

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

На правах рукописи

Петросян Ованес Леонович

**ДИНАМИЧЕСКОЕ И НЕПРЕРЫВНОЕ
ОБНОВЛЕНИЕ ИНФОРМАЦИИ В МОДЕЛЯХ
КОНФЛИКТНОГО УПРАВЛЕНИЯ**

Научная специальность 2.3.1.

Системный анализ, управление и обработка информации

Диссертация на соискание ученой степени
доктора физико-математических наук

Санкт-Петербург – 2021

Оглавление

ВВЕДЕНИЕ	7
Глава 1 ПРД-ядро: новое кооперативное решение для дифференциальных игр	28
1.1 Постановка задачи и предварительные сведения	29
1.1.1 Модель дифференциальной игры	29
1.1.2 Модель кооперативной дифференциальной игры	31
1.1.3 С-ядро	34
1.1.4 Непустота С-ядра в статических играх	35
1.1.5 Динамическая устойчивость кооперативного решения и процедура распределения дележа	36
1.2 ПРД-ядро и доминирование	40
1.2.1 Доминирование по процедурам распределения дележа	40
1.2.2 Построение ПРД-ядра	41
1.2.3 Свойства ПРД-ядра	47
1.3 Непустота ПРД-ядра на основе подхода линейного программирования	50
1.4 Дифференциальная игровая модель добычи ресурсов	51
1.4.1 Кооперативные стратегии и кооперативная траектория	52
1.4.2 Характеристическая функция	54
1.4.3 ПРД-ядро	57
1.4.4 Непустота ПРД-ядра	57
1.4.5 С-ядро и ПРД-ядро	58
Глава 2 Дискретные игры с динамическим обновлением информации	61
2.1 Модель дискретной игры с динамическим обновлением информации	61
2.1.1 Исходная дискретная игра	61
2.1.2 Усеченная подыгра	62

2.2	Некооперативная дискретная игра с динамическим обновлением информации	63
2.3	Кооперативная дискретная игра с динамическим обновлением информации	65
2.3.1	Результирующее кооперативное решение и соответствующие теоремы	68
2.3.2	Свойства результирующего кооперативного решения	71
2.4	Случайный информационный горизонт	78
2.5	Динамическая игровая модель олигополии с рекламой	81
2.5.1	Исходная игра	81
2.5.2	Некооперативные исходы в усеченной подыгре	83
2.5.3	Кооперативные исходы в усеченной подыгре	89
2.5.4	Характеристическая функция в усеченной подыгре	94
2.5.5	Численное моделирование динамической игровой модели олигополии с рекламой и обновлением информации	95
Глава 3 Дифференциальные игры с динамическим обновлением информации		106
3.1	Общий класс кооперативных дифференциальных игр с динамическим обновлением информации	107
3.1.1	Исходная дифференциальная игра	107
3.1.2	Усеченная подыгра	108
3.1.3	Концепция кооперативного решения с динамическим обновлением информации	112
3.1.4	Построение характеристической функции в игре с динамическим обновлением информации	116
3.1.5	Связь решения в усеченных подыграх и результирующего решения	120
3.1.6	Кооперативная игра добычи ограниченного ресурса с динамическим обновлением информации	130
3.2	Кооперативные дифференциальные игры с динамическим обновлением информации и стохастическим прогнозом	139
3.2.1	Исходная игра	139
3.2.2	Комбинированная усеченная подыгра	141
3.2.3	Концепция комбинированного кооперативного решения	146

3.2.4	Кооперативная дифференциальная игра добычи ограниченного ресурса с динамическим обновлением информации и стохастическим прогнозом	149
3.3	Кооперативные дифференциальные игры с динамическим обновлением информации и случайным горизонтом	157
3.3.1	Усеченная подыгра со случайной продолжительностью	157
3.3.2	Концепция кооперативного решения для игры с динамическим обновлением и случайным горизонтом	166
3.3.3	Кооперативная игра добычи ограниченного ресурса с динамическим обновлением информации и случайным горизонтом	170
3.4	Дифференциальная игра нефтяного рынка с динамическим обновлением информации	184
3.4.1	Введение	184
3.4.2	Модель некооперативной игры	189
3.4.3	Частично кооперативная игра	196
3.4.4	Соглашение о сотрудничестве	206
Глава 4 Дифференциальные игры с непрерывным обновлением информации		215
4.1	Модель дифференциальной игры с непрерывным обновлением информации	216
4.1.1	Исходная дифференциальная игра	216
4.1.2	Игра с непрерывным обновлением информации	216
4.2	Модель некооперативной дифференциальной игры с непрерывным обновлением информации	219
4.2.1	Равновесие по Нэшу в игре с непрерывным обновлением информации	219
4.2.2	Уравнения Гамильтона–Якоби–Беллмана с непрерывным обновлением информации	221
4.2.3	Принцип максимума Понтрягина с непрерывным обновлением информации	228
4.2.4	Автономный случай линейно-квадратичной игры с непрерывным обновлением информации	237

4.2.5	Неавтономный случай линейно-квадратичной игры с непрерывным обновлением информации	252
4.3	Модель кооперативной дифференциальной игры с непрерывным обновлением информации	280
4.3.1	Кооперативная игра с непрерывным обновлением информации	281
4.3.2	Уравнения Гамильтона–Якоби–Беллмана для кооперативной игры с непрерывным обновлением информации	294
4.3.3	Принцип максимума Понтрягина для кооперативной дифференциальной игры с непрерывным обновлением информации	305
4.3.4	Линейно-квадратичный случай кооперативной дифференциальной игры с непрерывным обновлением информации	328
4.4	Класс дифференциальных игр с нетрансферабельной полезностью и непрерывным обновлением информации	344
4.4.1	Парето-оптимальные стратегии с непрерывным обновлением информации	344
4.4.2	Уравнение Гамильтона–Якоби–Беллмана с непрерывным обновлением информации	345
4.4.3	Временная состоятельность в игре с непрерывным обновлением информации	347
4.5	Дифференциальная игра добычи ресурса с непрерывным и динамическим обновлением информации	356
4.5.1	Исходная игра добычи ресурса	356
4.5.2	Игра с динамическим обновлением информации	361
4.5.3	Игра с непрерывным обновлением информации	365
4.5.4	Численное моделирование	373
Глава 5 Обратная задача теории управления с непрерывным обновлением информации		376
5.1	Оптимальное управление с непрерывным обновлением информации	377
5.1.1	Исходная задача оптимального управления	377
5.1.2	Задача оптимального управления с непрерывным обновлением информации	377

5.1.3	Оптимальное управление с непрерывным обновлением информации	379
5.2	Линейно-квадратичное оптимальное управление с непрерывным обновлением информации	380
5.2.1	Постановка задачи для линейно-квадратичного случая с непрерывным обновлением информации	380
5.2.2	Оптимальное управление с непрерывным обновлением информации	381
5.3	Обратная задача оптимального управления с непрерывным обновлением информации	383
5.4	Численное моделирование	384
5.4.1	Модель рулевого управления одноколейным ТС	385
5.4.2	Оптимальное управление с непрерывным обновлением информации	387
5.4.3	Обратная задача оптимального управления	389
5.4.4	Обсуждение	390
5.5	Заключение	391
	Заключение	393
	Список литературы	396

ВВЕДЕНИЕ

Актуальность темы диссертации

Основными задачами современной теории игр являются конструирование и анализ принципов оптимального поведения участников в различных задачах конфликтного управления. Реально происходящие конфликты развиваются во времени, поэтому особую актуальность приобретают динамические модели. Дифференциальные игры являются удобными математическими моделями для описания конфликтно-управляемых процессов, происходящих в экономике, экологии, менеджменте и других сферах человеческой деятельности. Теория дифференциальных игр выделилась в отдельный раздел математики в пятидесятых годах XX в. Одной из первых работ в области дифференциальных игр принято считать работу Р. Айзекса [1], в которой в терминах состояний и управлений была сформулирована задача перехвата самолета управляемой ракетой, а также выведено основополагающее уравнение для нахождения решения. Вклад Р. Айзекса вместе с классическим исследованием Р. Беллмана [2] создали основу для использования результатов теории оптимального управления в задачах конфликтного управления с несколькими участниками. Первые интересные результаты в теории дифференциальных игр были получены Л. Берковицем [3], Г. Лейтманом [4], В. Флемингом [5], А. Фридманом [6] и др. Значительный вклад в дальнейшее развитие дифференциальных игр внесли отечественные ученые Л.С. Понтрягин [7; 8], Л.А. Петросян [9; 10], Н.Н. Красовский [11; 12], Б.Н. Пшеничный [13], работы которых в основном были связаны с дифференциальными играми преследования. Важнейшие результаты в области обоснования и методов нахождения решений антагонистических дифференциальных игр были получены в работах Красовского Н. Н. и А.И. Субботина [14; 15; 16]. Параллельно начала развиваться теория неантагонистических дифференциальных игр, в которых в качестве принципа оптимальности использовалось равновесие по Нэшу [17]. Особо следует отметить работы отечественных ученых, внесших

большой вклад в развитие неантагонистических дифференциальных игр: Э. М. Вайсборда, Р.В. Гамкрелидзе, Н. Л. Григоренко, В.И. Жуковского, А. Ф. Клейменова, А. Ф. Кононенко, А.В. Кряжимского, А. Б. Куржанского, В.Н. Лагунова, Н. Ю. Лукоянова, С.С. Кумкова, О. А. Малафеева, Мищенко Е.Ф., В.С. Пацко, Н. Н. Петрова, Субботиной Н.Н. , Тынянского Н.Т., Чикрия А.А., Чистякова С. В., Ченцова А. Г. [18; 19; 20; 21; 22; 23; 24; 25; 26; 27; 28; 29; 30; 31; 32; 33; 34; 35; 36; 37; 38; 39; 40; 41; 42; 43; 44; 45; 46] и многих других. Позднее работы, использующие методы дифференциальных игр, появились и в области моделирования конфликтно– управляемых экономических процессов, в том числе в задачах природоохранной политики, оптимальной эксплуатации природных ресурсов и пр. (см., например, [47; 48; 49]). Данная область развивается достаточно быстро, подробный анализ указанных работ можно найти в [50] (см. также [51]). Особо отметим работы Л.А. Петросяна, В.В. Захарова, Н.А. Зенкевича, В.В. Мазалова, А.Н. Реттеновой, С. Йоргенсена, Е. Докнера, Н. Лонга, Г. Соргера, Дж. Заккура, Я. Кравчика, Дж. Филара и др. [47; 52; 53; 54; 55; 56; 57; 58; 59; 60; 61; 62; 63; 64; 65; 66; 67; 68; 69; 70; 71; 72; 73; 74; 75], посвященные использованию теоретико-игрового подхода для решения проблемы охраны окружающей среды.

Большинство реальных конфликтных процессов управления развиваются динамически или непрерывно во времени, а их участники непрерывно получают обновленную информацию и адаптируются к ней. Для таких процессов предложен подход, позволяющий строить более реалистичные модели, а именно, подход с динамическим или непрерывным обновлением информации. Теория игр является классическим и фундаментальным инструментом, который можно использовать для моделирования поведения конфликтных процессов с несколькими участниками. В частности, учитывая динамический характер процессов, развивающихся во времени, это теория динамических и дифференциальных игр.

Классические динамические и дифференциальные игры основаны на предположении о том, что структура игры не изменяется на временном интервале определения игры, или что игроки имеют полную информацию об изменении структуры игры. Однако, при рассмотрении долгосрочных процессов эти предположения не соответствуют действительности. Чтобы моделировать поведение игроков в сценарии, близком к реальному, необходимо учитывать специфиче-

ское поведение игрока в следующем смысле: как правило, даже если игроки имеют долгосрочную информацию о процессе, они будут лучше использовать краткосрочную информацию для принятия решения о будущих действиях. В классе игр с динамическим и непрерывным обновлением информации предполагается, что игроки имеют или используют только информацию об уравнениях движения и функциях выигрыша, определенных на конечном временном интервале (информационном горизонте). Информационный горизонт определяет, насколько далеко игроки могут или хотят прогнозировать свои действия. Информация об уравнениях движения и функциях выигрыша обновляется по мере изменения текущего времени. Чтобы определить наилучшее возможное поведение игроков в этом типе динамической или дифференциальной игры, необходимо использовать специальный подход, который является предметом изучения в данной работе. Этот подход получил развитие в статьях автора диссертации и его соавторов.

Большинство реальных конфликтных процессов управления динамично или непрерывно развиваются во времени, а их участники постоянно получают обновленную информацию и адаптируются. Для таких процессов предложен подход, позволяющий строить более реалистичные модели, а именно, игры с динамическим обновлением информации [76; 77] и игры с непрерывным обновлением информации [78; 79]. Фундаментальные модели, ранее рассмотренные в теории дифференциальных игр, относятся к задачам на фиксированном временном интервале (у игроков есть вся информация для конечного временного отрезка) [26], задачам на бесконечном временном интервале с дисконтированием (у игроков есть информация для бесконечного временного интервала) [80] и задачам на случайном временном интервале (игроки имеют информацию для данного временного интервала, но конечный момент является случайной величиной) [81]. Более того, одна из первых работ по теории дифференциальных игр посвящена игре преследователя и убегающего (выигрыш игрока зависит от того, когда соперник пойман) [82]. Во всех вышеупомянутых моделях и предлагаемых решениях считается, что игроки в начале игры имеют всю информацию о динамике игры (уравнения движения) и о предпочтениях игроков (функции выигрыша). Однако этот подход не принимает во внимание тот факт, что во многих реальных процессах игроки в начальный момент не имеют всей информации об игре. Таким образом, существующие подходы

не могут быть напрямую использованы для построения достаточно большого набора теоретико-игровых моделей, соответствующих реальности.

В игровых моделях с динамическим обновлением информации предполагается, что игроки:

- 1 имеют информацию об уравнениях движения и функциях выигрыша на усеченном временном интервале длиной \bar{T} , который называется информационным горизонтом;
- 2 получают обновленную информацию об уравнениях движения и функциях выигрыша в фиксированные моменты времени $t_j = t_0 + j\Delta t$, $j = 0, \dots, l$, $l = \frac{T-t_0}{\Delta t}$ и, как следствие, динамически адаптируются к обновленной информации.

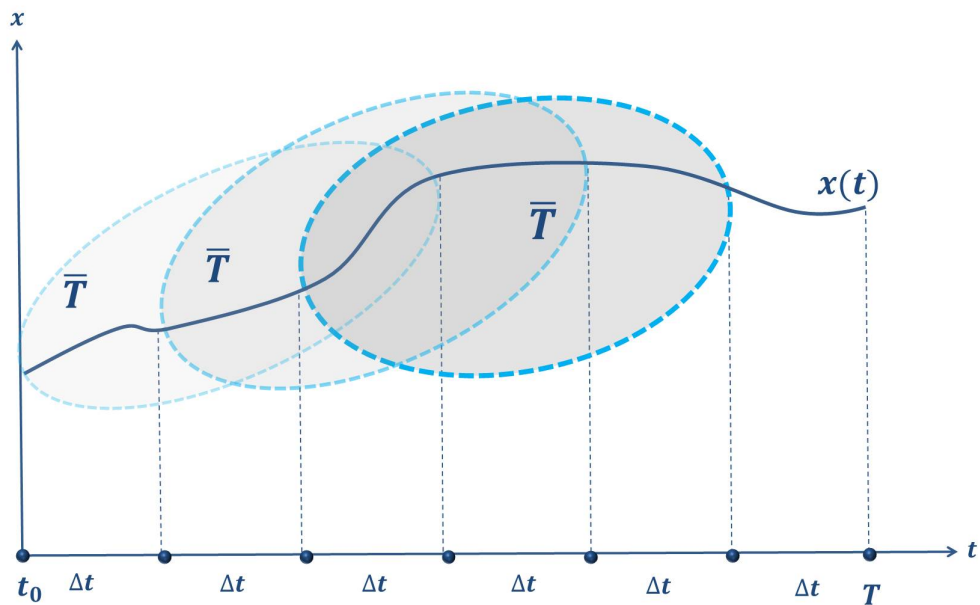


Рис. 1: Каждый синий овал показывает информацию, доступную игрокам на интервале $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, а именно, $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$, $l = \frac{T-t_0}{\Delta t}$.

В игровых моделях с непрерывным обновлением информации предполагается, что игроки:

- 1 имеют информацию об уравнениях движения и функциях выигрыша на усеченном временном интервале длиной \bar{T} , который называется информационным горизонтом;

2 постоянно получают обновленную информацию об уравнениях движения и функциях выигрыша и, как следствие, непрерывно адаптируются к обновленной информации.

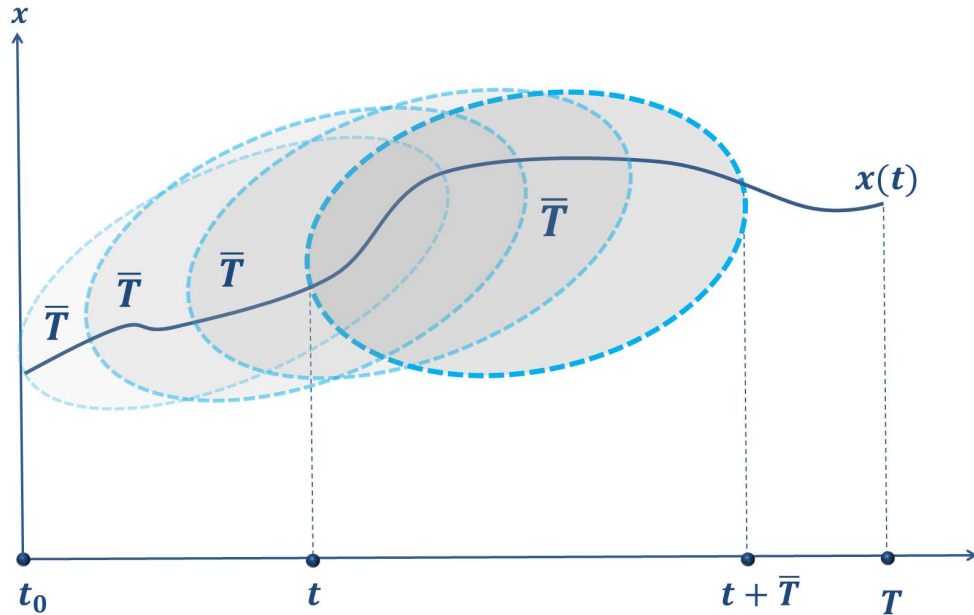


Рис. 2: Каждый синий овал показывает информацию, доступную игрокам в момент времени t , а именно, $[t, t + \bar{T}]$, где \bar{T} — информационный горизонт.

Очевидно, что получить равновесные по Нэшу или кооперативные стратегии сложно из-за отсутствия фундаментальных подходов к дифференциальным играм и задачам управления с динамическим и непрерывным обновлением информации. Классические методы, такие как динамическое программирование и уравнение Гамильтона–Якоби–Беллмана [2] или принцип максимума Понтрягина [83], не позволяют напрямую строить равновесные по Нэшу или кооперативные стратегии в задачах с обновлением информации. При описанных выше предположениях, возникают две основные проблемы:

- 1 Как определить концепцию решения, подобную равновесным по Нэшу, кооперативным и оптимальным по Парето стратегиям, характеристическую функцию и кооперативное решение для класса игр с непрерывным обновлением информации?
- 2 Как вывести соответствующие условия оптимальности для равновесных по Нэшу, кооперативных и оптимальных по Парето стратегий и характеристической функции?

Помимо проблемы построения оптимальных в некотором смысле стратегий, соответствующей траектории для класса игр с динамическим и непрерывным обновлением информации, существуют и проблемы, связанные с исследованием свойств динамической устойчивости и временной состоятельности кооперативного решения.

