbb{T}^d)$, порожденную игрой среднего поля. Пусть $s, r \in [0,T]$, $s \leq r$, $m(\cdot) \in C([s,r]; \mathcal{P}^1(\mathbb{T}^d))$, определим оператор $B^{s,r}_{m(\cdot)} : C(\mathbb{T}^d) \to C(\mathbb{T}^d)$ по правилу:

$$(B^{s,r}_{m(\cdot)}\psi)(y) \triangleq \sup\{\psi(x(r)) + z(r) - z(s) : (x(\cdot), z(\cdot)) \in \mathrm{Sol}(r, s, y, m(\cdot))\}.$$

Также для $\phi \in C(\mathbb{T}^d)$, $\nu \in \mathcal{P}^1(\mathbb{T}^d \times \mathbb{R})$ через $[\phi, \nu]$ обозначим усреднение функции $\phi(x) + z$ по вероятности ν:

$$[\phi, \nu] \triangleq \int_{\mathbb{T}^d \times \mathbb{R}} (\phi(x) + z)\nu(d(x, z)).$$

Определение 3.28 [8, Definition 3.5]**.** Для $s, r \in [0, T]$, $s \leq r$ определим многозначное отображение $\Psi^{r,s} : \mathcal{P}^1(\mathbb{T}^d) \times C(\mathbb{T}^d) \rightrightarrows \mathcal{P}^1(\mathbb{T}^d) \times C(\mathbb{T}^d)$ по правилу: $(\mu, \psi) \in \Psi^{r,s}(m, \phi)$ тогда и только тогда, когда существует вероятность $\chi \in \mathcal{P}^1(C([s, r]; \mathbb{T}^d \times \mathbb{R}))$, удовлетворяющая следующим свойствам при $\nu(t) \triangleq \hat{e}_{t\#}\chi$, и $m(t) \triangleq e_{t\#}\chi$:

(Ψ1) $m(s) = m$, $m(r) = \mu$;

(Ψ2) $\phi = B^{s,r}_{m(\cdot)}\psi$;

(Ψ3) $[\psi, \nu(r)] \geq [\phi, \nu(s)]$.

Определение выживаемости дается стандартно.

Определение 3.29 [8, Definition 3.7]**.** Будем говорить, что многозначное отображение $\mathcal{V} : [0, T] \times \mathcal{P}^1(\mathbb{T}^d) \rightrightarrows C(\mathbb{T}^d)$ выживает относительно динамики, порожденной игрой среднего поля, если для любых $s, r \in [0, T]$, $s \leq r$, $m \in \mathcal{P}^1(\mathbb{T}^d)$, $\phi \in \mathcal{V}(s, m)$ существуют $\mu \in \mathbb{T}^d$ и $\psi \in C(\mathbb{T}^d)$ такие, что

- $(\mu, \psi) \in \Psi^{r,s}(m, \phi)$;
- $\psi \in \mathcal{V}(r, \mu)$.

Следующее утверждение устанавливает связь между мультифункцией цены и отображениями, выживающими относительно динамики, задаваемой игрой среднего поля.

Теорема 3.30 [8, Theorem 3.10]**.** *Предположим, что многозначное отображение $\mathcal{V} : [0, T] \times \mathbb{R}^d \rightrightarrows C(\mathbb{T}^d)$ выживает относительно динамики, задаваемой игрой среднего поля и $\mathcal{V}(T, m) = \{\gamma(\cdot, m)\}$. Тогда $\mathcal{V}$ является мультифункцией цены.*