Отметим, что существует широкий спектр реальных конфликтных процессов управления, которые можно моделировать с помощью подхода с динамическим и непрерывным обновлением информации. Поэтому необходимо развивать этот подход для класса кооперативных и некооперативных игр, для частных случаев линейно-квадратичных игр, а также конкретных динамических и дифференциальных игровых моделей, чтобы показать значимость предлагаемого подхода. Также важно применять предложенный подход в задачах управления техническими, экологическими и социально-экономическими системами, чтобы сделать его ближе к практическим приложениям. Всё вышесказанное обосновывает актуальность темы диссертации.

Степень разработанности проблемы в литературе

Класс дифференциальных игр с динамическим и непрерывным обновлением информации имеет некоторое сходство с теорией управления с прогнозирующими моделями (Model Predictive Control, MPC), которая разрабатывается в рамках численного оптимального управления, см. [84; 85; 86; 87]. В подходе MPC текущее управляющее воздействие получается при решении задачи оптимального управления без обратной связи с конечным горизонтом в каждый момент выборки. Для линейных систем существует решение в явном виде [88; 89]. Однако в целом подход MPC требует решения нескольких задач оптимизации. В серии близких работ [90; 91; 92; 93] рассмотрен класс стабилизирующих управлений, и аналогичные подходы использованы для линейно-квадратичных задач оптимального управления. Однако в данной диссертации и статьях, посвященных подходу с непрерывным обновлением информации, основная цель иная — моделировать поведение игроков, когда информация об игре постоянно обновляется с течением времени.

Статьи [94; 95; 96; 97; 98; 99] более тесно связаны с подходом с непрерывным обновлением информации. В частности, статья [96] посвящена подходу движущегося горизонта для динамических игровых моделей. Новая концепция

решения, основанная на управлении движущимся горизонтом, введена для дифференциальных игр с ненулевой суммой и бесконечным горизонтом, а также рассмотрен линейно-квадратичный случай с программными и позиционными стратегиями. Другая статья [94] посвящена концепции равновесия для динамических игр как с дискретным, так и с непрерывным временем и применяющихся (симметричных и асимметричных) режимах игры. В статье [95] рассмотрена изменяющаяся во времени макроэкономическая модель, в которой некоторые параметры могут колебаться экзогенным образом в соответствии с цепью Маркова. В статье [97] разработаны теоретические основы общепринятой деловой практики принятия решений на скользящем горизонте. Основная идея представленного подхода заключается в следующем: полезность методов скользящего горизонта в значительной степени определяется тем фактом, что прогнозирование будущего является затратным. Подход, аналогичный непрерывному обновлению информации, исследован в статьях [98] и [99], посвященных повторяющимся играм со скользящими горизонтами планирования.

Содержание работы

Диссертация посвящена разработке теоретических моделей конфликтного управления (игр с динамическим и непрерывным обновлением информации). Первая глава посвящена построению и исследованию нового кооперативного решения ПРД-ядра для дифференциальных игр с динамическим и непрерывным обновлением информации. ПРД-ядро определяется на основе аксиом динамической устойчивости и недоминирования по ПРД. Доказывается сильная динамическая устойчивость этого решения. Приведена явная формула для построения ПРД-ядра. Представлен численный алгоритм для анализа свойства непустоты. Главы 2 и 3 посвящены применению подхода с динамическим обновлением информации к динамическим и дифференциальным играм. Глава 4 посвящена исследованию и применению подхода с непрерывным обновлением информации к классу дифференциальных игр. В целом в главах 2–4 рассматриваются как кооперативные, так и некооперативные постановки. Здесь впервые представлен широкий спектр условий оптимальности для кооперативных и равновесных по Нэшу стратегий как в программных, так и в позиционных стратегиях. Приведены методики построения соответствующих кооперативной и равновесной траектории с динамическим и непрерывным обновлением

информации. Кроме того, для модели кооперативной игры представлен алгоритм построения характеристической функции и кооперативного решения с динамическим и непрерывным обновлением информации. Доказана связь между кооперативными решениями, заданными на усеченных интервалах, определяемых информационным горизонтом. Другие результаты касаются связи игр с динамическим и непрерывным обновлением информации. Представлено доказательство сходимости стратегий и траекторий с динамическим обновлением информации к соответствующим стратегиям и траекториям с непрерывным обновлением информации. Пятая глава посвящена обратной задаче оптимального управления с непрерывным обновлением информации. В качестве приложения подход с непрерывным обновлением информации используется для обратной задачи оптимального управления системы помощи водителю.

Первая глава. В статье [100] (и позже в [101]) автор диссертации и его соавторы ввели понятие сильного динамически устойчивого подмножества S -ядра. Авторы построили новое кооперативное решение, используя геометрический подход, и доказали, что оно является подмножеством S -ядра и обладает сильной динамической устойчивостью. Позже это решение было названо ПРД-ядром, и оно может быть построено с использованием системы линейных ограничений для процедур распределения дележа. Эти условия определяются для каждого момента времени дифференциальной игры. Из непустоты множества, описываемого этими ограничениями и непустоты соответствующего множества ПРД в каждый момент времени следует, что ПРД-ядро не пусто. В статье [102] методика, предложенная в [103], применена для исследования непустоты ПРД-ядра для каждого момента времени. Если для каждого момента времени оно непусто, можно заключить, что ПРД-ядро не пусто. Полученные результаты могут быть использованы для построения ПРД-ядра и проверки его непустоты в качестве числового примера. Кроме того, частный случай этого подхода представлен для дифференциальных игр 3-х лиц. Можно аналитически построить условия непустоты ПРД-ядра в зависимости от характеристической функции. Кроме того, можно определить аналитическую формулу для селекторов ПРД-ядра, в частности, формулу для процедур распределения дележа для селекторов ПРД-ядра. Позже в статье [104] ПРД-ядро изучено как кооперативное решение для дифференциальных игр, построенных

исключительно с использованием аксиом динамической устойчивости и доминирования по ПРД. В работах выше описан довольно новый подход, который позволяет строить динамически устойчивые кооперативные решения. Этот подход использует свойство динамической устойчивости в качестве основного аксиоматического свойства для определения кооперативного решения. Этот подход является предметом исследования в работе [104]. Стоит отметить, что использование свойства динамической устойчивости для динамических кооперативных игр, теории социального выбора и теории дизайна механизмов в качестве аксиомы является многообещающим. Еще одно важное свойство, рассматриваемое в работе [104], называется доминированием по процедуре распределения дележа (ПРД). Согласно этому свойству, соответствующее кооперативное решение строится с использованием процедур распределения дележа, которые являются недоминируемыми. Будем говорить, что исходная ПРД является недоминируемой, если не существует другой ПРД, коалиции S и момента времени, таких что мгновенные выплаты, соответствующие этой ПРД, выше для игроков из коалиции S в данный момент времени, чем их мгновенные выплаты, соответствующие исходной ПРД.

ПРД-ядро может быть эффективно использовано специально для класса игр с динамическим и непрерывным обновлением информации. При построении кооперативного решения с играх с обновлением информации на каждом шаге или моменте моделирования используется процедура распределения дележа соответствующая кооперативному решению заданному на временном интервале информационного горизонта. Далее на основе ПРД на каждом шаге строится кооперативное решение на всем временном интервале, на котором определена игра. Однако, ПРД-ядро изначально строится с помощью ПРД, что позволяет сократить количество шагов вычислений для получения кооперативного решения с динамическим обновлением информации. В частности, в работе [105] ПРД-ядро используется для модели дифференциальной игры с динамическим обновлением и случайным горизонтом. В работе [106] исследуется и доказыва-ется связь ПРД-ядра заданного на горизонте обновления информации и ПРД-ядра заданного для дифференциальной игры с динамическим обновлением информации.

Вторая и третья главы. Класс игр с динамическим обновлением информации стал первым в потоке работ, посвященных обновлению информации.

Этот класс игр в основном изучался в статьях автора диссертации и его соавторов, см. [76; 77; 105; 106; 107; 108; 109; 110; 111; 112; 113; 114; 115; 116]. Данные работы заложили основу для дальнейших исследований в классе игр с динамическим обновлением информации в предположении, что информация об уравнениях движения и функциях выигрыша обновляется в дискретные моменты времени, а интервал, на котором игроки владеют информацией, определяется значением информационного горизонта.

Первой работой, посвященной этому классу игр, является [76]. В данной статье построена модель кооперативной дифференциальной игры с заданной продолжительностью и динамическим обновлением информации. Введены концепция усеченной подыгры, а также результирующие кооперативные стратегии, условно-кооперативная траектория и результирующее кооперативное решение. Доказана теорема о том, что произвольное результирующее кооперативное решение является Δt -динамически устойчивым в этом классе игр. Игры с динамическим обновлением информации, стохастическим прогнозом и динамической адаптацией представлены в [77]. В статье [112] описанный выше подход применен к игровым моделям с бесконечным горизонтом. Статья [105] посвящена специальному классу дифференциальных игр с динамическим обновлением информации и случайным горизонтом. Рассмотрена кооперативная постановка, использовано новое кооперативное решение (ПРД-ядро), и представлена новая процедура обновления параметров случайного горизонта. В статье [106] подробно описаны кооперативные дифференциальные игры с динамическим обновлением информации. Предложен подход к построению кооперативных стратегий, соответствующей траектории, характеристической функции и кооперативного решения с динамическим обновлением информации. Доказана связь между кооперативными решениями на усеченном временном интервале и на всем временном интервале. В статье [108] изучена зависимость выигрышей игроков от значения информационного горизонта. В статье [109] исследованы свойства кооперативного решения в классе игровых моделей с динамическим обновлением информации. Статья [116] посвящена построению специального класса уравнений Гамильтона–Якоби–Беллмана для некооперативной динамической игровой модели, определяющей различные типы информационных структур. Полученные результаты могут быть использованы для построения моделей, в которых игроки используют различные

информационные структуры. В статьях [110; 111] и [107] игровые модели с динамическим обновлением информации применены к олигополистической модели нефтяного рынка. Численное моделирование проведено в Matlab с использованием реальных данных по ценам нефти марок Brent и Light. В статье [113] исследовался класс динамических игр с обновлением информации, в кооперативной и некооперативной постановках. В статьях [114] и [115] дискретная модель с динамическим обновлением информации применена к задачам рекламы.

Четвертая глава. Класс дифференциальных игр с непрерывным обновлением информации рассмотрен в статьях [78; 79; 117; 118; 119; 120; 121; 122; 123; 124]; здесь предполагается, что процесс обновления информации постоянно протекает во времени.

Впервые этот класс игр рассмотрен в статьях [78] и [79]. В [78] выведена система уравнений Гамильтона–Якоби–Беллмана для равновесия Нэша в позиционных стратегиях с непрерывным обновлением информации. Статья [79] посвящена классу автономных линейно-квадратичных дифференциальных игр с непрерывным обновлением информации, для которых рассмотрены позиционные стратегии. Доказана сходимости равновесных по Нэшу стратегий и траекторий с динамическим и непрерывным обновлением информации. Далее в статье [118] рассмотрен класс кооперативных дифференциальных игр с трансферабельной полезностью с использованием уравнений Гамильтона–Якоби–Беллмана. Кроме того, построена характеристическая функция с непрерывным обновлением информации и представлены несколько теорем, а также доказано свойство сильной динамической устойчивости кооперативного решения с непрерывным обновлением информации. Другой результат, связанный с уравнениями Гамильтона–Якоби–Беллмана с непрерывным обновлением информации, посвящен классу кооперативных дифференциальных игр с нетрансферабельной полезностью, см. [119]. В статье [124] подробно исследована игровая модель добычи ресурсов в кооперативной и некооперативной постановках. В статье [120] явный вид равновесия по Нэшу для дифференциальной игры с непрерывным обновлением информации выведен с использованием принципа максимума Понтрягина. В статье [123] изучена кооперативная постановка с использованием принципа максимума Понтрягина. Другие статьи, касающиеся класса линейно-квадратичных дифференциальных игр с непрерывным обновлением

информации, посвящены программным равновесным по Нэшу стратегиям [117] и кооперативной постановке в форме характеристической функции с непрерывным обновлением информации [121]. В последней статье, посвященной линейно-квадратичной игре с непрерывным обновлением информации [122], исследован неавтономный случай, когда сама модель обновления зависит от текущего времени t . Более того, результаты сходимости получены и для неавтономного случая.

Пятая глава. Другие потенциально важные результаты связаны с постановкой и решением обратной задачи оптимального управления с непрерывным обновлением информации. Прикладные обратные задачи оптимального управления играют все большую роль по мере того, как растет взаимодействие человека с различными программными инструментами и техническими интерфейсами. Важным моментом для обеспечения оптимальной реакции технической программной системы на действия пользователя является определение его типа и целей. Одним из подходов к этой проблеме может быть решение обратной задачи оптимального управления, в которой тип человека и его цели определяются путем задания подынтегральной функции его целевой функции. В этом направлении была подготовлена первая статья [125], посвященная применению подхода с непрерывным обновлением информации к классу обратных задач оптимального управления. Подход может быть использован для определения модели поведения водителя, необходимой в системах помощи водителю. Целью данного исследования является определение профиля поведения водителя для построения адаптивной системы помощи в управлении автомобилем в различных режимах движения (экономичный, стандартный, спортивный). Подход с непрерывным обновлением информации позволяет точнее моделировать поведение человека за счет предположений о постоянном обновлении информации о динамической системе и целевой функции.

Цель и задачи диссертации

Целью диссертации является исследование нового класса моделей конфликтного управления (дифференциальных и динамических игр с непрерывным и динамическим обновлением информации), построение условий оптимальности для некооперативных и кооперативных стратегий с непрерывным и динамическим обновлением информации, алгоритма построения соответствующей

траектории, характеристической функции и кооперативного решения с непрерывным обновлением информации. Наиболее важно, что для этого необходимо изучить особые свойства кооперативных решений с непрерывным обновлением информации, включая динамическую устойчивость и состоятельность во времени. Более того, целью включает исследование соответствия между случаями динамического и непрерывного обновления информации путем доказательства сходимости оптимальных стратегий и соответствующих траекторий и изучения соответствия между кооперативными решениями на интервале обновления и во всей игре с динамическим или непрерывным обновлением информации. Еще одна составляющая — найти приложение для подхода с непрерывным обновлением информации и представить для него конкретные результаты моделирования.

Указанная выше цель исследования достигается путем решения следующих **задач диссертации:**

- разработать новый класс моделей конфликтного управления (некооперативных и кооперативных динамических и дифференциальных игр с динамическим и непрерывным обновлением информации);
- разработать методы построения равновесия по Нэшу, Парето-оптимальных, кооперативных стратегий, а также характеристической функции и кооперативного решения игр с динамическим и непрерывным обновлением информации;
- получить условия оптимальности для кооперативных стратегий и равновесия по Нэшу с динамическим и непрерывным обновлением информации;
- разработать новое кооперативное решение для класса игр с динамическим и непрерывным обновлением информации;
- изучить сходимость решений с динамическим обновлением информации к решениям с непрерывным обновлением информации;
- разработать алгоритмы построения характеристической функции и кооперативного решения с динамическим и непрерывным обновлением информации;

- изучить свойства Δt -сильной динамической устойчивости и сильной динамической устойчивости кооперативных решений с динамическим и непрерывным обновлением информации;
- исследовать обратную задачу оптимального управления с непрерывным обновлением информации.

Теоретическая новизна диссертации

Научной новизной диссертационного исследования является разработка теории новых моделей конфликтного управления (динамических и дифференциальных игр с динамическим и непрерывным обновлением информации), создание и адаптация ключевых условий оптимальности для кооперативных и равновесных стратегий в этом классе игр, создание и исследование кооперативных решений, созданных специально для класса игр с динамическим и непрерывным обновлением информации, исследование свойств кооперативных и некооперативных решений с динамическим и непрерывным обновлением информации, исследование свойств динамической устойчивости и временной состоятельности кооперативных решений с непрерывным обновлением информации, исследование сходимости кооперативных и некооперативных решений с динамическим и непрерывным обновлением информации. Также новизной исследования является построение и исследование некоторого класса обратных задач оптимального управления с непрерывным обновлением информации.

Теоретическая и практическая значимость

Теоретическая значимость работы заключается в том, что она развивает и описывает не только класс игр с непрерывным и динамическим обновлением информации, но и представляет условия оптимальности в виде уравнений Гамильтона–Якоби–Беллмана, принципа максимума Понтрягина, а также решение для линейно-квадратичного случая для программных и позиционных стратегий соответственно. Использование классических подходов для этих целей ограничено из-за характера и предположений о непрерывном обновлении информации. Следующие теоретические результаты связаны с исследованием свойства динамической устойчивости и временной состоятельности кооперативного решения с непрерывным и динамическим обновлением информации:

доказано, что кооперативное решение игры с непрерывным (динамическим) обновлением информации всегда обладает сильной динамической устойчивостью (Δt -динамической устойчивостью) и состоятельностью во времени (Δt -состоятельностью во времени) в зависимости от типа кооперации. Другие теоретические результаты связаны с доказательством связи между решениями, определенными в подыграх, и решениями, определенными на всем временном интервале, то есть решениями с непрерывным обновлением информации. Доказана равномерная и поточечная сходимость оптимальных стратегий с динамическим обновлением информации и соответствующих им траекторий, а также оптимальных стратегий с непрерывным обновлением информации и соответствующих им траекторий, для частного случая линейно-квадратичной игровой модели.

Теоретическая значимость работы подтверждена проектами, поддержанными научными фондами и Санкт-Петербургским государственным университетом. А именно, автор выступил в качестве руководителя

- проекта №18-71-00081 «Разработка методологического инструментария математической теории динамически устойчивых соглашений для цифровых интеллектуальных систем для формирования соглашения о сотрудничестве как шаг к искусственному интеллекту в этой области» (Российский научный фонд),

а также в качестве участника

- проекта №18-00-00725 COMFI «Развитие теоретико-игровых методов анализа конфликтов и сотрудничества в изменяющейся среде» (Российский фонд фундаментальных исследований),
- проекта НИР №9.38.205.2014 «Новые конструктивные подходы в негладком анализе и недифференцируемой оптимизации и их приложения» (Санкт-Петербургский государственный университет),
- проекта НИР №9.38.76.2011 «Развитие теории и конструктивных методов негладкого анализа и их применение к решению задач математического моделирования» (Санкт-Петербургский государственный университет).

Практическая значимость работы определяется тем, что реально существующие конфликтно-управляемые процессы развиваются во времени непре-

рывно и участники этих процессов обычно не имеют или не могут использовать полную информацию о них. Поэтому в теории игр важно ввести класс динамических и непрерывных игр с обновлением информации, актуальных для моделирования экологических и социально-экономических систем. Еще одним существенным практическим вкладом диссертации является построение и исследование класса обратных задач оптимального управления с непрерывным обновлением информации. Эти задачи могут быть использованы для анализа профиля человека в технических системах человеко-машинного типа. Результаты проиллюстрированы для модели системы помощи водителю и применены к реальным данным, полученным с помощью симулятора вождения в Институте систем управления Технологического института Карлсруэ, Германия.

Методы исследования

В диссертации используются **методы исследования** динамической и дифференциальной теории игр (построение динамических и дифференциальных игр, подыгр), некооперативной теории игр (ситуация равновесия по Нэшу, связанная траектория), теории кооперативных игр (кооперативные стратегии, характеристическая функция, динамическая устойчивость и временная состоятельность кооперативных решений, процедура распределения дележа), теории оптимального управления (динамическое программирование, принцип максимума Понтрягина и линейно-квадратичный частный случай), математического анализа (сходимость кооперативных, некооперативных стратегий и траекторий), а также теории оптимизации и вероятностей (распределения случайных величин, случайных процессов).

Результаты, выносимые на защиту

Основные результаты, полученные в диссертации и выносимые на защиту:

- 1 Описаны и изучены новые классы моделей конфликтного управления (динамических игр с обновлением информации, дифференциальных игр с динамическим обновлением информации и дифференциальных игр с непрерывным обновлением информации). Изучены кооперативные постановки с трансферабельной и нетрансферабельной полезностями, а также

некооперативная постановка. Приведены методы построения равновесных по Нэшу, оптимальных по Парето и кооперативных стратегий, характеристической функции и кооперативного решения игр с динамическим и непрерывным обновлением информации.

- 2 Новые условия оптимальности для кооперативных и равновесных по Нэшу стратегий, как программных, так и позиционных, представлены в форме уравнений Гамильтона–Якоби–Беллмана. Сформулирован принцип максимума Понтрягина с непрерывным обновлением информации. Приведены условия оптимальности для случая линейно-квадратичных автономных и неавтономных моделей с непрерывным обновлением информации. Доказаны соответствующие теоремы. Эти условия оптимальности использованы для решения классических дифференциальных игровых моделей для случая непрерывного обновления информации.
- 3 Представлены и доказаны новые результаты о связи решений для игровых моделей с динамическим и непрерывным обновлением информации в линейно-квадратичном автономном и неавтономном случаях, а также для специальной классической дифференциальной игровой модели добычи ресурсов. Доказана сходимости кооперативных и равновесных по Нэшу стратегий и траекторий с динамическим обновлением информации к соответствующим стратегиям и траекториям с непрерывным обновлением информации.
- 4 Разработаны алгоритмы построения характеристической функции и кооперативного решения с динамическим и непрерывным обновлением информации. Доказана связь между кооперативными решениями, заданными на усеченных временных интервалах, определяемых информационным горизонтом и кооперативными решениями с динамическим и непрерывным обновлением информации. Установлены свойства Δt -сильной динамической устойчивости и сильной динамической устойчивости для произвольных кооперативных решений соответственно при динамическом и непрерывном обновлении информации.
- 5 Определено новое кооперативное решение (ПРД-ядро) для дифференциальных игр с динамическим и непрерывным обновлением информации,

основанное на аксиомах динамической устойчивости и недоминирования по ПРД. Доказана сильная динамическая устойчивость этого решения. Представлена явная формула для построения ПРД-ядра. Представлен численный алгоритм для анализа свойства непустоты. Продемонстрировано применение ПРД-ядра для модели дифференциальной игры с динамическим обновлением информации.

- 6 Сформулирована и решена обратная задача оптимального управления с непрерывным обновлением информации. Подход с непрерывным обновлением информации использован для прикладной обратной задачи оптимального управления системы помощи водителю.

Апробация полученных результатов

Результаты, представленные в диссертации, были доложены и **апробированы** на следующих международных конференциях и семинарах (с достаточным количеством международных участников):

- CPS2012–2013,2015: XLIII, XLIV, и XLVI международные конференции «Процессы управления и устойчивость», Факультет прикладной математики и процессов управления Санкт-Петербургского государственного университета, Санкт-Петербург (Россия);
- GTM2013–2020: 7-я, 8-я, 9-я, 10-я, 11-я, 12-я, 13-я, и 14-я международные конференции по теории игр и менеджменту, Санкт-Петербург (Россия);
- Второй Российский экономический конгресс, Суздаль (Россия), 18–22 февраля 2013 г.;
- SING11-GTM2015: Европейское совещание по теории игр, Санкт-Петербург (Россия), 8–10 июля 2015 г.;
- SING14,15: Европейские совещания по теории игр, Байройт (Германия), 2018 г. и Турку (Финляндия), 2019 г.;
- SCP2015: Конференция «Устойчивость и процессы управления» памяти проф. Владимира Зубова, Санкт-Петербург (Россия), 5-9 октября 2015 г.;

- STAB2016, 2020: XV Международная конференция «Устойчивость и колебания нелинейных систем управления» (конференция Пятницкого), Москва (Россия), 2016 г. и 2020 г. (онлайн);
- GAM-SOC2015: Семинар по теории игр и социальному выбору, Будапешт (Венгрия), 11–12 декабря 2015 г.;
- Международная конференция «Конструктивный негладкий анализ и смежные вопросы», посвященной памяти В.Ф. Демьянова, Санкт-Петербург (Россия), 22–27 мая 2017 г.;
- ISDG18: 18-й Международный симпозиум по динамическим играм и приложениям, Гренобль (Франция), 9–12 июля 2018 г.;
- ISDG-GTM2019: Международная конференция по теории игр, Санкт-Петербург (Россия), 3–5 июля 2019 г.;
- MOTOR2019, 2020: Математическая теория оптимизации и исследование операций, Екатеринбург (Россия), 2019 г. и Новосибирск (Россия), 2020 г. (онлайн);
- IFAC2020: Всемирный конгресс по оптимальному управлению, Берлин (Германия), 12–17 июля 2020 г. (онлайн);
- 31-я Международная конференция по теории игр в Стоуни-Брук. Семинар по дифференциальным играм. Примечание: автор диссертации организовал семинар и сделал презентацию ([http : //www.gtcenter.org/?page = Workshops.html](http://www.gtcenter.org/?page=Workshops.html)), Стоуни Брук, Нью-Йорк (США), 20–24 июля 2020 г. (онлайн).

Публикации. Автором опубликовано более 38 научных работ (26 из них Web of Science/Scopus), из них 28 работы по теме исследования [76; 77; 78; 79; 100; 101; 102; 104; 106; 107; 108; 110; 112; 113; 114; 115; 116; 117; 118; 119; 120; 125; 126; 127; 128; 129], из которых 6 работ опубликованных в научных изданиях, включенных в перечень рецензируемых научных изданий, рекомендованных ВАК РФ [100; 102; 112; 115; 127; 128], 22 работ - в изданиях, индексируемых в международных наукометрических базах SCOPUS/Web of Science [76; 77; 78; 79; 101; 104; 106; 107; 108; 109; 110; 112; 113; 114; 115;

117; 118; 119; 120; 125; 126; 129]. Основные научные результаты, выносимые на защиту, опубликованы в рецензируемых научных изданиях и отражены в следующих работах: [76; 77; 78; 79; 105; 106; 107; 108; 109; 110; 111; 112; 113; 114; 115; 116; 117; 118; 119; 120; 123; 126; 127; 128; 129] - пункт 1, [78; 79; 117; 118; 119; 120; 123; 125; 129] - пункт 2, [79] - пункт 3, [76; 77; 106; 109; 112; 113; 117; 118; 119; 123; 126; 127; 128] - пункт 4, [100; 101; 102; 104] - пункт 5, [125] - пункт 6.

Диссертация состоит из введения, пяти глав, заключения, списка обозначений, списка литературы и приложений. Определения, утверждения, теоремы, примеры и примечания нумеруются в соответствии с главой и разделом, в котором они расположены. Следствия пронумерованы в соответствии с теоремами, к которым они относятся. Объем работы составляет 417 страниц. Библиография включает 237 источников. В работе представлены 100 рисунков, 8 таблиц и 0 приложений.

Благодарности

Автор выражает глубочайшую благодарность д.физ.-мат.н., заведующему кафедрой Математического моделирования энергетических систем СПбГУ, проф. В.В. Захарову за долгую и терпеливую работу с автором, поддержка которого существенным образом помогла закончить диссертацию.

Автор считает своим долгом поблагодарить всех сотрудников факультета прикладной математики и процессов управления СПбГУ за то, что они были научной семьей, вырастившей, будем надеяться, ученого, достойного наследия В.И. Зубова. Автор выражает благодарность докторам инженерии М. Фладу и Ж. Инга за то, что они познакомили автора с классом обратных задач оптимального управления; проф. Х. Гао за постановку больших научных целей и поддержку в ненаучных начинаниях; Д. Чаи за то, что он интенсивно продвигал ненаучную работу и познакомил с другими научными направлениями, представляющими интерес для автора в настоящее время; Х. Вэй за дружеское отношение, доверие и помощь в научном и личностном развитии, за хороший пример руководители и просто хорошего человека; д.физ.-мат.н., проф. А.Ю. Крылатову за отличный пример непреодолимого духа и поддержку многими жизненными советами в начале научной карьеры автора; доц. к.физ.-мат.н. И.С. Блеканову за его открытость и человечность,

интересные научные дискуссии и совместные начинания в последний период написания диссертации; Ли Инь за дружбу и помощь автору в различных совместных делах и бескорыстную поддержку на пути автора; всем дорогим для автора аспирантам и студентам, которые усердно работали с ним бок о бок в течение многих лет; Л. Ши за его бесконечную энергию; Ц. Ванг за ее сильный дух и желание стать к.физ.-мат.н.; Д. Жу за его упорный труд, понимание и веру своему научному руководителю; И. Кучкарову за его упорный многолетний труд бок о бок с автором над подходом с непрерывным обновлением информации; многим другим людям, с которыми автор начал работать недавно и которых не упомянул здесь. Автор благодарит Г.В. Янченкову за то, что она всегда была на его стороне и поддерживала его в школьное время. Автор глубоко благодарит свою семью за то, что они не теряют надежды на сына, за постоянную помощь, поддержку, советы в каждом начинании, за нетривиальную, интересную и незабываемую жизнь. Особенно автор глубоко благодарит свою маму Нину Ю. Митину за ее бесконечную заботу, за то, что она подарила жизнь, и за то, что она всю жизнь была примером красивой и совершенной трудолюбивой женщины. Автор глубоко благодарит своего отца Леона А. Петросяна, чья непоколебимая мудрость и важные советы всегда помогали решать большинство проблем на его пути, за то, что он был тем, с кем можно было обсудить все жизненные и научные проблемы и, наконец, за отличный пример лучшего человека и лучшего ученого. Автор благодарит своего брата Евгения В. Митина и сестру Наиру Л. Казакову-Фрезе за их особое место в жизни автора, за их советы, помощь и большую поддержку в разные и трудные периоды жизни. Автор благодарит всю семью своей сестры, ее мужа Торстена Фрезе, сына Карла Фрезе за то, что они были хорошими друзьями, и, наконец, ее дочь Татьяну В. Казакову, с которой автор всегда брал пример в жизни.

Глава 1

ПРД-ядро: новое кооперативное решение для дифференциальных игр

Эта глава посвящена новому кооперативному решению для динамических и дифференциальных игр, а именно ПРД-ядру. Предлагается новый подход к построению решений для динамических и дифференциальных игр, в котором динамическая устойчивость используется в качестве основной аксиомы для кооперативного решения. Еще одним новым важным свойством и подходом для построения ПРД-ядра, является свойство доминируемости по ПРД, которое позволяет выбрать недоминируемые процедуры распределения дележей и получить кооперативное решение или множество дележей. Этот подход показывает потенциал свойства динамической устойчивости как основной аксиомы при решении задач в различных областях, таких как теория социального выбора и теория дизайна механизмов. Общая процедура определения кооперативного решения также является новой, поскольку ПРД-ядро строится с использованием процедур распределения дележа, а не самих дележей. Это делает ПРД-ядро полезным кооперативным решением для класса игр с динамическим и непрерывным обновлением информации.

При построении кооперативного решения с играх с обновлением информации на каждом шаге или моменте моделирования используется процедура распределения дележа соответствующая кооперативному решению заданному на временном интервале информационного горизонта. Далее на основе ПРД на каждом шаге строится кооперативное решение на всем временном интервале, на котором определена игра. Однако, ПРД-ядро изначально строится с помощью ПРД, что позволяет сократить количество шагов вычислений для получения кооперативного решения с динамическим обновлением информации. Ниже,

в Главе 3 представлены результаты применения ПРД-ядра для дифференциальных игр с динамическим обновлением информации [105], [106]. В [105] ПРД-ядро используется для модели дифференциальной игры с динамическим обновлением и случайным горизонтом. В работе [106] исследуется и доказывается связь ПРД-ядра заданного на горизонте обновления информации и ПРД-ядра заданного для дифференциальной игры с динамическим обновлением информации.

1.1 Постановка задачи и предварительные сведения

1.1.1 Модель дифференциальной игры

В 50-х годах XX века теория дифференциальных игр развивалась как отдельный класс прикладной математики. Одной из первых работ в области дифференциальных игр является книга Р. Айзекса [130], в которой сформулированы понятия состояния и управления, задачи перехвата самолетов управляемой ракетой, а также получено фундаментальное уравнение для нахождения решения. Среди других работ, описывающих класс дифференциальных игр с нулевой суммой, стоит упомянуть [3; 82; 83; 131; 132]. Для рассмотрения ситуаций, в которых несколько участников (игроков) имеют разные цели или функции выигрыша (и, следовательно, действуют индивидуально), разработаны модели некооперативных дифференциальных игр. В качестве принципа оптимальности в некооперативных дифференциальных играх чаще всего используется равновесие по Нэшу в программных или позиционных стратегиях, см. [26; 41; 80; 133]. Интересные приложения теории дифференциальных игр могут быть найдены в работе [134].

В этом разделе дается общее описание дифференциальной игры. Основными понятиями этой модели являются тип игры, функции выигрыша игроков, уравнения движения и концепция кооперативного решения. Тип игровой модели отражает то, что с ней планируется делать. Здесь рассматривается кооперативная игровая модель. Нужно определить, каким образом распределять суммарный кооперативный выигрыш между игроками. Функции выигрыша игроков задают цели игроков в зависимости от состояния игры и их стратегий и рассчитываются на некотором конкретном временном интервале (в нашем случае, временном отрезке). Уравнения движения описывают динамику состояния игры в соот-

ветствии со стратегиями игроков. В случае кооперативной игровой модели, решение точно определяет множество дележей, которое будет использоваться для распределения суммарного выигрыша между игроками.

Рассмотрим дифференциальную игру n лиц $\Gamma(x_0, T - t_0)$ с заданными продолжительностью $T - t_0$ и начальным условием x_0 . Игровая динамика задается системой дифференциальных уравнений

$$\begin{aligned} \dot{x} &= f(x, u_1, \dots, u_n), \quad x \in R^n, \quad u_i \in U_i \subset \text{comp}R^k, \quad t \in [t_0, T], \quad i = \overline{1, n}, \\ x(t_0) &= x_0, \end{aligned} \quad (1.1)$$

где программное управление $u_i(t)$, удовлетворяющее системе (1.1), является стратегией игрока i , и $\text{comp}R^k$ обозначает компактное множество в k -мерном пространстве действительных чисел (k — целое число).

Пусть $N = \{1, \dots, n\}$ — множество игроков. Выигрыш игрока i определяется следующим образом:

$$K_i(x_0, T - t_0; u_1, \dots, u_n) = \int_{t_0}^T h_i(x(\tau), u_1(\tau), \dots, u_n(\tau)) d\tau, \quad i = \overline{1, n}, \quad (1.2)$$

где $h_i(x, u_1, \dots, u_n) \geq 0$, $i = \overline{1, n}$, и $f(x, u_1, \dots, u_n)$ — интегрируемые функции, а $x(t)$ представляет собой решение системы (1.1) при управлениях $u(t) = (u_1(t), \dots, u_n(t))$. Набор стратегий $u(t, x) = (u_1(t, x), \dots, u_n(t, x))$ называется допустимым, если задача (1.1) имеет единственное и продолжимое на отрезке $[t_0, T]$ решение. Здесь и ниже в диссертации, используются условия существования, единственности и продолжимости А.Ф. Филиппова [135]:

- 1 Функция $f(t, x, u)$ в правой части уравнений движения (1.1) непрерывна на множестве $[t_0, T] \times X \times U_1 \times \dots \times U_n$.
- 2 Функция $f(t, x, u)$ в правой части уравнений движения удовлетворяет условиям Липшица по x равномерно по u с некоторой константой $k_1 > 0$:

$$\|f(t, x', u) - f(t, x'', u)\| \leq k_1 \|x' - x''\|, \quad \forall t \in [t_0, T], \quad x', x'' \in X, \quad u \in U.$$

- 3 Существует некоторая константа k_2 при которой функция $f(t, x, u)$ удо-

влетворяет неравенству

$$\|f(t, x, u)\| \leq k_2(1 + \|x\|), \quad \forall t \in [t_0, T], \quad x \in X, u \in U.$$

4 Для любых $t \in [t_0, T]$ и $x \in X$, множество

$$G(x) = \{f(t, x, u) | u \in U\}$$

является выпуклым компактом на R^l .

1.1.2 Модель кооперативной дифференциальной игры

Теория кооперативных игр исследует вопросы определения оптимальных параметров кооперативных и стратегических соглашений. Основная проблема теории кооперативных игр с трансферабельной полезностью состоит в том, чтобы определить порядок распределения общего выигрыша в случаях, когда все игроки сотрудничают. Правило распределения выигрыша между игроками называется дележом. В теории кооперативных игр с нетрансферабельными полезностями, основная проблема заключается в определении соглашения о стратегиях или исходе игры, благоприятном для всех игроков.

В рамках классической теории кооперативных игр с нетрансферабельными полезностями изучены многочисленные кооперативные решения или правила распределения, в частности, С-ядро. Концепция С-ядра предложена Д. Гиллиссом [136] и является обобщением контрактной кривой Эджворта [137]. Эджворт описал рынок с двумя продуктами и двумя участниками, определив С-ядро как часть множества Парето оптимальных решений. С-ядро представляет собой множество недоминируемых дележей, каждый из которых может использоваться в качестве решения в игре. Интересное обсуждение возможности применения кооперативных решений для прикладных задач обсуждаются в работе [138].

Также важно построить кооперативные решения для класса динамических и дифференциальных игр. Решения для таких моделей могут быть использованы для моделирования кооперативных и стратегических соглашений, в которых условия определяются в течение длительного периода времени [139; 140]. Подробное описание динамических кооперативных игр представлено в [141]. Еще больше кооперативных дифференциальных и динамических игровых моделей

рассмотрено в статьях [9; 142; 143; 144; 145].

В кооперативной дифференциальной игре с трансферабельной полезностью возникают две проблемы:

- 1 Определить набор стратегий игроков, который максимизирует суммарный выигрыш или определить стратегии, соответствующие кооперативному поведению. Эти стратегии $u^* = (u_1^*, \dots, u_n^*)$ называются оптимальными, а соответствующая траектория называется кооперативной траекторией и обозначается через $x^*(t)$.
- 2 Определить правила распределения максимального кооперативного выигрыша игроков, соответствующего оптимальным стратегиям $u^*(t)$, и определить оптимальную траекторию $x^*(t)$. А именно, определить кооперативное решение как подмножество множества дележей.

Пусть $u^* = (u_1^*, \dots, u_n^*)$ — вектор оптимальных программных стратегий игроков, т.е. вектор управлений без обратной связи, который максимизирует кооперативный выигрыш игроков:

$$u^* = (u_1^*, \dots, u_n^*) = \arg \max_{u_1, \dots, u_n} \sum_{i=1}^n K_i(x_0, T - t_0; u_1, \dots, u_n). \quad (1.3)$$

Предположим, что максимум в (1.3) достигается на множестве допустимых стратегий.

Чтобы определить распределение максимального суммарного выигрыша между игроками, необходимо ввести понятие характеристической функции коалиции $S \subseteq N$. Характеристическая функция показывает силу коалиции и, таким образом, позволяет учитывать вклад игроков в каждую коалицию.