Далее мы опишем инфинитезимальную форму условия выживаемости. Она использует производную в силу динамики игры среднего поля от многозначной функции $\mathcal{V}$. Прежде всего, несколько злоупотребляя обозначениями, для $m \in \mathcal{P}^1(\mathbb{T}^d)$ и $c > 0$ положим $\mathcal{L}^c(m)$ равным множеству вероятностей $\beta \in \mathcal{P}^1(\mathbb{T}^d \times \mathbb{R}^{d+1})$, сосредоточенных на $\mathbb{T}^d \times B_c \times [-c, c]$ с маргинальным распределением на $\mathbb{T}^d$ равным m. Для $s, r \in [0, T]$, $s \leq r$ через $A^{s,r}_m$ обозначим оператор на $C(\mathbb{T}^d)$, действующий по правилу

$$(A^{s,r}_m \phi)(x) \triangleq \sup\{\phi(x + (r - s)a) + (r - s)b : (a, b) \in F(s, x, m)\}. \tag{50}$$

Здесь F определяется равенством (46). Для $\tau \geq 0$ оператор сдвига $\Theta^\tau : \mathbb{T}^d \times \mathbb{R}^{d+1} \to \mathbb{T}^d \times \mathbb{R}$ зададим формулой

$$\Theta^\tau(x, a, b) \triangleq (x + \tau a, \tau b). \tag{51}$$

Определение 3.31 [8, Definition 4.1]**.** Вероятность $\beta \in \mathcal{L}^c(m)$ принадлежит $\mathcal{D}^c_F\mathcal{V}(t,m,\phi)$, если существуют последовательности $\{\tau_n\}_{n=1}^\infty \subset (0,+\infty)$, $\{\beta_n\}_{n=1}^\infty \subset \mathcal{L}^c(m)$ и $\{\phi_n\}_{n=1}^\infty \subset C(\mathbb{T}^d)$, удовлетворяющие следующим свойствам при $\nu_n \triangleq \Theta^{\tau_n}{}_\#\beta_n$ и $m_n \triangleq \mathrm{p}_\#\nu_n$:

1. $\tau_n, W_1(\beta,\beta_n) \to 0$ при $n \to \infty$;
2. $\phi_n \in \mathcal{V}(t+\tau_n, m_n)$;
3. $$\lim_{n\to\infty} \frac{\|A_m^{t,t+\tau_n}\phi_n - \phi\|}{\tau_n} = 0;$$
4. $$\lim_{n\to\infty} \frac{[\phi_n,\nu_n] - [\phi,\widehat{m}]}{\tau_n} \geq 0;$$
5. $$\int_{\mathbb{T}^d\times\mathbb{R}^{d+1}} \mathrm{dist}(v; F(t,x,m))\beta(d(x,v)) = 0.$$

Здесь $\widehat{m}$ –вероятность на $\mathbb{T}^d\times\mathbb{R}$, определенная по правилу: для $\phi \in C(\mathbb{T}^d\times\mathbb{R})$

$$\int_{\mathbb{T}^d\times\mathbb{R}} \phi(x,z)\widehat{m}(d(x,x)) = \int_{\mathbb{T}^d} \phi(x,0)m(dx).$$

Пусть $M, C > 0$, через $\mathrm{BL}_{M,C}$ обозначим множество функций $\phi \in \mathrm{Lip}_C(\mathbb{T}^d)$ таких, что $\|\phi\| \leq M$.

Теорема 3.32 [8, Theorem 4.2]**.** *Предположим, что полунепрерывное многозначное отображение $\mathcal{V} : [0,T]\times\mathcal{P}^1(\mathbb{T}^d) \rightrightarrows C(\mathbb{T}^d)$ непустозначно и существуют константы M и C такие, что для всех $t \in [0,T]$, $m \in \mathcal{P}^1(\mathbb{T}^d)$,*

$$\mathcal{V}(t,m) \subset \mathrm{BL}_{M,C}(\mathbb{T}^d).$$

Тогда $\mathcal{V}$ выживает относительно динамики, задаваемой игрой среднего поля тогда и только тогда, когда существует константа $c > 0$ такая, что для любых $t \in [0,T]$, $m \in \mathcal{P}^1(\mathbb{T}^d)$, $\phi \in \mathcal{V}(t,m)$

$$\mathcal{D}^c_F\mathcal{V}(t,m,\phi) \neq \varnothing.$$

Эта теорема влечет инфинитезимальное условие на мультифункцию цены.