Предположим, что в игре $\Gamma(x_0, T - t_0)$ характеристическая функция $V(S; x_0, T - t_0)$, $S \subseteq N$ построена любым подходящим способом (например, как описано в [146]). Считаем, что выполнены условия супераддитивности:

$$\begin{aligned} V(S_1 \cup S_2; x_0, T - t_0) &\geq V(S_1; x_0, T - t_0) + V(S_2; x_0, T - t_0), \\ \forall S_1, S_2 \subseteq N, S_1 \cap S_2 &= \emptyset. \end{aligned}$$

Обозначим через $L(x_0, T - t_0)$ множество дележей [147] в игре $\Gamma(x_0, T - t_0)$:

$$L(x_0, T - t_0) = \left\{ \xi(x_0, T - t_0) = (\xi_1(x_0, T - t_0), \dots, \xi_n(x_0, T - t_0)) : \right. \\ \left. \sum_{i=1}^n \xi_i(x_0, T - t_0) = V(N; x_0, T - t_0), \right. \\ \left. \xi_i(x_0, T - t_0) \geq V(\{i\}; x_0, T - t_0), i \in N \right\},$$

где $V(\{i\}; x_0, T - t_0)$ — значение характеристической функции $V(S; x_0, T - t_0)$ для коалиции $S = \{i\}$.

Обозначим через $M(x_0, T - t_0)$ произвольное кооперативное решение или подмножество множества дележей $L(x_0, T - t_0)$:

$$M(x_0, T - t_0) \subseteq L(x_0, T - t_0).$$

Предположим, что в начале игры $\Gamma(x_0, T - t_0)$ в момент t_0 игроки договорились о выборе некоторого подмножества $L(x_0, T - t_0)$ или какого-либо кооперативного решения. Однако предположим, что в какой-то момент \bar{t} игроки решили пересмотреть выбранное кооперативное решение или правило распределения кооперативного выигрыша. Чтобы смоделировать их поведение, необходимо определить понятие подыгры $\Gamma(x^*(t), T - t)$ вдоль кооперативной траектории $x^*(t)$, начинающейся в момент времени $t \in [t_0, T]$.

Для каждой подыгры $\Gamma(x^*(t), T - t)$, $t \in [t_0, T]$ вдоль траектории $x^*(t)$, определим супераддитивную характеристическую функцию $V(S; x^*(t), T - t)$, $S \subseteq N$, так же, как это было сделано для исходной игры $\Gamma(x_0, T - t_0)$:

$$\forall S, A \subseteq N, S \cap A = \emptyset :$$

$$V(S \cup A; x^*(t), T - t) \geq V(S; x^*(t), T - t) + V(A; x^*(t), T - t). \quad (1.4)$$

Также возможно определить понятие дележа $\xi(x^*(t), T - t)$ для подыгры $\Gamma(x^*(t), T - t)$ вдоль кооперативной траектории $x^*(t)$, $t \in [t_0, T]$. Множество всевозможных дележей в подыгре $\Gamma(x^*(t), T - t)$ обозначим через $L(x^*(t), T - t)$,

$t \in [t_0, T]$:

$$L(x^*(t), T - t) = \left\{ \begin{aligned} &\xi(x^*(t), T - t) = (\xi_1(x^*(t), T - t), \dots, \xi_n(x^*(t), T - t)) : \\ &\sum_{i=1}^n \xi_i(x^*(t), T - t) = V(N; x^*(t), T - t), \\ &\xi_i(x^*(t), T - t) \geq V(\{i\}; x^*(t), T - t), \quad i \in N \end{aligned} \right\}. \quad (1.5)$$

Свойство супераддитивности (1.4) для характеристической функции $V(S; x^*(t), T - t)$ гарантирует непустоту множества дележей $L(x^*(t), T - t)$, $t \in [t_0, T]$. Кооперативное решение подыгры $\Gamma(x^*(t), T - t)$ обозначим соответственно через $M(x^*(t), T - t)$.

1.1.3 С-ядро

В теории кооперативных игр, основной проблемой является «справедливое» распределение максимального суммарного выигрыша $V(N; x_0, T - t_0)$ между игроками гранд-коалиции $N = \{1, \dots, n\}$.

Предположим, что игроки в кооперативной дифференциальной игре $\Gamma(x_0, T - t_0)$ (подыгре $\Gamma(x^*(t), T - t)$, $t \in [t_0, T]$ вдоль кооперативной траектории $x^*(t)$) договорились о правиле распределения выигрыша $\xi(x_0, T - t_0)$ (дележ $\xi(x^*(t), T - t)$), в котором ни один из дележей не доминирует $\xi(x_0, T - t_0)$ ($\xi(x^*(t), T - t)$) [147]. Такое правило распределения устойчиво, поскольку не существует дележа, который был бы лучше для каждой коалиции в каждый момент времени $t \in [t_0, T]$.

Определение 1.1.0.1 *Множество недоминируемых дележей в кооперативной дифференциальной игре $\Gamma(x^*(t), T - t)$ будем называть С-ядром и обозначать через $C(x^*(t), T - t)$, $t \in [t_0, T]$.*

Имеет место следующая теорема.

Теорема 1.1.1 *Дележ $\xi(x^*(t), T - t)$ принадлежит С-ядру $C(x^*(t), T - t)$, тогда и только тогда, когда для всех $S \subseteq N$ выполнены неравенства*

$$V(S; x^*(t), T - t) \leq \sum_{i \in S} \xi_i(x^*(t), T - t), \quad t \in [t_0, T].$$

1.1.4 Непустота С-ядра в статических играх

Важно изучить свойство непустоты кооперативного решения — определить условия, при которых кооперативное решение не будет пустым, поскольку его применимость зависит от величины класса игр, в которых это решение может быть использовано. Г. Скарф [148] показал, что С-ядро не пусто для класса выпуклых игр в форме характеристической функции. Характеристическая функция — это функция коалиции или подмножества игроков в игре, которая задает прибыль коалиции. Результаты Скарфа обобщены в статьях Л. Биллера [149] и Л. Шепли [150]. Необходимые и достаточные условия непустоты С-ядра сформулированы Бондаревой [151] и Шепли [152]; основная роль в доказательстве принадлежала концепции сбалансированной игры. К сожалению, на основе этой концепции невозможно применить конструктивный метод выбора конкретных дележей из С-ядра. В. Захаров [103] предложил необходимые и достаточные условия непустоты С-ядра, которые упрощают проверку принадлежности С-ядру для таких одноточечных решений (дележей) как вектор Шепли, индекс влияния Банцафа и других. В работах [153] и [154] на основе этого подхода исследованы геометрические свойства нескольких кооперативных решений. Этот подход подразумевает, что свойство непустоты С-ядра может быть сформулировано с помощью задачи линейного программирования, построенной с использованием значений характеристической функции.

Приведем основные результаты, касающиеся условий непустоты С-ядра в статических играх. Необходимые и достаточные условия непустоты С-ядра сформулированы О. Бондаревой [151] и Л. Шепли [152]. Эти условия основаны на концепции сбалансированной игры, но применение такого подхода для конкретной игровой модели затруднительно.

В статье [155] Г. Оуэн показал, что в игре (N, v) существует непустое С-ядро тогда и только тогда, когда оптимальное значение задачи линейного программирования

$$\begin{aligned} \sum_{i \in N} \xi_i &\longrightarrow \min \\ \sum_{i \in S} \xi_i &\geq v(S), \quad \forall S \subseteq N, \quad S \neq \emptyset \end{aligned}$$

равно $v(N)$.

В статьях [103; 153; 154] также использована задача линейного программирования для установления непустоты S -ядра. Рассмотрим следующую задачу линейного программирования:

$$\begin{aligned} \sum_{i \in N} \xi_i &\longrightarrow \min \\ \sum_{i \in S} \xi_i &\geq v(S), \quad \forall S \subseteq N, \quad S \neq N, \quad \emptyset. \end{aligned} \quad (1.6)$$

Предположим, что $\xi^0 = (\xi_1^0, \dots, \xi_n^0)$ — произвольное оптимальное решение задачи линейного программирования (1.6). Множество всех оптимальных решений оптимизационной задачи (1.6) обозначим через $X^0(v)$. В [103] показано, что необходимые и достаточные условия непустоты S -ядра могут быть формализованы следующим образом:

Теорема 1.1.2 *S -ядро в кооперативной игре с трансферабельной полезностью (N, v) непусто тогда и только тогда, когда выполнено неравенство*

$$\sum_{i \in N} \xi_i^0 \leq v(N), \quad (1.7)$$

где $\xi^0 \in X^0(v)$ — решение задачи линейного программирования (1.6).

1.1.5 Динамическая устойчивость кооперативного решения и процедура распределения дележа

Теория кооперативных дифференциальных игр помогает разработать условия кооперативного соглашения, а именно: кооперативные стратегии, соответствующую траекторию, общий выигрыш вдоль кооперативной траектории, правила распределения общего выигрыша между игроками и свойство динамической устойчивости решения. Естественный подход к исследованию кооперативных дифференциальных игр — это попытка перенести результаты классической статической кооперативной теории [156] на теорию дифференциальных игр. Однако, чтобы использовать результаты классической кооперативной теории, необходимо дополнительно изучить свойства динамической устойчивости и сильной динамической устойчивости кооперативных решений. Динамическая устойчивость кооперативного решения — свойство, при котором игрокам невыгодно отклоняться от выбранного кооперативного решения во время игры.

Проблема динамической устойчивости арбитражного решения Нэша в динамической игре о переговорах обсуждалась в [157]. Использование динамически неустойчивых кооперативных решений в экономике, экологии и менеджменте делает эти решения нереализуемыми, поскольку для получения выгоды игроки могут пересмотреть кооперативное решение.

Впервые концепции динамической устойчивости и сильной динамической устойчивости в кооперативных дифференциальных играх сформулированы Л. Петросяном соответственно в статьях [158] и [159]. Чтобы избежать нарушения условий кооперативного соглашения или свойства динамической устойчивости, в статье [160] предложена схема оплаты, названная процедурой распределения дележа (ПРД). Понятие сильной динамической устойчивости сформулировано в [161]. Работы [145; 162; 163; 164] посвящены исследованию свойства динамической устойчивости кооперативных решений. В недавней статье [165] авторы впервые определили понятие временной состоятельности для динамических кооперативных дифференциальных игр с нетрансферабельной полезностью. В той же статье авторы получили состоятельные во времени решения для класса кооперативных стохастических дифференциальных игр с нетрансферабельной полезностью. Другой метод построения состоятельного во времени решения представлен в [166].

Предположим, что ПРД в кооперативной дифференциальной игре $\Gamma(x_0, T - t_0)$ являются интегрируемыми векторными функциями, которые формируют некоторый дележ из множества дележей:

$$\beta(t) : \int_{t_0}^T \beta(\tau) d\tau \in L(x_0, T - t_0), \quad (1.8)$$

или

$$\int_{t_0}^T \beta_i(\tau) d\tau \geq V(\{i\}; x_0, T - t_0), \quad i \in N,$$

$$\sum_{i \in N} \int_{t_0}^T \beta_i(\tau) d\tau = V(N; x_0, T - t_0).$$

Следовательно, в приведенном выше определении, ПРД основывается не на самом дележе, а генерирует его. Определим так называемую «соответствующую ПРД», концепция которой близка к первоначальному определению ПРД в

статье [167].

Определение 1.1.0.2 *Интегрируемая функция $\beta(t) = (\beta_1(t), \dots, \beta_n(t))$, $t \in [t_0, T]$ называется соответствующей процедурой распределения дележа (ПРД) для $\xi(x_0, T - t_0) \in L(x_0, T - t_0)$, если выполняются равенства*

$$\xi_i(x_0, T - t_0) = \int_{t_0}^T \beta_i(\tau) d\tau, \quad i \in N. \quad (1.9)$$

Фактически, соответствующая ПРД $\beta(t)$ зависит от $\xi(x_0, T - t_0)$ и не является единственной для этого дележа. Она может быть представлена в виде

$$\beta(t) = \beta(t, \xi(x_0, T - t_0)),$$

или

$$\beta_i(t) = \beta_i(t, \xi_i(x_0, T - t_0)), \quad i \in N.$$

Из (1.9) мы имеем для $t \in [t_0, T]$, $i \in N$:

$$\xi_i(x_0, T - t_0) = \int_{t_0}^t \beta_i(\tau) d\tau + \int_t^T \beta_i(\tau) d\tau,$$

или

$$\int_t^T \beta_i(\tau) d\tau = \xi_i(x_0, T - t_0) - \int_{t_0}^t \beta_i(\tau) d\tau.$$

Другими словами, ПРД разбивает в момент времени t дележи на две части: выплаты игроку i на временном интервале $[t_0, t]$ и на временном интервале $(t, T]$.

Определение 1.1.0.3 *Кооперативное решение $M(x_0, T - t_0)$ в игре $\Gamma(x_0, T - t_0)$ называется динамически устойчивым, если для каждого дележа $\xi(x_0, T - t_0) \in M(x_0, T - t_0)$ существует соответствующая ПРД $\beta(t) = \beta(t, \xi(x_0, T - t_0))$, такая что:*

$$\int_t^T \beta(\tau) d\tau \in M(x^*(t), T - t), \quad t \in [t_0, T], \quad (1.10)$$

или, в эквивалентной форме,

$$\xi(x_0, T - t_0) - \int_{t_0}^t \beta(\tau) d\tau \in M(x^*(t), T - t), \quad t \in [t_0, T]. \quad (1.11)$$

Заметим, что из условия (1.10) следует равенство

$$\sum_{i \in N} \int_t^T \beta_i(\tau) d\tau = V(N; x^*(t), T - t), \quad t \in [t_0, T]. \quad (1.12)$$

Очевидно, если $M(x^*(t), T - t) \neq \emptyset \forall t \in [t_0, T]$, то для любой дифференцируемой по t функции $\xi(x^*(t), T - t) \in M(x^*(t), T - t)$ ($\xi(x^*(t_0), T - t_0) = \xi(x_0, T - t_0)$) ПРД $\beta(t)$ может быть задана в виде

$$\begin{aligned} \beta(t) &= -\frac{d}{dt} \xi(x^*(t), T - t), \quad t \in [t_0, T], \quad i \in N, \\ \xi(x^*(t_0), T - t_0) &= \xi(x_0, T - t_0). \end{aligned} \quad (1.13)$$

В этом случае, дележ $\xi(x_0, T - t_0)$ рассчитывается по формуле

$$\xi(x_0, T - t_0) = \int_{t_0}^t \beta(\tau) d\tau + \xi(x^*(t), T - t), \quad t \in [t_0, T].$$

Определим дележ в текущей кооперативной игре $\Gamma(x^*(t), Tt)$ с характеристической функцией $V(S; x^*(t), Tt)$, который соответствует данной ПРД $\beta(t) = \beta(t, \xi(x_0, T - t_0))$, в виде

$$\xi(x^*(t), T - t) = \int_t^T \beta(\tau) d\tau. \quad (1.14)$$

Из Определения 1.1.0.3 получаем:

$$\xi(x^*(t), T - t) \in M(x^*(t), T - t). \quad (1.15)$$

Будем называть дележ (1.14) динамическим дележом, сгенерированным соответствующей ПРД $\beta(t) = \beta(t, \xi(x_0, T - t_0))$.

1.2 ПРД-ядро и доминирование

Рассмотрим развитие игры в момент времени $t \in (t_0, T)$. Предположим, что в момент времени t_0 игроки согласились осуществить дележ $\xi(x_0, T - t_0) = (\xi_1(x_0, T - t_0), \dots, \xi_n(x_0, T - t_0))$. Тогда, согласно соответствующей ПРД $\beta(t)$, до момента времени t игрок $i \in N$ получает выигрыш

$$\int_{t_0}^t \beta_i(\tau) d\tau.$$

Однако для некоторых игроков ПРД $\beta(t, \xi(x_0, T - t_0))$ не будет выгодна, если существует другая процедура распределения дележа $\bar{\beta}(t, \xi(x_0, T - t_0))$, согласно которой игрок i на временном интервале $[t_0, t]$ получает более высокий выигрыш:

$$\int_{t_0}^t \bar{\beta}_i(\tau) d\tau > \int_{t_0}^t \beta_i(\tau) d\tau. \quad (1.16)$$

В таком случае, ПРД $\beta(t)$ может считаться менее выгодной для игрока i по крайней мере на временном интервале $[t_0, t]$. Важно отметить, что понятие доминирования по ПРД может применяться к процедурам распределения дележа, не обязательно определенным для единственного дележа. ПРД определяет, как должен быть построен динамический дележ. Поэтому понятие доминирования по ПРД имеет смысл не только для фиксированного дележа.

1.2.1 Доминирование по процедурам распределения дележа

В этом разделе рассмотрим ПРД $\beta(t)$, заданную формулой (1.8). Предположим, что функция $V(S; x^*(t), T - t)$, $S \subseteq N$, непрерывно дифференцируема по $t \in [t_0, T]$. Определим функцию $U(S; x^*(t), T - t)$ следующим образом:

$$U(S; x^*(t), T - t) = -\frac{d}{dt} V(S; x^*(t), T - t), \quad t \in [t_0, T], \quad S \subseteq N. \quad (1.17)$$

Определение 1.2.0.1 Будем говорить, что ПРД $\beta(t)$ доминирует над ПРД $\bar{\beta}(t)$ по коалиции $S \subseteq N$ в момент времени $\bar{t} \in [t_0, T]$ (и обозначать

доминирование через $\beta(t) \stackrel{S, \bar{t}}{>} \bar{\beta}(t)$, если выполняются неравенства:

$$\begin{aligned} \beta_i(\bar{t}) &> \bar{\beta}_i(\bar{t}), \quad i \in S, \\ \sum_{i \in S} \beta_i(\bar{t}) &\leq U(S; x^*(\bar{t}), T - \bar{t}). \end{aligned} \quad (1.18)$$

Определение 1.2.0.2 Будем говорить, что ПРД $\beta(t)$ не доминируема, если в любой момент времени $\bar{t} \in [t_0, T]$ не существует $\bar{\beta}(t)$, которая доминирует над $\beta(t)$ по коалиции $S \subseteq N$:

$$\bar{\beta}(t) \not\stackrel{S, \bar{t}}{>} \beta(t), \quad \forall \bar{\beta}(t), S. \quad (1.19)$$

1.2.2 Построение ПРД-ядра

В статье [101] автор диссертации и его соавторы сначала представили и рассмотрели подмножество множества дележей в кооперативной дифференциальной игре, которое получило название подъядра. Это подмножество было построено с использованием набора процедур распределения дележа, удовлетворяющих системе неравенств и равенств. Такой подход не является классическим для теории дифференциальных игр, поскольку он использует ПРД для дележа, а не наоборот. В статье [102] автор диссертации и его соавторы переопределили понятие подъядра для динамического случая, назвав его ПРД-ядром, а также сформулировали необходимые и достаточные условия существования ПРД-ядра вдоль кооперативной траектории игры. Концепция решения для ПРД-ядра определяется с помощью понятия доминирования по ПРД и свойств или аксиом динамической устойчивости; см. ниже. Доказаны необходимые и достаточные условия для принадлежности динамического дележа ПРД-ядру, позволяющие представить ПРД-ядро в виде системы неравенств, введенных в статьях [100; 101].

Предположим, что игроки в игре $\Gamma(x_0, T - t_0)$ согласовали правило распределения для суммарного выигрыша гранд-коалиции N (дележ $\xi(x_0, T - t_0)$), используя кооперативное решение ПРД-ядра:

Определение 1.2.0.3 Динамическим ПРД-ядром $IDP - core(x^*(t), T - t)$ вдоль кооперативной траектории $x^*(t)$, $t \in [t_0, T]$ ($IDP - core(x_0, T - t_0)$)

будем называть решение в кооперативной дифференциальной игре $\Gamma(x^*(t), T - t)$ (соответственно, в игре $\Gamma(x_0, T - t_0)$), которое включает в себя все динамически устойчивые дележи, сгенерированные недоминируемыми ПРД $\beta(\tau)$, $\tau \in [t, T]$ (1.8), $t \in [t_0, T]$ (соответственно, ПРД $\beta(t)$, $t \in [t_0, T]$):

$$IDP - core(x^*(t), T - t) = \left\{ \xi(x^*(t), T - t) = \int_t^T \beta(\tau) d\tau : \right. \\ \left. \xi(x^*(t), T - t) \text{ и соответствующая ПРД } \beta(\tau) \text{ удовлетворяет (1.10),} \right. \\ \left. \nexists \bar{\beta}(t), S, \bar{t} : \bar{\beta}(t) \stackrel{S, \bar{t}}{>} \beta(t) \right\}. \quad (1.20)$$

Отметим, что $IDP - core(x^*(t), T - t)$ включает такие дележи из S -ядра $C(x^*(t), T - t)$ кооперативной игры $\Gamma(x_0, T - t_0)$, для которых существует соответствующая недоминируемая ПРД, и эта ПРД генерирует динамический дележ, принадлежащий $C(x^*(t), T - t)$ для каждого $t \in [t_0, T]$.

Теорема 1.2.1 Пусть $C(x^*(t), T - t)$ не пусто для любого $t \in [t_0, T]$. Динамический дележ $\xi(x^*(t), T - t)$ в кооперативной дифференциальной игре $\Gamma(x_0, T - t_0)$ принадлежит динамическому ПРД-ядру $IDP - core(x^*(t), T - t)$ тогда и только тогда, когда для соответствующей $\beta(t) = \beta(t, \xi(x_0, T - t_0))$ выполнены следующие условия $\forall t \in [t_0, T]$:

$$\sum_{i \in S} \beta_i(t) \geq U(S; x^*(t), T - t), \quad \forall S \subset N, \quad (1.21)$$

$$\sum_{i \in N} \beta_i(t) = U(N; x^*(t), T - t). \quad (1.22)$$

Доказательство Достаточность. Пусть для соответствующей ПРД $\beta(t) = \beta(t, \xi(x_0, T - t_0))$ условия (1.21)–(1.22) выполняются при любом $t \in [t_0, T]$. Интегрируя (1.21)–(1.22) на данном интервале, получаем:

$$\sum_{i \in S} \xi(x^*(t), T - t) \geq V(S; x^*(t), T - t), \quad \forall S \subset N, \\ \sum_{i \in N} \xi(x^*(t), T - t) = V(N; x^*(t), T - t). \quad (1.23)$$

Следовательно, дележ имеет вид

$$\xi(x_0, T - t_0) = \int_{t_0}^T \beta(t) dt \in C(x^*(t_0), T - t_0) = C(x_0, T - t_0).$$

Покажем, что $\beta(t) = \beta(t, \xi(x_0, T - t_0))$ недоминируем. Докажем от противного: предположим, что для некоторого $\bar{S} \subset N$ существуют такие $\bar{t} \in [t_0, T]$ и $\bar{\beta}(t)$, что:

$$\begin{aligned} \xi(x_0, T - t_0) &= \int_{t_0}^T \bar{\beta}(t) dt \in L(x_0, T - t_0), \\ \bar{\beta}_i(\bar{t}) &> \beta_i(\bar{t}), \quad i \in \bar{S}, \\ \sum_{i \in \bar{S}} \bar{\beta}_i(\bar{t}) &\leq U(\bar{S}; x^*(\bar{t}), T - \bar{t}). \end{aligned} \quad (1.24)$$

Следовательно,

$$\sum_{i \in \bar{S}} \beta_i(\bar{t}) < U(\bar{S}; x^*(\bar{t}), T - \bar{t}).$$

Это неравенство противоречит (1.21). Таким образом, $\beta(t) = \beta(t, \xi(x_0, T - t_0))$ является недоминируемым на временном интервале $[t_0, T]$. Обращаем внимание, что $\beta(t)$ является недоминируемым на любом временном подынтервале $[\tau, T]$, $\tau \in [t_0, T]$ в подыгре $\Gamma(x^*(\tau), T - \tau)$.

Для $S = \{i\}$ условие (1.21) представляется в виде

$$\beta_i(t) \geq U(\{i\}; x^*(t), T - t), \quad i \in N.$$

Интегрируя эти неравенства и равенство (1.22) и учитывая (1.9), (1.17), получаем:

$$\begin{aligned} \xi_i(x^*(t), T - t) &= \int_t^T \beta_i(\tau) d\tau \geq V(\{i\}; x^*(t), T - t), \quad i \in N, \\ \sum_{i \in N} \xi_i(x^*(t), T - t) &= \sum_{i \in N} \int_t^T \beta_i(\tau) d\tau = V(N; x^*(t), T - t). \end{aligned} \quad (1.25)$$

Таким образом, динамический выигрыш

$$\xi(x^*(t), T - t) = \int_t^T \beta(\tau) d\tau$$

является дележом в текущей игре $\Gamma(x^*(t), T - t)$, сгенерированным соответствующей ПРД $\beta(t) = \beta(t, \xi(x_0, T - t_0))$.

В силу непустоты $C(x^*(t), T - t)$ для любого $t \in [t_0, T]$, ПРД $\beta(t) = \beta(t, \xi(x_0, T - t_0))$, удовлетворяющая (1.21)–(1.22), генерирует вектор выигрыша

$$\xi(x^*(t), T - t) = \int_t^T \beta(\tau) d\tau,$$

принадлежащий $C(x^*(t), T - t)$. Итак, $\xi(x^*(t), T - t)$ удовлетворяет (1.10). Следовательно, дележ $\xi(x_0, T - t_0)$ динамически устойчив и лежит в ПРД-ядре $IDP - core(x_0, T - t_0)$.

Необходимость. Пусть дележ $\xi(x_0, T - t_0)$ в кооперативной дифференциальной игре $\Gamma(x_0, T - t_0)$ принадлежит ПРД-ядру $IDP - core(x_0, T - t_0)$. Следовательно, согласно Определению 1.2.0.3 дележ $\xi(x_0, T - t_0)$, сгенерированный недоминируемой ПРД $\beta(t) = \beta(t, \xi(x_0, T - t_0))$, принадлежит $C(x_0, T - t_0)$ и является динамически устойчивым. В соответствии с Определением 1.2.0.3, динамическая устойчивость дележа $\xi(x_0, T - t_0)$ означает, что существует ПРД $\beta(t)$, такая что

$$\xi(x^*(t), T - t) = \int_t^T \beta(\tau) d\tau \in C(x^*(t), T - t), \quad t \in [t_0, T].$$

Покажем, что это включение имеет место для недоминируемой ПРД $\beta(t) = \beta(t, \xi(x_0, T - t_0))$, соответствующей дележу $\xi(x_0, T - t_0)$. Допустим, это не так. Тогда существуют коалиция $\bar{S} \subset N$ и некоторый момент времени $\bar{t} \in [t_0, T]$ такие, что выполняется неравенство

$$\sum_{i \in \bar{S}} \int_{\bar{t}}^T \beta(t) dt < V(\bar{S}; x^*(\bar{t}), T - \bar{t}). \quad (1.26)$$

Обращаем внимание, что $\beta(\bar{t}) = \beta(\bar{t}, \xi(x_0, T - t_0))$ принадлежит множеству недоминируемых дележей в кооперативной игре с характеристической функцией

$U(S; x^*(\bar{t}), T - \bar{t})$, $S \subseteq N$. Следовательно, согласно Теореме 1.1.1 для $\bar{S} \subset N$ должно выполняться следующее условие:

$$\sum_{i \in \bar{S}} \beta_i(\bar{t}) dt \geq U(\bar{S}; x^*(\bar{t}), T - \bar{t}). \quad (1.27)$$

Интегрируя данное неравенство на временном интервале $[\bar{t}, T]$, получаем

$$\sum_{i \in \bar{S}} \int_{\bar{t}}^T \beta_i(t) dt \geq V(\bar{S}; x^*(\bar{t}), T - \bar{t}), \quad (1.28)$$

что противоречит (1.27). Теорема доказана. \square

Утверждение 1.2.0.1 *Если С-ядро $C(x^*(t), T - t)$ пусто для некоторого $t = \bar{t} \in [t_0, T]$, то ПРД-ядро $IDP - core(x^*(t_0), T - t_0)$ также пусто.*

Доказательство Предположим, что можно найти динамически устойчивый дележ $\xi(x_0, T - t_0) \in IDP - core(x^*(t_0), T - t_0)$ и соответствующую недоминируемую ПРД $\beta(t)$, которая удовлетворяет (1.21) и (1.22). Интегрируя неравенства (1.21) и (1.22), получаем

$$\begin{aligned} \sum_{i \in S} \int_{\bar{t}}^T \beta_i(\tau) d\tau &\geq V(S; x^*(\bar{t}), T - \bar{t}), \quad S \subset N, \\ \sum_{i \in N} \int_{\bar{t}}^T \beta_i(\tau) d\tau &= V(N; x^*(\bar{t}), T - \bar{t}). \end{aligned} \quad (1.29)$$

Таким образом,

$$\xi(x^*(\bar{t}), T - \bar{t}) = \int_{\bar{t}}^T \beta(\tau) d\tau \in C(x^*(\bar{t}), T - \bar{t}). \quad (1.30)$$

Это противоречит пустоте С-ядра $C(x^*(\bar{t}), T - \bar{t})$. Следовательно, ПРД-ядро $IDP - core(x^*(t_0), T - t_0)$ является пустым. Предложение доказано. \square

Утверждение 1.2.0.2 *Если С-ядро $C(x^*(t), T - t)$ непусто для любого $t \in [t_0, T]$, то ПРД-ядро $IDP - core(x^*(t_0), T - t_0) = C(x^*(t_0), T - t_0)$.*

Доказательство Рассмотрим дележ $\xi(x_0, T - t_0) \in C(x^*(t_0), T - t_0)$, который не принадлежит ПРД-ядру $IDP - core(x^*(t_0), T - t_0)$. То есть $\xi(x_0, T - t_0)$ принадлежит $C(x^*(t_0), T - t_0)$, но является динамически неустойчивым. Согласно определению динамической устойчивости, дележ $\xi(x_0, T - t_0)$ динамически неустойчив, если не существует ПРД $\beta(t) = \beta(t, \xi(x_0, T - t_0))$, такой что для любого $t \in [t_0, T]$ динамический дележ $\xi(x^*(t), T - t)$, сгенерированный этой ПРД, принадлежит C -ядру $C(x^*(t), T - t)$.

Но, как следует из Теоремы 1.2.1, если соответствующая ПРД $\beta(t) = \beta(t, \xi(x_0, T - t_0))$ удовлетворяет условиям (1.21) и (1.22), то $\xi(x_0, T - t_0)$ принадлежит ПРД-ядру $IDP - core(x_0, T - t_0)$. То есть дележ $\xi(x_0, T - t_0)$ является динамически устойчивым по Определению 1.2.0.3. Другими словами, для любого $t \in [t_0, T]$ справедливо включение

$$\xi(x^*(t), T - t) = \int_t^T \beta(\tau) d\tau \in C(x^*(t), T - t).$$

Предложение доказано. □

Замечание 1 Согласно Предложению 1.2.0.2, если C -ядро не пусто, то ПРД-ядро и C -ядро совпадают в текущей игре или, что эквивалентно, дележи из C -ядра совпадают с дележами из ПРД-ядра. Система неравенств (1.21)–(1.22) позволяет из множества процедур распределения дележа выделить подмножество, в котором обеспечивается динамическая устойчивость и свойство недоминирования по ПРД для дележей из C -ядра. Для других подмножеств множества ПРД, такой результат, вообще говоря, неверен. Также обращаем внимание, что множество процедур распределения дележа (1.21)–(1.22) может быть пустым. В следующем разделе рассмотрим подход к проверке непустоты и построению ПРД из ПРД-ядра.

Пусть характеристическая функция $V(S; x^*(t), T - t)$, $t \in [t_0, T]$, определена некоторым подходящим образом (например, как в [146]). Предположим, что это строго монотонно убывающая функция для любого $t \in [t_0, T]$, убывающая быстрее линейного закона. Построим C -ядро $C(x_0, T - t_0)$ для начального момента времени $t = t_0$. Затем выберем некоторый дележ из C -ядра, $\xi(x_0, T - t_0) \in C(x_0, T - t_0)$. Согласно Определению 1.1.0.2, выберем соответствующую процедуру распределения дележа $\beta(t) = \beta(t, \xi(x_0, T - t_0))$

(ПРД) для $\xi(x_0, T - t_0)$ подходящим образом. Как следует из Предложения 1.2.0.2, для дележей $\xi(x_0, T - t_0) \in IDP - core(x^*(t_0), T - t_0) = C(x^*(t_0), T - t_0)$ мы всегда можем найти ПРД $\beta'(t)$, удовлетворяющую условиям (1.21)–(1.22), такую что

$$\int_{t_0}^T \beta'(t) dt = \xi(x_0, T - t_0).$$

С другой стороны, условия (1.21)–(1.22) могут не выполняться для $\beta(t) = \beta(t, \xi(x_0, T - t_0))$. Например, в случае, если $\beta(t) = \beta(t, \xi(x_0, T - t_0))$ выбрана следующим образом:

$$\begin{aligned} \beta(t) &= c = const, \quad t \in [t_0, t'], \quad (t' \in (t_0, T)), \\ \beta(t) &= \frac{\xi(x_0, T - t_0) - c(t' - t_0)}{T - t'}, \quad t \in [t', T], \end{aligned}$$

а производная $V(S; x^*(t), T - t)$

$$U(S; x^*(t), T - t) > 0, \quad t \in [t_0, T].$$

Этот случай подробно рассмотрен для модельного примера в Разделе 5.

1.2.3 Свойства ПРД-ядра

Здесь представлены несколько свойств, связанных с динамической устойчивостью ПРД-ядра.

Утверждение 1.2.0.3 Если C -ядро $C(x^*(t), T - t)$ непусто для любого $t \in [t_0, T]$, то все дележи из $C(x^*(t_0), T - t_0)$ динамически устойчивы.

Доказательство Рассмотрим дележ $\xi(x_0, T - t_0)$ из C -ядра $C(x^*(t_0), T - t_0)$. Согласно Предложению 1.2.0.2, $IDP - core(x^*(t_0), T - t_0) = C(x^*(t_0), T - t_0)$ и, следовательно, $\xi(x_0, T - t_0) \in IDP - core(x^*(t_0), T - t_0)$. Из определения ПРД-ядра следует, что $\xi(x_0, T - t_0)$ является динамически устойчивым. Предложение доказано. \square

Ниже показано, что ПРД-ядро как кооперативное решение является сильно динамически устойчивым, а не только построено с использованием аксиомы

динамической устойчивости.

В статье [101] показано, что ПРД-ядро обладает свойством сильной динамической устойчивости.

Определение 1.2.0.4 Кооперативное решение $IDP - core(x_0, T - t_0)$ сильно динамически устойчиво в игре $\Gamma(x_0, T - t_0)$, если

$$1 \quad IDP - core(x^*(t), T - t) \neq \emptyset, \quad \forall t \in [t_0, T],$$

$$2 \quad \forall \xi(x_0, T - t_0) \in IDP - core(x_0, T - t_0) \text{ существует ПРД } \beta(\tau), \tau \in [t_0, T],$$

такая что

$$\xi(x_0, T - t_0) = \int_{t_0}^T \beta(\tau) d\tau, \text{ и выполнены следующие условия:}$$

$$IDP - core(x_0, T - t_0) \supseteq \int_{t_0}^t \beta(\tau) d\tau \oplus IDP - core(x^*(t), T - t), \quad \forall t \in [t_0, T].$$

Здесь операция \oplus определяется следующим образом: для $a \in R^n$ и $D \subseteq R^n$, $a \oplus D = \{a + d : d \in D\}$.

Свойство сильной динамической устойчивости ПРД-ядра имеет следующий смысл: когда игра развивается по кооперативной траектории $x^*(t)$, для любого дележа $\xi(x_0, T - t_0) \in IDP - core(x_0, T - t_0)$ существует такая ПРД, что отклонение от выбранного дележа в момент времени t к другому дележу $\hat{\xi}(x^*(t), T - t) \in IDP - core(x^*(t), T - t)$ приводит к общему выигрышу в игре $\Gamma(x_0, T - t_0)$, соответствующему некоторому дележу $\bar{\xi}(x_0, T - t_0)$ из исходного ПРД-ядра $IDP - core(x_0, T - t_0)$. Следовательно, если кооперативное решение сильно динамически устойчиво, то его пересмотр во время игры приводит к одному из решений исходного ПРД-ядра и, следовательно, пересмотр нецелесообразен.

Сконструированное ПРД-ядро $IDP - core(x_0, T - t_0)$ является сильно динамически устойчивым кооперативным решением.

Теорема 1.2.2 Предположим, что $C(x^*(t), T - t) \neq \emptyset$ и ПРД-ядро $\bar{C}(x^*(t), T - t) \neq \emptyset, \forall t \in [t_0, T]$. Тогда ПРД-ядро $IDP - core(x_0, T - t_0)$ является сильно динамически устойчивым кооперативным решением в игре $\Gamma(x_0, - t_0)$.

Доказательство Поскольку S -ядро и ПРД-ядро непусты, для любого дележа $\xi(x_0, T - t_0) \in \overline{C}(x_0, T - t_0) \subseteq C(x_0, T - t_0)$ существует ПРД $\beta_i(t)$, такая что $\xi_i(x_0, T - t_0) = \int_{t_0}^T \beta_i(t) dt, i \in N$. Игра развивается вдоль кооперативной траектории $x^*(t), t \in [t_0, T]$, на которой в каждый момент времени игроки находятся в текущих подыграх $\Gamma_v(x^*(t), T - t)$.

Пусть в начале игры $\Gamma_v(x_0, T - t_0)$ игроки выбрали дележ $\xi(x_0, T - t_0) \in \overline{C}(x_0, T - t_0)$. Предположим, что в момент времени $t_{br}, t_{br} \in [t_0, T]$, происходит отклонение от дележа $\xi(x^*(t), T - t)$ в пользу другого дележа $\tilde{\xi}(x^*(t_{br}), T - t_{br})$ из ПРД-ядра $\overline{C}(x^*(t_{br}), T - t_{br})$. Покажем, что результирующий дележ $\hat{\xi}(x_0, T - t_0) = \int_{t_0}^{t_{br}} \beta(t) dt + \tilde{\xi}(x^*(t_{br}), T - t_{br})$ также принадлежит ПРД-ядру $\overline{C}(x_0, T - t_0)$. Другими словами, докажем существование ПРД $\hat{\beta}(t)$, такой что $\int_{t_0}^T \hat{\beta}(t) dt = \hat{\xi}(x_0, T - t_0)$ и $\hat{\beta}(t) \in B(t)$ удовлетворяет условиям (1.21)–(1.22). По определению $\overline{C}(x_0, T - t_0)$, дележам $\xi(x_0, T - t_0)$ и $\tilde{\xi}(x^*(t_{br}), T - t_{br})$ соответствуют ПРД $\beta(t) \in B(t), t \in [t_0, T]$, и $\tilde{\beta}(t) \in B(t), t \in [t_{br}, T]$, (1.21)–(1.22).