Следствие 3.33 [8, Corollary 4.3]**.** *Пусть полунепрерывное сверху по включению многозначное отображение $\mathcal{V} : [0,T]\times\mathcal{P}^1(\mathbb{T}^d) \rightrightarrows C(\mathbb{T}^d)$ непустозначно. Предположим, также, что для всех $t \in [0,T]$, $m \in \mathcal{P}^1(\mathbb{T}^d)$, $\phi \in \mathcal{V}(t,m)$*

- $\mathcal{V}(t,m) \subset \mathrm{BL}_{M,C}(\mathbb{T}^d)$*, где константы M и C не зависят от t и m;*
- $\mathcal{V}(T,m) = \{\gamma(\cdot,m)\}$*;*
- $\mathcal{D}^c_F\mathcal{V}(t,m,\phi) \neq \varnothing$*, где константа c не зависит от t, m и ϕ.*

Тогда $\mathcal{V}$ является мультифункцией цены для игры среднего поля (41), (42).

Список литературы

[AD01] N.U. Ahmed and X. Ding. Controlled McKean-Vlasov equation. *Commun. Appl. Anal.*, 5:183–206, 2001.

[AGS05] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows: in metric spaces and in the space of probability measures.* Lectures in Mathematics. ETH Zurich. Birkhäuser, Basel, 2005.

[Aub09] J.-P. Aubin. *Viability theory.* Birkhäuser, Boston, 2009.

[Ave15] Yu. Averboukh. Universal Nash equilibrium strategies for differential games. *J. Dyn. Control Syst.*, 21:329–350, 2015.

[BD96] M. Bardi and I. Capuzzo Dolcetta. *Optimal control and viscosity solutions of Hamilton–Jacobi–Bellman equations.* Birkhäuser, Basel, 1996.

[BK12] V.I. Bogachev and A. V. Kolesnikov. The Monge-Kantorovich problem: achievements, connections, and perspectives. *Russ. Math. Surv+*, 67(5):785–890, 2012.

[BS71] N. Barabanova and A. Subbotin. On continuous evasion strategies in game problems on the encounter of motions. *J. Appl. Math. Mech.*, 34(5):765–772, 1971.

[BS04a] A. Bressan and W. Shen. Semi-cooperative strategies for differential games. *Internat. J. Game Theory*, 32:561–59, 2004.

[BS04b] A. Bressan and W. Shen. Small BV solutions of hyperbolic noncooperative differential games. *SIAM J. Control Optim.*, 43:194–215, 2004.

[CD18] Rene Carmona and Francois Delarue. *Probabilistic theory of mean field games with applications. Vol. I, II.* Springer, 2018.

[CDLL19] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions. *The master equation and the convergence problem in mean field games.* Princeton University Press, Princeton, 2019.

[CLSW98] F. H. Clarke, Y. S. Ledyaev, R. Stern, and P. R. Wolenski. *Nonsmooth analysis and control theory.* Springer, Heidelberg, 1998.

[CP03] P. Cardaliaguet and S. Plaskacz. Existence and uniqueness of a Nash equilibrium feedback for a simple nonzero-sum differential game. *Internat. J. Game Theory*, 32:33–71, 2003.

[Fis17] M. Fischer. On the connection between symmetric N-player games and mean field games. *Ann. Appl. Probab.*, 27:757–810, 2017.

[Fri70] A. Friedman. Existence of value and of saddle points for differential games of pursuit and evasion. *J. Differential Equations*, 7(1):92–110, 1970.

[FS06] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions.* Springer, New York, 2006.

[GHL03] X. Guo and O. Hernández-Lerma. Zero-sum games for continuous-time Markov chains with unbounded transitions and average payoff rates. *J. Appl. Probab.*, 40:327–345, 2003.