Тогда дележ $\hat{\beta}(t)$ имеет вид

$$\hat{\beta}(t) = \begin{cases} \beta(t), & t \in [t_0, t_{br}], \\ \tilde{\beta}(t), & t \in (t_{br}, T]. \end{cases}$$

Чтобы установить включение $\hat{\beta}(t) \in B(t), \forall t \in [t_0, T]$, необходимо доказать, что функция $\hat{\beta}(t)$ интегрируема. Функции $\beta(t)$ и $\tilde{\beta}(t)$ интегрируемы по построению, поскольку $\beta(t), \tilde{\beta}(t) \in B(t)$ соответственно на временных интервалах $[t_0, T]$ и $[t_{br}, T]$. Согласно определению интегрируемых функций $\beta(t), \tilde{\beta}(t)$, существует предел интегрируемых сумм на временных интервалах $[t_0, t_{br}]$ и $(t_{br}, T]$. Следовательно, сумма этих двух интегрируемых сумм также сходится. Учитывая $\hat{\beta}(t) \in B(t), \forall t \in [t_0, T]$, по определению множества $\overline{C}(x_0, T - t_0)$ мы приходим к включению $\hat{\xi}(x_0, T - t_0) \in \overline{C}(x_0, T - t_0)$. Теорема доказана. \square

1.3 Непустота ПРД-ядра на основе подхода линейного программирования

Данный раздел посвящен подходу линейного программирования, описанному в разделе 2.4, применительно к свойству непустоты С-ядра. ПРД-ядро может быть построено с использованием системы линейных ограничений для процедур распределения дележа. Эти ограничения определяются для каждого момента времени в игре. Из непустоты множества, описываемого этими ограничениями, следует, непустота ПРД-ядра.

Рассмотрим следующую задачу линейного программирования для фиксированного t :

$$\begin{aligned} \sum_{i \in N} \beta_i &\longrightarrow \min \\ \sum_{i \in S} \beta_i &\geq U(S; x^*(t), T - t), \quad \forall S \subseteq N, S \neq N, S \neq \emptyset. \end{aligned} \quad (1.31)$$

Предположим, что $\beta_i^0 = (\beta_1^0, \dots, \beta_n^0)$ — оптимальное решение задачи линейного программирования (1.31) с фиксированным значением t . Множество оптимальных решений задачи (1.31) обозначим через Y^0 .

Тогда верна следующая теорема:

Теорема 1.3.1 *Множество ПРД, удовлетворяющих условиям (1.21), (1.22), $t \in [t_0, T)$, непусто тогда и только тогда, когда $\forall t \in [t_0, T)$ справедливо неравенство*

$$\sum_{i \in N} \beta_i^0 \leq U(N; x^*(t), T - t), \quad (1.32)$$

где $\beta^0 \in Y^0$ — любое решение задачи линейного программирования (1.31).

Доказательство Начнем доказательство с достаточного условия. Предположим, что условие (1.32) выполнено. Тогда, в силу (1.31), для любого $t \in [t_0, T]$ существует $\hat{\beta}^0$, такое что для

$$\beta_i = \beta_i^0 + \frac{U(N; x^*(t), T - t)}{n} - \frac{\sum_{i \in N} \beta_i^0}{n}, \quad i \in N, \quad (1.33)$$

выполняются условия (1.21) и (1.22) при любом фиксированном $t \in [t_0, T]$. Если это так, мы можем составить интегрируемую функцию времени $\hat{\beta}^0(t)$, для которой будут верны условия (1.21) и (1.22).

Доказательство необходимости. Если ПРД-ядро непусто, то существует хотя бы одна интегрируемая функция $\beta(t)$, удовлетворяющая условиям (1.21) и (1.22). В результате для решения (1.31) должно выполняться условие (1.32). \square

1.4 Дифференциальная игровая модель добычи ресурсов

Рассмотрим теоретико-игровую модель добычи невозобновляемых ресурсов с асимметричными игроками [57; 168]. Количество ресурса зависит от интенсивностей добычи, которые выбирают игроки. В игре участвуют n асимметричных игроков, функции полезности которых зависят от текущего количества ресурсов и интенсивностей добычи. Другой пример похожей модели можно найти в статье [102].

Обозначим через $x(t) \in R^1$ количество ресурса в момент времени t и через $u_i(t, x)$ интенсивность добычи ресурсов, выбранную игроком i в момент времени t . В качестве класса стратегий будем рассматривать позиционные стратегии (с обратной связью), представляющие собой функции времени t и состояния x . Предполагаем, что из $u_i(t, x) \geq 0 \forall t$ и $x(t) = 0$ следует $u_i(t, x) = 0$. Количество ресурса $x(t)$ как функция t зависит от $u_i(t, x)$ следующим образом:

$$\begin{aligned} \dot{x} &= - \sum_{i=1}^n a_i u_i(t, x), \quad a_i > 0, \quad i = 1, \dots, n, \\ x(t_0) &= x_0. \end{aligned} \tag{1.34}$$

Функция выигрыша описывает доход игрока i :

$$K_i(x_0, T - t_0) = \int_{t_0}^T \log(u_i(\tau, x)) d\tau, \quad i = 1, \dots, n. \tag{1.35}$$

1.4.1 Кооперативные стратегии и кооперативная траектория

Рассмотрим кооперативную версию игры по добыче невозобновляемых ресурсов [168]. Здесь игроки объединяются в гранд-коалицию и максимизируют суммарный выигрыш, действуя как один игрок. Соответствующая задача оптимального управления формализуется следующим образом:

$$\sum_{i=1}^n K_i(x_0, T - t_0) = \sum_{i=1}^n \int_{t_0}^T \log(u_i(\tau, x)) d\tau \rightarrow \max_{u_i, i=\overline{1, n}} \quad (1.36)$$

при ограничениях вида

$$\begin{aligned} \dot{x} &= - \sum_{i=1}^n a_i u_i(t, x), \\ x(t_0) &= x_0 > 0, \\ u(t, x) &\geq 0. \end{aligned} \quad (1.37)$$

Для решения задачи оптимизации (1.36), (1.37) используем принцип динамического программирования, предложенный Беллманом. Функцию Беллмана определим как максимальное значение суммарного выигрыша игроков (1.35) в подыгре $\Gamma(x, T - t)$, начиная с момента времени t в позиции x :

$$W(t, x) = \max_{u_i, i=\overline{1, n}} \left\{ \sum_{i=1}^n K_i(x, T - t) \right\} = \max_{u_i, i=\overline{1, n}} \left\{ \sum_{i=1}^n \int_t^T \log u(\tau, x) d\tau \right\} \quad (1.38)$$

в силу (1.37) при $x_0 = x$ и $t_0 = t$.

Доказано, что если существует непрерывно дифференцируемая функция $W(t, x)$, удовлетворяющая уравнению Гамильтона–Якоби–Беллмана

$$\begin{aligned} -W_t(t, x) &= \max_{u_i, i=\overline{1, n}} \left\{ \sum_{i=1}^n \log u_i(t, x) - W_x(t, x) \left(\sum_{i=1}^n a_i u_i(t, x) \right) \right\}, \\ \lim_{t \rightarrow T-0} W(t, x) &= 0, \end{aligned} \quad (1.39)$$

то стратегии $u_i^*(t, x)$, полученные при максимизации правой части (1.39), доставляют максимум функционалу в задаче оптимизации (1.36), (1.37).

Из условия оптимальности первого порядка для (1.39) получаем:

$$u_i^* = \frac{1}{a_i W_x(t, x)}.$$

Подстановка этой формулы в (1.39) дает

$$\begin{aligned} W_t(t, x) &= n \log W_x(t, x) + \log A^{[N]} + n, & A^{[N]} &= \prod_{i=1}^n a_i \\ \lim_{t \rightarrow T-0} W(t, x) &= 0. \end{aligned} \quad (1.40)$$

Рассмотрим функцию Беллмана вида

$$W(t, x) = A(t) \log x + B(t).$$

Подставляя это выражение в (1.39) имеем:

$$\begin{aligned} \dot{A} \log x + \dot{B} &= n \log A - n \log x + \log A^{[N]} + n, \\ \lim_{t \rightarrow T-0} A(t) &= \lim_{t \rightarrow T-0} B(t) = 0. \end{aligned} \quad (1.41)$$

Решением (1.41) являются функции

$$\begin{aligned} A(t) &= n(T - t), \\ B(t) &= -(T - t) \left(\log A^{[N]} + n \log n(T - t) \right). \end{aligned} \quad (1.42)$$

При подстановке $A(t)$ и $B(t)$ в функцию Беллмана получаем

$$W(t, x) = n(T - t) \log \frac{x}{n(T - t)} - (T - t) \log A^{[N]}, \quad t \in [t_0, T]. \quad (1.43)$$

Соответствующее оптимальное управление (кооперативная стратегия) имеет вид

$$u_i^*(t, x) = \frac{1}{a_i W_x(t, x)} = \frac{x}{a_i n(T - t)}, \quad t \in [t_0, T]. \quad (1.44)$$

Подставляя найденное оптимальное управление в уравнения движения (1.37), приходим к дифференциальному уравнению для траектории, соответствующей

оптимальному управлению:

$$\begin{aligned}\dot{x} &= -\frac{x}{T-t}, \\ x(t_0) &= x_0.\end{aligned}\tag{1.45}$$

Решение имеет вид

$$x^*(t) = x_0 \frac{T-t}{T-t_0}, \quad t \in [t_0, T].\tag{1.46}$$

Будем называть траекторию $x^*(t)$ и стратегию (управление) $u^*(t, x)$ кооперативными.

Чтобы определить значение максимального суммарного выигрыша игроков, соответствующее задаче оптимизации (1.36), (1.37) в подыгре вдоль кооперативной траектории $x^*(t)$ (1.46), необходимо заменить выражение для кооперативной траектории выражением для функции Беллмана (1.43):

$$W(t, x^*(t)) = n(T-t) \log \frac{x_0}{n(T-t_0)} - (T-t) \log A^{[N]}, \quad t \in [t_0, T].\tag{1.47}$$

1.4.2 Характеристическая функция

Чтобы построить правило распределения максимального суммарного выигрыша между игроками, необходимо определить характеристическую функцию для каждой коалиции $S \subseteq N$:

$$V(S; x, T-t) = \begin{cases} \sum_{i=1}^n K_i(x, T-t), & S = N \\ W_S(t, x), & S \subset N, \\ 0, & S = \emptyset, \end{cases}\tag{1.48}$$

где $W_S(t, x)$ определяется как максимальный суммарный выигрыш коалиции S при условии, что игроки из коалиции $N \setminus S$ используют стратегии из фиксированного допустимого равновесия по Нэшу $u^{NE} = (u_1^{NE}, \dots, u_n^{NE})$ (п.1.1.1) в начальной игре.

Можно показать, что в случае некооперативной игры равновесные по Нэшу

стратегии имеют вид

$$u_i^{NE}(t, x) = \frac{x}{a_i(T-t)}, \quad i \in N. \quad (1.49)$$

Введем функцию Беллмана $W_S(t, x)$ как максимальный суммарный выигрыш игроков из коалиции S в подыгре $\Gamma(x, T-t)$, начиная с момента времени t в позиции x :

$$W_S(t, x) = \max_{u_i, i \in S} \sum_{i \in S} \left\{ \int_t^T \log u_i d\tau \right\} \quad (1.50)$$

$$\text{subject to } \dot{x}(\tau) = - \sum_{i \in N} a_i u_i \quad (1.51)$$

$$u_i = u_i^{NE}, \quad i \in N \setminus S. \quad (1.52)$$

Уравнение Гамильтона–Якоби–Беллмана для этой задачи имеет следующий вид:

$$-\frac{\partial W_S(t, x)}{\partial t} = \max_{u_i, i \in S} \left\{ \sum_{i \in S} \log u_i(t, x) - \frac{\partial W_S(t, x)}{\partial x} \left(\sum_{j=1}^n a_j u_j(t, x) \right) \right\}, \quad (1.53)$$

$$\lim_{t \rightarrow T-0} W_S(t, x) = 0.$$

Из условия оптимальности первого порядка для (1.53) получаем:

$$u_i^* = \frac{1}{a_i \frac{\partial W_S(t, x)}{\partial x}}. \quad (1.54)$$

Подстановка этой формулы в (1.53) дает

$$\frac{\partial W_S}{\partial t} = k \log \frac{\partial W_S}{\partial x} + \log A^{[S]} + k + \frac{\partial W_S}{\partial x} \sum_{j \in N \setminus S} \frac{x}{T-t}, \quad A^{[S]} = \prod_{i=1}^n a_i$$

$$\lim_{t \rightarrow T-0} W_S(t, x) = 0, \quad (1.55)$$

где $k = |S|$ и $n = |N|$. Рассмотрим функцию Беллмана вида

$$W_S(t, x) = A(t) \log x + B(t).$$

Подставляя это выражение в (1.55), имеем:

$$\begin{aligned} \dot{A} \log x + \dot{B} &= k \log A - k \log x + \log A^{[S]} + k + (n - k) \frac{A}{T-t}, & (1.56) \\ \lim_{t \rightarrow T-0} A(t) &= \lim_{t \rightarrow T-0} B(t) = 0. \end{aligned}$$

Решением (1.56) являются функции

$$\begin{aligned} A(t) &= k(T - t), \\ B(t) &= -k(T - t) \left(\frac{\log A^{[S]}}{k} + \log k(T - t) + n - k \right). \end{aligned} \quad (1.57)$$

Тогда решение оптимизационной задачи (1.50) принимает вид

$$W_S(t, x) = k(T - t) \left[\log \frac{x}{T - t} - \log k - \frac{\log A^{[S]}}{k} - n + k \right]. \quad (1.58)$$

Согласно определению получаем следующую характеристическую функцию для коалиции $S \neq N$:

$$V(S, x, T - t) = W_S(t, x).$$

Чтобы распределить максимальный суммарный выигрыш игроков (1.47) между ними вдоль кооперативной траектории $x^*(t)$ (1.46), а именно для подыгры, начинающейся в момент времени t на кооперативной траектории $x^*(t)$ (1.46), необходимо определить характеристическую функцию вдоль кооперативной траектории. Подставим выражение для $x^*(t)$ (1.46) в выражение для характеристической функции $V(S, T - t, x)$, $S \subset N$ (4.188). В результате имеем:

$$W_S(t, x^*(t)) = k(T - t) \left[\log \frac{x_0}{T - t_0} - \log k - \frac{\log A^{[S]}}{k} - n + k \right]. \quad (1.59)$$

В случае $S = N$, характеристическая функция вычисляется в соответствии с (1.47).

1.4.3 ПРД-ядро

Предположим, что все игроки объединяются в гранд-коалицию N . Тогда они могут гарантировать себе суммарный выигрыш $V(N; x^*(t), T - t)$. Чтобы распределить максимальный суммарный выигрыш между игроками, будем использовать понятие дележа $\xi(x, T - t)$. В частности, применим ПРД-ядро как кооперативное решение в игре. Согласно Теореме 1.2.1, ПРД-ядро можно построить с помощью условий для ПРД $\beta_i(t)$, $i \in N$:

$$\begin{aligned} \sum_{i \in S} \beta_i(t) &\geq -kt \left[\log \frac{x_0}{T - t_0} - \log k - \frac{\log A^{[S]}}{k} - n + k \right], \quad \forall S \subset N, \\ \sum_{i \in N} \beta_i(t) &= -nt \log \frac{x_0}{n(T - t_0)} - (T - t) \log A^{[N]}, \quad \forall t \in [t_0, T]. \end{aligned} \quad (1.60)$$

1.4.4 Непустота ПРД-ядра

Для исследования условий непустоты решим задачу линейного программирования, представленную в статье [102], для $t \in [t_0, T]$ с фиксированным шагом Δt :

$$\begin{aligned} \sum_{i \in N} \beta_i &\longrightarrow \min \\ \sum_{i \in S} \beta_i &\geq -kt \left[\log \frac{x_0}{T - t_0} - \log k - \frac{\log A^{[S]}}{k} - n + k \right], \quad \forall S \subset N. \end{aligned} \quad (1.61)$$

В результате вектор-функция $\beta^0 = (\beta_1^0, \dots, \beta_n^0)$ находится с помощью численных методов, а соответствующие условия (при которых ПРД-ядро непусто) имеют вид

$$\sum_{i \in N} \beta_i^0 \leq U(N; x^*(t), T - t). \quad (1.62)$$

Построим ПРД $\hat{\beta}^0(t)$ с помощью $\beta^0(t)$ и покажем, что она удовлетворяет условиям (1.21),(1.22):

$$\hat{\beta}_i^0(t) = \beta_i^0(t) + \frac{U(N; x^*(t), T - t) - \sum_{i \in N} \beta_i^0(t)}{n}. \quad (1.63)$$

1.4.5 С-ядро и ПРД-ядро

Согласно Теореме 1.2.1, соответствующий ПРД $\hat{\beta}^0(t)$ дележ

$$\xi(x_0, T - t_0) = \int_{t_0}^T \hat{\beta}^0(t) dt \quad (1.64)$$

принадлежит С-ядру $C(x_0, T - t_0)$, потому что для данных параметров ПРД $\hat{\beta}^0(t)$, $t \in [t_0, T]$, удовлетворяет условиям (1.21)–(1.22) или $\xi(x_0, T - t_0)$ принадлежит ПРД-ядру $IDP - core(x_0, T - t_0)$. Но при использовании С-ядра $C(x_0, T - t_0)$ вместо ПРД-ядра $IDP - core(x_0, T - t_0)$ в качестве кооперативного решения в игре, мы можем выбрать любую ПРД для дележа (1.64), например:

$$\begin{aligned} \beta(t) &= c = const, \quad t \in [t_0, t'], \quad (t' \in (t_0, T)), \\ \beta(t) &= \frac{\xi(x_0, T - t_0) - c(t' - t_0)}{T - t'}, \quad t \in [t', T]. \end{aligned} \quad (1.65)$$

Однако, она не обязательно удовлетворяет условиям (1.21)–(1.22) в какой-то момент времени и может оказаться доминируемой. Следовательно, соответствующий этой ПРД дележ $\xi(x_0, T - t_0)$ динамически неустойчив.

На рисунке 1 показано множество, заданное системой ограничений (1.21), (1.22). Здесь: сплошная линия — решение $\beta^0(t)$ соответствующей задачи линейного программирования (1.31) как функция времени, а пунктирная линия — ПРД $\hat{\beta}^0(t)$ и ПРД $\beta(t)$ (1.65), соответствующая дележу (1.64).

Функция $\hat{\beta}_i^0(t)$ удовлетворяет ограничениям (1.21), (1.22). Видно, что ПРД-ядро в этой игровой модели непусто, и условия (1.32) Теоремы 1.3.1 выполнены.

С помощью рисунка 2 можно проверить условия непустоты (1.32) Теоремы 1.3.1. Сплошной линией показана сумма значений $\beta_i^0(t)$, $i = \overline{1, 3}$:

$$S_{\beta^0}(t) = \beta_1^0(t) + \beta_2^0(t) + \beta_3^0(t). \quad (1.66)$$

Пунктирной линией показаны значения характеристической функции для

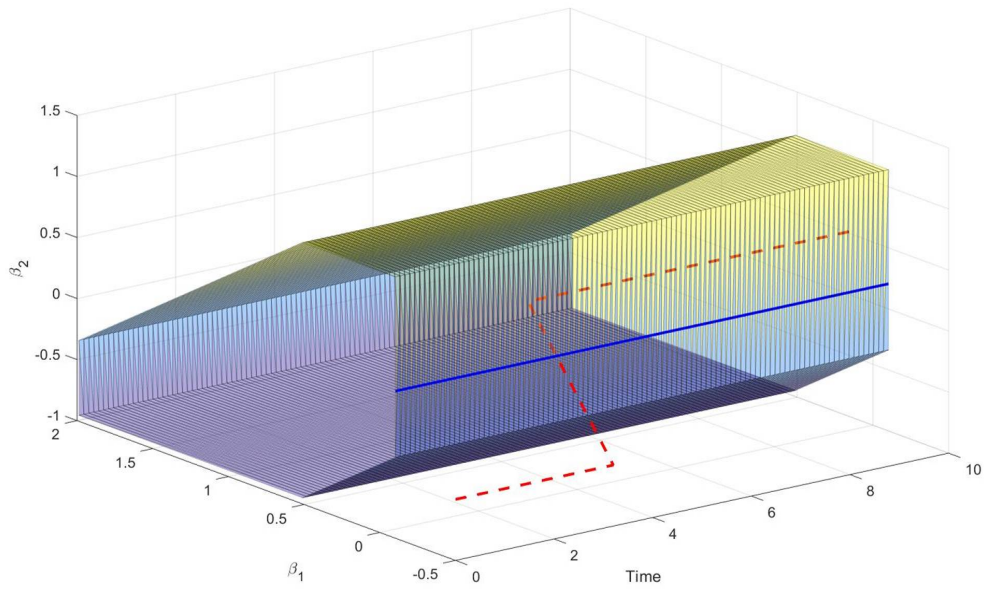


Рис. 1.1: Оси: β_1 , β_3 , t . β_2 можно найти из неравенства (1.22).

гранд-коалиции:

$$U(N; x^*(t), T - t) = U(\{1, 2, 3\}; x^*(t), T - t),$$

где $U(\{1, 2, 3\}; x^*(t), T - t)$ задана в (1.17). На рисунке 2 видно, что

$$S_{\beta^0}(t) \leq U(\{1, 2, 3\}; x^*(t), T - t) \quad \forall t \in [t_0, T].$$

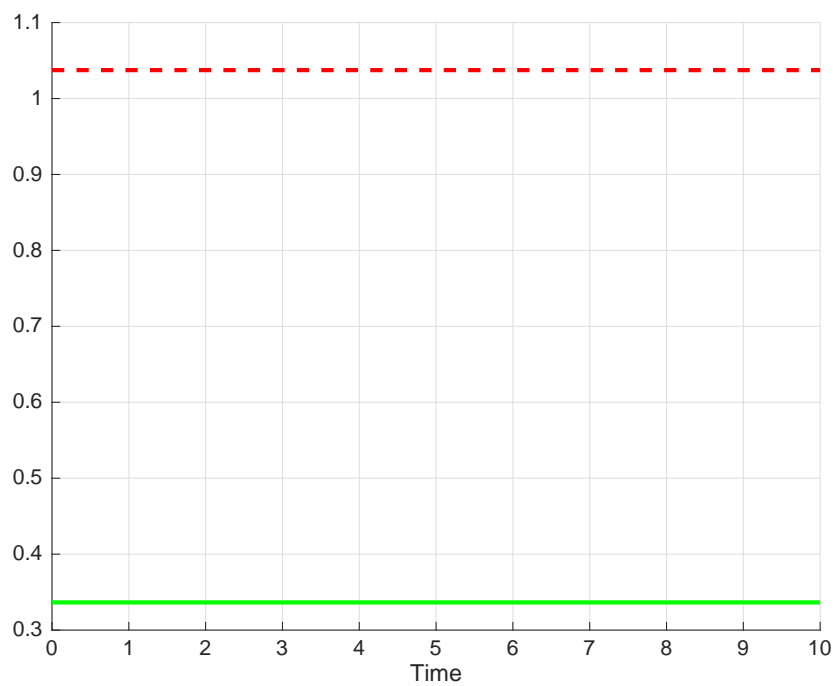


Рис. 1.2: $U(\{1, 2, 3\}; x^*(t), T - t)$ (1.17) — пунктирная линия, $S_{\beta^0}(t)$ (1.66) — сплошная линия.

Глава 2

Дискретные игры с динамическим обновлением информации

Информация играет одну из главных ролей в теории игр. Как правило, в классических дифференциальных играх рассматривается только случай полной информации: информация о структуре игры известна и не изменяется в течение всей игры. В этой главе изучается случай, когда при динамическом обновлении информации игроки используют только ограниченную информацию и обновляют ее на каждом шаге. Чтобы моделировать поведение игроков в таких условиях, вводится новый подход с динамическим обновлением информации.

2.1 Модель дискретной игры с динамическим обновлением информации

2.1.1 Исходная дискретная игра

Рассмотрим динамическую N -шаговую игру n лиц с ненулевой суммой и дискретным временем общего вида $\Gamma(x^0, N)$ [169] с начальным состоянием $x^0 \in R^m$. Динамика состояния игры описывается следующим разностным уравнением:

$$\begin{aligned}x_{k+1} &= f_k(x_k, u_1^k, u_2^k, \dots, u_n^k), \\x_1 &= x^0,\end{aligned}\tag{2.1}$$

где $u_i^k \in U^i \subset R^{m_i}$ — вектор управления игрока i на шаге k , $x_k \in X$ — состояние игры, а $k \in \{1, 2, \dots, N\} \equiv T$.

Функция выигрыша игрока i имеет вид

$$K_i(x^0, N, U) = \sum_{j=1}^N g_i^j[x_j, u_1^j, u_2^j, \dots, u_n^j, x_{j+1}] \left(\frac{1}{1+r} \right)^{j-1}, \text{ for } i \in \{1, 2, \dots, n\} \equiv I, \quad (2.2)$$

где r — коэффициент дисконтирования.

Эта игра — с полной информацией; однако в реальных игровых ситуациях игроки не имеют полной информации о структуре игры на временном интервале, на котором она определена. Далее используем подход с динамическим обновлением информации, чтобы моделировать поведение игроков в случае, когда информация об игре обновляется динамически.

2.1.2 Усеченная подыгра

Предположим, что на каждом шаге k игроки имеют полную информацию об уравнениях движения и функциях выигрыша на шагах $k + \bar{T}$, где \bar{T} — фиксированное значение, называемое информационным горизонтом. На шаге k информация об игре обновляется, и на шаге $k + 1$ игроки имеют полную информацию о структуре игры на временном интервале вплоть до шага $k + 1 + \bar{T}$.

Определение 2.1.0.1 Пусть $k = 1, \dots, N - \bar{T}$. Усеченная подыгра $\Gamma_k(x_{k,0}, k, k + \bar{T})$ определена на временном интервале $[k, k + \bar{T}]$. Уравнения движения и функция выигрыша усеченной подыгры на временном интервале $[k, k + \bar{T}]$ совпадают с уравнениями движения и функцией выигрыша исходной игры на том же временном интервале. Уравнения движения и начальное условие усеченной подыгры имеют следующий вид:

$$\begin{aligned} x_{k+1} &= f_k(x_k, u_1^k, u_2^k, \dots, u_n^k), \\ x_{1,0} &= x_1 = x^0. \end{aligned} \quad (2.3)$$

(Для усеченной подыгры $\Gamma_k(x_{k,0}, k, k + \bar{T})$, используется обозначение $x_{k,l} = x_{k+l}$, $u_i^{k,l} = u_i^{k+l}$, $l = \bar{0}, \bar{T}$.)

Функция выигрыша игрока i в усеченной подыгре $\Gamma_k(x_{k,0}, k, k + \bar{T})$ задается в виде

$$K_i(x_{k,0}, k, k + \bar{T}, U) = \sum_{j=0}^{\bar{T}} g_i^{k+j} [x_{k,j}, u_1^{k,j}, u_2^{k,j}, \dots, u_n^{k,j}, x_{k,j+1}] \left(\frac{1}{1+r} \right)^{k+j-1}, \quad (2.4)$$

для $i \in \{1, 2, \dots, n\} \equiv I$.

2.2 Некооперативная дискретная игра с динамическим обновлением информации

В последнее время разрабатывается теория некооперативных игр, касающаяся стратегического взаимодействия между несколькими лицами, принимающими решения, с целевыми функциями, зависящими от выбора всех игроков. Предполагается, что игроки не сотрудничают, и важно моделировать их поведение, используя некоторый принцип оптимальности. В области некооперативной теории игр получено много результатов; общее описание можно найти в фундаментальной книге [170]. В статьях [17; 171] предложена знаменитая концепция решения, называемая равновесием по Нэшу.

Рассматривая каждую усеченную подыгру, обозначим через $u_k^{NE,l}(x) = (u_1^{NE,k,l}(x), \dots, u_n^{NE,k,l}(x))$, $l \in [0, \bar{T}]$, ситуацию в игре, образующую равновесие по Нэшу в позиционных стратегиях для усеченной подыгры $\Gamma_k(x_{k,0}, k, k + \bar{T})$, а через $\bar{V}_i^k(l, x)$ — выигрыш игрока $i \in I$ в ситуации равновесия по Нэшу в позиционных стратегиях в подыгре, заданной на временном интервале $[l, k + \bar{T}]$, $l = k, \dots, k + \bar{T}$:

$$\bar{V}_i^k(l, x) = \sum_{j=l}^{k+\bar{T}} g_i^{k+j} [x_{k,j}, u_1^{NE,k,j}, u_2^{NE,k,j}, \dots, u_n^{NE,k,j}, x_{k,j+1}] \left(\frac{1}{1+r} \right)^{k+j-1}.$$

Теорема 2.2.1 ([172]) *Предположим, что существуют функции $\bar{V}_i^k(l, x)$, удовлетворяющие (2.5). Тогда стратегии, максимизирующие правую часть*

(2.5), образуют ситуацию равновесия по Нэшу в позиционных стратегиях:

$$\begin{aligned}
\bar{V}_i^k(l, x) &= \max_{u_i^{k,l}} \left\{ g_i^k [x, u_1^{NE,k,l}(x), \dots, u_{i-1}^{NE,k,l}, u_i^{k,l}, u_{i+1}^{NE,k,l}, \dots, u_n^{NE,k,l}(x), \right. \\
&\quad \left. \bar{f}_k(x, u_i^{k,l}) \right] \left(\frac{1}{1+r} \right)^{k+l-1} + \bar{V}_i^k(l+1, \bar{f}_k(x, u_i^{k,l})) \left. \right\} = \\
&= g_i^k [x, u_1^{NE,k,l}(x), \dots, u_n^{NE,k,l}(x), f_k(x_k, u_1^{NE,k,l}(x), \dots, u_n^{NE,k,l}(x))] \times \\
&\quad \times \left(\frac{1}{1+r} \right)^{k+l-1} + \bar{V}_i^k(l+1, f_k(x_k, u_1^{NE,k,l}(x), \dots, u_n^{NE,k,l}(x))),
\end{aligned} \tag{2.5}$$

где $\bar{V}_i^k(\bar{T} + 1, x) = 0$, $i \in I$.

Обозначим через $u_k^{NE,l}(x)$, $l \in [0, \bar{T}]$, стратегии, полученные при максимизации правой части (2.5), а через $x_k^{NE} = (x_{k,0}^{NE}, \dots, x_{k,\bar{T}}^{NE})$ — траекторию, полученную с использованием стратегий $u_k^{NE,l}(x)$. Представленное решение относится только к усеченной подыгре, а не ко всей игре, определенной на временном интервале $[0, N]$. Равновесие по Нэшу не может быть построено с использованием классических подходов, поскольку информация об игре обновляется динамически, и на разных интервалах она разная. Чтобы построить результирующие равновесные стратегии и соответствующую траекторию для класса игр с динамическим обновлением информации, введем понятие результирующих стратегий и соответствующей результирующей траектории:

Определение 2.2.0.1 *Результирующие равновесные по Нэшу позиционные стратегии $\{\hat{u}_i^{NE,j}\}_{j=1}^N, \forall i \in I$ строятся следующим образом:*

$$\hat{u}_i^{NE,j} = \begin{cases} u_i^{NE,j,0}, & j \in \{1, \dots, N - \bar{T}\}, \\ u_i^{NE, N - \bar{T}, j - N - \bar{T}}, & j \in \{N - \bar{T} + 1, \dots, N\}. \end{cases}$$

Определение 2.2.0.2 *Результирующая некооперативная траектория $\{\hat{x}_j^{NE}\}_{j=1}^N$ представляет собой комбинацию из x_k^{NE} для каждой усеченной подыгры $\Gamma_k(x_{k,0}^{NE}, k, k + \bar{T})$:*

$$\{\hat{x}_j^{NE}\}_{j=0}^N = \begin{cases} x_{j,0}^{NE}, & j = \overline{1, N - \bar{T}}, \\ x_{N - \bar{T}, j - (N - \bar{T})}^{NE}, & j = \overline{N - \bar{T} + 1, N}. \end{cases} \tag{2.6}$$

Вдоль результирующей некооперативной траектории игроки получают выигрыш, вычисляемый по следующей формуле:

$$\begin{aligned} \hat{V}(\hat{x}_k^{NE}, N - k) &= \\ &= \begin{cases} \sum_{j=k}^{N-1-\bar{T}} [\bar{V}_i^j(0, x_{j,0}^{NE}) - \bar{V}_i^j(1, x_{j,1}^{NE})] + \bar{V}_i^{N-\bar{T}}(0, x_{N-\bar{T},0}^{NE}), & k = \overline{1, N - \bar{T}}, \\ \bar{V}_i^{N-\bar{T}}(k - N + \bar{T}, x_{N-\bar{T},k-N+\bar{T}}^{NE}), & k = \overline{N - \bar{T} + 1, N}. \end{cases} \quad (2.7) \end{aligned}$$

2.3 Кооперативная дискретная игра с динамическим обновлением информации

Теория кооперативных динамических игр предлагает подходящие (в социальном отношении) и эффективные (в групповом смысле) решения для различных задач принятия решений со стратегическими действиями. Одна из фундаментальных проблем теории кооперативных динамических игр — формулировка оптимального или кооперативного поведения игроков и определение правила распределения (или кооперативного решения), представляющего собой множество дележей, выбранных игроками в игре. Предположим, что игроки решают сотрудничать в каждой усеченной подыгре. Обозначим через $\Gamma_k^c(x_{k,0}, k, k + \bar{T})$ усеченную кооперативную подыгру, определенную на временном интервале $[k, k + \bar{T}]$ с начальным условием $x_{k,0}$. Максимальный кооперативный выигрыш всех игроков в этой игре можно определить, максимизируя функционал вида

$$\sum_{i \in I} K_i(x_{k,0}, k, k + \bar{T}, U) = \sum_{i \in I} \sum_{j=0}^{\bar{T}} g_i^{k+j} [x_{k,j}, u_1^{k,j}, u_2^{k,j}, \dots, u_n^{k,j}, x_{k,j+1}] \left(\frac{1}{1+r} \right)^{k+j-1}, \quad (2.8)$$

при условии

$$x_{k+1} = f_k(x_k, u_1^k, u_2^k, \dots, u_n^k), \quad (2.9)$$

$$x_{1,0} = x_1 = x^0. \quad (2.10)$$

Пусть функция Беллмана для каждой усеченной подыгры $\Gamma_k^c(x_{k,0}, k, k + \bar{T})$

имеет следующий вид [173]:

$$W^k(0, x_{k,0}) = \max_{u_i(k), i \in I} \left\{ \sum_{i \in I} \sum_{j=0}^{\bar{T}} g_i^{k+j} [x_{k,j}, u_1^{k,j}, \dots, u_n^{k,j}, f_k(x_{k,j}, u_1^{k,j}, \dots, u_n^{k,j})] \left(\frac{1}{1+r} \right)^{k+j-1} \right\}, \quad (2.11)$$

где $u_i(k) = \{u_i^{k,0}, \dots, u_i^{k,\bar{T}}\}$.

Теорема 2.3.1 *Предположим, что существует функция $W^k(l, x)$ в усеченной подыгре $\Gamma_k^c(x_{k,0}, k, k + \bar{T})$, $l \in [0, \bar{T}]$, такая, что выполняются следующие рекурсивные соотношения:*

$$W^k(l, x) = \max_{u_i(k), i \in I} \left\{ \sum_{i \in I} g_i^k [x, u_1^{k,l}, \dots, u_n^{k,l}, f_k(x, u_1^{k,l}, \dots, u_n^{k,l})] \left(\frac{1}{1+r} \right)^{k+l-1} + W^k(l+1, f_k(x, u_1^k(x), \dots, u_n^k(x))) \right\}, \quad (2.12)$$

Предположим, что максимум в (2.12) достигается при управлении $u_i^(k)$. Тогда $u_i^*(k)$ оптимально в задаче (2.8), (2.9).*

Определение 2.3.0.1 *Результирующие кооперативные стратегии $\{\hat{u}_i^{*,j}\}_{j=1}^N$, $\forall i \in I$ строятся следующим образом:*

$$\hat{u}_i^{*,j} = \begin{cases} u_i^{*,j,0}, & j \in \{1, \dots, N - \bar{T}\}, \\ u_i^{*,N-\bar{T},j-N-\bar{T}}, & j \in \{N - \bar{T} + 1, \dots, N\}. \end{cases}$$

Аналогичным образом определим результирующую кооперативную траекторию. Обозначим через u^* оптимальные стратегии, полученные максимизацией правой части (2.12). Соответствующая траектория задается в виде $x_k^* = (x_{k,0}^*, \dots, x_{k,\bar{T}}^*)$.

Определение 2.3.0.2 *Результирующая кооперативная траектория $\{\tilde{x}_l^*\}_{j=1}^N$ представляет собой комбинацию кооперативных траекторий x_k^* для каждой*

усеченной кооперативной подыгры $\Gamma_k^c(x_k^*, k, k + \bar{T})$, определенной на последовательных временных интервалах $[k, k + \bar{T}]$:

$$\{\tilde{x}_j^*\}_{j=1}^N = \begin{cases} x_{j,0}^*, & j = \overline{1, N - \bar{T}}, \\ x_{N-\bar{T}, j-(N-\bar{T})}^*, & j = \overline{N - \bar{T} + 1, N}. \end{cases} \quad (2.13)$$

На временном интервале $[k, k + 1]$ результирующая кооперативная траектория совпадает с кооперативной траекторией x_k^* в усеченной кооперативной подыгре $\Gamma_k^c(x_k^*, k, k + \bar{T})$. На шаге $k + 1$ обновляется информация о структуре игры. На временном интервале $[k + 1, k + 2]$ траектория \tilde{x}_k^* совпадает с кооперативной траекторией x_{k+1}^* в усеченной кооперативной подыгре $\Gamma_{k+1}^c(x_{k+1}^*, k + 1, k + 1 + \bar{T})$ и т. д.

Характеристическая функция коалиции — важное понятие в теории кооперативных игр. По аналогии с [174], ниже характеристическая функция определяется как кооперативный выигрыш игроков из коалиции $S \subseteq I$ в равновесии по Нэшу в игре со следующим набором игроков: коалиция S (действующая как один игрок) и игроки из множества $I \setminus S$ (действующие как индивидуумы). Для каждой коалиции $S \subseteq I$, определим значения характеристической функции в каждой усеченной подыгре, как это было сделано в [112]:

$$V^k(S; x_{k,0}^*) = \begin{cases} \sum_{i \in I} K_i^k(x_{k,0}^*; u), & S = I, \\ \tilde{V}_k(S, x_{k,0}^*), & S \subseteq I, \\ 0, & S = \emptyset, \end{cases} \quad (2.14)$$

где $\tilde{V}_k(S, x_{k,0}^*)$ — кооперативный выигрыш игроков из коалиции $S \setminus I$ в равновесии по Нэшу в позиционных стратегиях $u^{NE} = (\bar{u}_1^{NE}, \dots, \bar{u}_n^{NE})$ в игре со следующим набором игроков: коалиция S (действующая как один игрок) и игроки из множества $I \setminus S$, т.е. в игре с $|I \setminus S| + 1$ игроками. При $|S| = 1$ кооперативный случай превращается в некооперативный. В других случаях игроки из коалиции S действуют как один игрок, максимизируя свой кооперативный выигрыш.

2.3.1 Результирующее кооперативное решение и соответствующие теоремы

Дележ $\xi^k(x_k^*, k, k + \bar{T})$ для каждой усеченной кооперативной подыгры $\Gamma_{k+1}^c(x_k^*, k, k + \bar{T})$ определяется как произвольный вектор, удовлетворяющий условиям:

индивидуальная рациональность: $\xi_i^k(x_k^*, k, k + \bar{T}) \geq V(\{i\}; x_k^*, k, k + \bar{T})$,

групповая рациональность: $\sum_{i \in I} \xi_i^k(x_k^*, k, k + \bar{T}) = V(I; x_k^*, k, k + \bar{T})$.