[GS15] W. Gangbo and A. Święch. Existence of a solution to an equation arising from the theory of mean field games. *J. Differential Equations*, 259:6573–6643, 2015.

[HL95] S. Hamadèène and J. P. Lepeltier. Backward equations, stochastic control and zero-sum stochastic differential games. *Stoch. Stoch. Rep.*, 54(3-4):221–231, 1995.

[HM19] S. Hamadène and P. Mannucci. Regularity of Nash payoffs of Markovian nonzero-sum stochastic differential games. *Stochastics*, 91(5):695–715, 2019.

[HMC05] M. Huang, R.P. Malhamé, and P.E. Caines. Nash equilibria for large population linear stochastic systems with weakly coupled agents. In E. K. Boukas and R.P. Malhamé, editors, *Analysis, control and optimization of complex dynamic systems*, pages 215–252. Springer, 2005.

[KK10] N. N. Krasovskii and A. N. Kotelnikova. An approach-evasion differential game: stochastic guide. *Proc. Steklov Inst. Math.*, 269(1 Supplement):S191–S213, 2010.

[Kle93] A.F. Kleimenov. *Nonzero-sum differential games.* Nauka, Ekaterinburg, 1993. in Russian.

[KLY11] V.N. Kolokoltsov, J.J. Li, and W. Yang. Mean field games and nonlinear Markov processes. Preprint at arXiv:1112.3744v2, 2011.

[KS70] N.N. Krasovskii and A.I. Subbotin. An alternative for the game problem of convergence. *J. Appl. Math. Mech.*, 34:948–965, 1970.

[KS88] N. N. Krasovskii and A. I. Subbotin. *Game-theoretical control problems.* Springer, New York, 1988.

[KTY14] V. N. Kolokoltsov, M. Troeva, and W. Yang. On the rate of convergence for the mean-field approximation of controlled diffusions with large number of players. *Dyn. Games Appl.*, 4:208–230, 2014.

[KY13] V. Kolokoltsov and W. Yang. Existence of solutions to path-dependent kinetic equations and related forward-backward systems. *Open Journal of Optimization*, 2:39–44, 2013.

[Lac17] D. Lacker. Limit theory for controlled McKean-Vlasov dynamics. *SIAM J. Control Optim.*, 55:1641–1672, 2017.

[Lev13] Y. Levy. Continuous-time stochastic games of fixed duration. *Dyn. Games Appl.*, 3:279–312, 2013.

[LL06a] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. I. Le cas stationnaire (French) [Mean field games. I. The stationary case]. *C. R. Math. Acad. Sci. Paris*, 343:619–625, 2006.

[LL06b] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. II. Horizon fini et contrôle optimal (French) [Mean field games. II. Finite horizon and optimal control]. *C. R. Math. Acad. Sci. Paris*, 343:679–684, 2006.

[Man04] P. Mannucci. Nonzero-sum stochastic differential games with discontinuous feedback. *SIAM J. Control Optim.*, 43:1222–1233, 2004.

[Man14] P. Mannucci. Nash p for nonzero-sum stochastic differential games with separate Hamiltonians. *Dyn. Games Appl.*, 4:329–344, 2014.

[SC81] A. I. Subbotin and A. G. Chentsov. *Optimization of guarantee in control problems*. Nauka, Moscow, 1981. in Russian.

[Sub95] A. I. Subbotin. *Generalized solutions of first-order PDEs. The dynamical perspective*. Birkhauser, Boston, 1995.

[YZ99] J. Yong and X. Yu. Zhou. *Stochastic controls. Hamiltonian systems and HJB equations*. Springer, New York, 1999.

[Zac64] L. E. Zachrisson. Markov games. In M. Dresher, L. S. Shapley, and A. W. Tucker, editors, *Advances in game theory*, pages 211–253. Princeton University Press, Princeton, 1964.