Обозначим через $E_k(x^*, k, k + \bar{T})$ множество всех возможных дележей для каждой усеченной подыгры. Предположим, что для каждой усеченной подыгры непустое кооперативное решение определяется следующим образом:

$$M_k(x^*, k, k + \bar{T}) \subseteq E_k(x^*, k, k + \bar{T}). \quad (2.15)$$

Это может быть С-ядро, N-ядро, или вектор Шепли.

Чтобы гарантировать динамическую устойчивость [169] в каждой кооперативной усеченной подыгре, необходимо использовать понятие процедуры распределения дележа. Теорема о динамической устойчивости была представлена в 1977 г. Л. Петросяном [175].

Следуя непрерывному временному анализу Янга и Петросяна [176] для кооперативных дифференциальных игр, сформулируем дискретную версию процедуры распределения дележа [177], чтобы согласованные дележи были динамически устойчивыми. Обозначим через $B_i^k(x_i^{*,k})$ платеж, который игрок i получает на шаге k в соответствии с кооперативным соглашением вдоль кооперативной траектории $\{x_i^{*,k}\}_k^N$. Схема выплат, включающая $B_i^k(x_i^{*,k})$, представляет собой ПРД в том смысле, что дележ игрока i на шагах от k до N может быть записан в виде

$$\xi_i^k(x_i^{*,k}) = B_i^k(x_i^{*,k}) + \sum_{j=k+1}^N B_i^j(x_i^{*,j}), \quad i \in I. \quad (2.16)$$

Определение 2.3.0.3 Решение $M(x_k^*, k)$ называется динамически устойчивым, если для любого дележа $\xi^k(x_k^*, k) \in M(x_k^*, k)$ существует ПРД $B^k(x_i^{*,k}) =$

$(B_1^k(x_i^{*,k}), \dots, B_n^k(x_i^{*,k}))$, такая что при $k = \overline{1, N}$ выполнено условие

$$\sum_k^N B^k(x_i^{*,k}) \in M(x_k^*, k) \quad \left(\sum_k^N B^k(x_i^{*,k}) = \xi^k(x_i^{*,k}) \right).$$

Выбранный принцип оптимальности или кооперативное решение является динамически устойчивым, если условие (2.3.0.3) справедливо в любой момент времени на протяжении всей игры вдоль кооперативной траектории.

Чтобы построить результирующее кооперативное решение или кооперативное решение в игре с динамическим обновлением информации $\Gamma(x^0, N)$, необходимо использовать особый подход, поскольку стандартный подход применим только к каждой усеченной подыгре с динамическим обновлением информации, определенной на временном интервале $[k, k + \bar{T}]$. Кроме того, ПРД позволяет доказать теорему о динамической устойчивости кооперативного решения с динамическим обновлением информации, а также определять решения на временном интервале $[k, k + \bar{T}]$.

Предположим, что в каждой усеченной подыгре $\Gamma_k^c(x_k^*, k, k + \bar{T})$ выбирается кооперативное решение $M_k(x^*, k, k + \bar{T}) \neq \emptyset$ вдоль кооперативной траектории x_k^* . Пусть дележ

$$\xi^k(x_k^*, k, k + \bar{T}) \in M_k(x^*, k, k + \bar{T})$$

и соответствующая ПРД при $k \in \overline{1, N - \bar{T}}$ выбираются таким образом, чтобы выполнить условия динамической устойчивости:

$$\beta^{k,l}(x_{k,l}^*) = [\beta_1^{k,l}(x_{k,l}^*), \dots, \beta_n^{k,l}(x_{k,l}^*)], \text{ при } l \in \overline{0, \bar{T}}.$$

Соответствующая ПРД $\beta^{k,l}(x_{k,l}^*)$ может быть получена в следующем виде [169]:

$$\beta^{k,l}(x_{k,l}^*) = \xi^k(x_{k,l}^*, k + l, k + \bar{T}) - \xi^k(x_{k,l+1}^*, k + l + 1, k + \bar{T}), \quad l \in \overline{0, \bar{T}}, \quad (2.17)$$

где $\xi^k(x_{k,\bar{T}+1}^*, k + \bar{T} + 1, k + \bar{T}) = 0$.

Однако динамическая устойчивость сохраняется только на временном интервале $[k, k + \bar{T}]$ для каждой усеченной подыгры $\Gamma_k^c(x_k^*, k, k + \bar{T})$, $k \in \{1, \dots, N\}$.

Чтобы гарантировать динамическую устойчивость для всего временного интервала $[1, N]$, введем результирующую ПРД.

Определение 2.3.0.4 Результирующая ПРД $\hat{\beta}^k(\tilde{x}^*)$ задается для множества выбранных дележей в каждой усеченной подыгре $\xi^k(x_{k,0}^*, k, k + \bar{T}) \in M_k(x^*, k, k + \bar{T})$ с помощью соответствующего $\beta^k(x_k^*)$ следующим образом:

$$\left\{ \hat{\beta}^k(x^*) \right\}_{k=1}^N = \begin{cases} \beta^{k,0}(x_{k,0}^*), & k = \overline{1, N - \bar{T}}, \\ \beta^{N-\bar{T}, k-N+\bar{T}}(x_{N-\bar{T}, k-N+\bar{T}}^*), & k = \overline{N - \bar{T} + 1, N}. \end{cases} \quad (2.18)$$

Используя результирующую ПРД $\hat{\beta}^k(\tilde{x}^*)$, введем еще одно понятие.

Определение 2.3.0.5 Результирующим дележом $\hat{\xi}(\tilde{x}^*, N - k)$ называется вектор, заданный с помощью результирующей ПРД $\hat{\beta}^k(x^*)$ следующим образом для шага $k = \overline{1, N}$:

$$\hat{\xi}(\tilde{x}^*, N - k) = \sum_{l=k}^N \hat{\beta}^l(\tilde{x}^*); \quad (2.19)$$

в частности, $\hat{\xi}(\tilde{x}^*, N) = \sum_{l=1}^N \hat{\beta}^l(\tilde{x}^*)$.

На основе понятия результирующего дележа можно определить правило распределения для кооперативного выигрыша:

Определение 2.3.0.6 Результирующим решением $\hat{M}_k(\tilde{x}^*, N - k)$ называется множество результирующих дележей $\hat{\xi}(\tilde{x}^*, N - k)$ для всевозможных результирующих ПРД, $\hat{\beta}^k(\tilde{x}^*)$, заданных различными дележами в каждой усеченной подыгре.

Любой результирующий дележ $\hat{\xi}(\tilde{x}^*, N) \in \hat{M}(\tilde{x}^*, N)$ и результирующая ПРД $\hat{\beta}^k(\tilde{x}^*)$ распределяют кооперативный выигрыш игроков вдоль результирующей кооперативной траектории \tilde{x}_l^* в исходной игре с заданной продолжительностью, или выполняется следующее условие:

$$\sum_{i \in I} \sum_{j=1}^k \hat{\beta}_i^j(\tilde{x}_j^*) = \sum_{i \in I} \sum_{j=1}^k g_i^j[\tilde{x}_j^*, \tilde{u}^*, \tilde{x}_{l+1}^*] \left(\frac{1}{1+r} \right)^{j-1}. \quad (2.20)$$

Результирующее решение $\hat{M}_k(\tilde{x}^*, N - k)$ динамически устойчиво по построению. Однако можно также доказать, что произвольное результирующее кооперативное решение сильно динамически устойчиво в игре с динамическим обновлением информации. Под произвольным решением мы понимаем решение, состоящее, например, из некоторых векторов из S -ядра на первом шаге, вектора Шепли на втором шаге и т.д.

Определение 2.3.0.7 Кооперативное решение $M(x^0, N)$ называется сильно динамически устойчивым, если для $\forall k = \overline{1, N}$ и каждого $\xi(x^0, N) \in M(x^0, N)$ соответствующая ПРД $\beta^k(x_k^*)$ удовлетворяет условию

$$\sum_{j=1}^k \beta^j(x_j^*) \oplus M(x_{k,0}^*, N - k) \subseteq M(x^0, N), \quad (2.21)$$

где $a \oplus A = \{a + a' : a' \in A\}$.

Теорема 2.3.2 Произвольное результирующее кооперативное решение $\hat{M}(\tilde{x}^*, N)$ сильно динамически устойчиво в игре с динамическим обновлением информации.

2.3.2 Свойства результирующего кооперативного решения

2.3.2.1 Построение характеристической функции в игре с динамическим обновлением информации

В предыдущем разделе было введено понятие результирующего кооперативного решения или, другими словами, определен способ распределения кооперативного выигрыша вдоль выбранной траектории в игре с динамическим обновлением информации. Однако не было показано, что построенное множество — это множество дележей или множество векторов, удовлетворяющих условиям индивидуальной и групповой рациональности. Чтобы обосновать результирующее кооперативное решение, нужно определить характеристическую функцию в игре с динамическим обновлением информации.

Определение 2.3.0.8 Результирующая характеристическая функция $\bar{V}(S; \tilde{x}_k^*, N - k)$, $S \subseteq I$ в игре с динамическим обновлением информации $\Gamma(\tilde{x}_k^*, N - k)$ — функция, вычисленная с помощью значений характеристических функций $V(S; x_k^*, k, k + \bar{T})$ в каждой усеченной подыгре $\Gamma_k^c(x_{k,0}, k, k + \bar{T})$ вдоль результирующей кооперативной траектории \tilde{x}_k^* при $k \in \{1, \dots, N\}$:

$$\bar{V}(S; \tilde{x}_k^*, N - k) = \begin{cases} \sum_{l=k}^{N-\bar{T}-1} [V(S; x_l^*, l, l + \bar{T}) - V(S; x_{l+1}^*, l + 1, l + \bar{T})] + \\ \quad + V(S; x_{N-\bar{T}}^*, N - \bar{T}, N), \quad k = \overline{1, N - \bar{T}}, \\ V(S; x_k^*, k, N), \quad k = \overline{N - \bar{T} + 1, N}. \end{cases} \quad (2.22)$$

Используя полученную характеристическую функцию, можно показать, что результирующий дележ $\hat{\xi}(\tilde{x}^*, N)$ является дележом в игре с динамическим обновлением информации.

Теорема 2.3.3 Результирующий дележ $\hat{\xi}(\tilde{x}^*, N)$ является дележом в игре с динамическим обновлением информации $\Gamma(x^0, N)$, если при $k \in \{1, \dots, N - \bar{T}\}$ выполняется следующее условие:

$$\begin{aligned} \xi_i^k(x_{k,l}^*, k + l, k + \bar{T}) - \xi_i^k(x_{k,l+1}^*, k + l + 1, k + \bar{T}) &\geq \\ &\geq V(\{i\}; x_{k,l}^*, k + l, k + \bar{T}) - V(\{i\}; x_{k,l+1}^*, k + l + 1, k + \bar{T}), \quad i \in I, \end{aligned} \quad (2.23)$$

где $l = \overline{0, \bar{T}}$.

Доказательство Необходимо установить, что справедливы следующие условия:

$$\sum_{i=1}^n \hat{\xi}_i(\tilde{x}_k^*, N - k) = \bar{V}(I; \tilde{x}_k^*, N - k), \quad (2.24)$$

$$\hat{\xi}_i(\tilde{x}_k^*, N - k) \geq \bar{V}(\{i\}; \tilde{x}_k^*, N - k), \quad i \in I. \quad (2.25)$$

Введем обозначение

$$y(k) = \begin{cases} N - \bar{T} + 1, & k = \overline{1, N - \bar{T}}, \\ k, & k = \overline{N - \bar{T} + 1, N}. \end{cases}$$

Согласно Определению 2.3.0.5, правая часть (2.24) может быть записана в виде

$$\begin{aligned}
\sum_{i=1}^n \hat{\xi}_i(\tilde{x}_k^*, N - k) &= \sum_{i=1}^n \sum_{j=k}^N \hat{\beta}_i^j(\tilde{x}_k^*) = \\
&= \sum_{i=1}^n \left[\sum_{j=k}^{N-\bar{T}} (\bar{\xi}_i^j(x_{j,0}^*, j, j + \bar{T}) - \bar{\xi}_i^k(x_{j,1}^*, j + 1, j + \bar{T})) + \right. \\
&+ \left. \sum_{j=y(k)}^N (\bar{\xi}_i^{N-\bar{T}}(x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - \bar{\xi}_i^{N-\bar{T}}(x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N)) \right] = \quad (2.26) \\
&= \sum_{j=k}^{N-\bar{T}} \left[V(I; x_{j,0}^*, j, j + \bar{T}) - V(I; x_{j,1}^*, j + 1, j + \bar{T}) \right] + \\
&+ \sum_{j=y(k)}^N \left[V(I; x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - V(I; x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N) \right].
\end{aligned}$$

Правая часть (2.26) равна $\bar{V}(I; \tilde{x}_k^*, N - k)$ по Определению 2.3.0.8; следовательно, (2.24) имеет место. Далее, рассмотрим выражение (2.25), правая часть которого может быть записана в виде

$$\begin{aligned}
\hat{\xi}_i(\tilde{x}_k^*, N - k) &= \sum_{j=k}^{N-\bar{T}} (\bar{\xi}_i^j(x_{j,0}^*, j, j + \bar{T}) - \bar{\xi}_i^k(x_{j,1}^*, j + 1, j + \bar{T})) + \\
&+ \sum_{j=y(k)}^N (\bar{\xi}_i^{N-\bar{T}}(x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - \bar{\xi}_i^{N-\bar{T}}(x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N)). \quad (2.27)
\end{aligned}$$

Условие (2.25) выполняется на каждом шаге, если при $k = \overline{1, N}$ справедливо неравенство

$$\begin{aligned}
&\bar{\xi}_i^j(x_{j,0}^*, j, j + \bar{T}) - \bar{\xi}_i^k(x_{j,1}^*, j + 1, j + \bar{T}) \geq \\
&V(\{i\}; x_{j,0}^*, j, j + \bar{T}) - V(\{i\}; x_{j,1}^*, j + 1, j + \bar{T}), j = \overline{k, N - \bar{T}}, \\
&\bar{\xi}_i^{N-\bar{T}}(x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - \bar{\xi}_i^{N-\bar{T}}(x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N) \geq \\
&V(\{i\}; x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - V(\{i\}; x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N), j = \overline{N - \bar{T} + 1, N}. \quad (2.28)
\end{aligned}$$

Из условия (2.28) следует условие

$$\begin{aligned} \xi_i^k(x_{k,l}^*, k+l, k+\bar{T}) - \xi_i^k(x_{k,l+1}^*, k+l+1, k+\bar{T}) \geq \\ V(\{i\}; x_{k,l}^*, k+l, k+\bar{T}) - V(\{i\}; x_{k,l+1}^*, k+l+1, k+\bar{T}), \quad l = \overline{0, \bar{T}}, \quad i \in I, \end{aligned} \quad (2.29)$$

для каждой усеченной подыгры $k = \{1, \dots, N - \bar{T}\}$.

Теорема доказана. □

2.3.2.2 Связь между решениями усеченных подыгр и результирующими решениями

В этом разделе получен следующий результат: если в каждой усеченной подыгре игроки выбирают дележи $\xi^k(x_k^*, k, k + \bar{T}) \in E_k(x_k^*, k, k + \bar{T})$ на основе $V(S; x_k^*, k, k + \bar{T})$ при $k \in \{1, \dots, N - \bar{T}\}$, $S \subseteq I$, по одному и тому же правилу, то результирующий дележ $\hat{\xi}(\tilde{x}^*, N - k)$ соответствует дележу, выбранному по тому же правилу с использованием результирующей характеристической функции $\bar{V}(S; x_k^*, N - k)$, $S \subseteq I$. Далее, этот факт доказывается для ряда классических принципов оптимальности.

Предположим, что в каждой усеченной подыгре $\Gamma_k^c(x_k^*, k, k + \bar{T})$ игроки выбирают в качестве дележа вектор Шепли $Sh^k(x_k^*, k, k + \bar{T})$. Тогда соответствующий результирующий дележ $\hat{\xi}(\tilde{x}^*, N - k)$ совпадает с вектором Шепли $\hat{Sh}(\tilde{x}_k^*, N - k)$, вычисленным с помощью результирующей характеристической функции $\bar{V}(S; \tilde{x}_k^*, N - k)$, $S \subseteq I$.

Теорема 2.3.4 *Предположим, что в каждой усеченной подыгре $\Gamma_k^c(x_k^*, k, k + \bar{T})$:*

$$\xi^k(x_{k,0}^*, k, k + \bar{T}) = Sh^k(x_k^*, k, k + \bar{T}),$$

где $k \in \{1, \dots, N - \bar{T}\}$. Тогда соответствующий результирующий дележ $\hat{\xi}(\tilde{x}_k^*, N - k)$ совпадает с $\hat{Sh}(\tilde{x}_k^*, N - k)$:

$$\hat{\xi}(\tilde{x}_k^*, N - k) = \hat{Sh}(\tilde{x}_k^*, N - k), \quad k \in [1, N],$$

где $\hat{Sh}(\tilde{x}_k^*, N - k)$ — вектор Шепли, вычисленный с помощью результирующей характеристической функции $\bar{V}(S; \tilde{x}_k^*, N - k)$, $S \subseteq I$.

Доказательство Согласно Определению 2.3.0.5 и (2.17), результирующий дележ вычисляется с помощью векторов Шепли $Sh^k(x_k^*, k, k + \bar{T})$ в каждой усеченной подыгре:

$$\begin{aligned}
\hat{\xi}(x_k^*, N - k) &= \sum_{j=k}^{N-\bar{T}} [Sh^j(x_{j,0}^*, j, j + \bar{T}) - Sh^j(x_{j,1}^*, j + 1, j + \bar{T})] + \\
&+ \sum_{j=y(k)}^N [Sh^{N-\bar{T}}(x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - Sh^{N-\bar{T}}(x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N)] \\
&= \sum_{S \subset I, i \in S} \frac{(|I| - |S|)! (|S| - 1)!}{|I|!} \left\{ \sum_{j=k}^{N-\bar{T}} [(V(S; x_{j,0}^*, j, j + \bar{T}) - \right. \\
&- V(S \setminus i; x_{j,0}^*, j, j + \bar{T})) - (V(S; x_{j,1}^*, j + 1, j + \bar{T}) - V(S \setminus i; x_{j,1}^*, j + 1, j + \bar{T}))] + \\
&+ \sum_{j=y(k)}^N [(V(S; x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - V(S \setminus i; x_{N-\bar{T},j-N+\bar{T}}^*, j, N)) - \\
&- (V(S; x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N) - V(S \setminus i; x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N))] \left. \right\} = \\
&= \sum_{S \subset I, i \in S} \frac{(|I| - |S|)! (|S| - 1)!}{|I|!} \left\{ \sum_{j=k}^{N-\bar{T}} [(V(S; x_{j,0}^*, j, j + \bar{T}) - \right. \\
&- V(S; x_{j,1}^*, j + 1, j + \bar{T})) - (V(S \setminus i; x_{j,0}^*, j, j + \bar{T}) - V(S \setminus i; x_{j,1}^*, j + 1, j + \bar{T}))] + \\
&+ \sum_{j=y(k)}^N [(V(S; x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - V(S; x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N)) - \\
&- (V(S \setminus i; x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - V(S \setminus i; x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N))] \left. \right\}, \tag{2.30}
\end{aligned}$$

где $\hat{Sh}(x_k^*, N - k)$ — вектор Шепли, вычисленный с помощью результирующей

характеристической функции $\bar{V}(S, \tilde{x}_k^*, N - k)$:

$$\begin{aligned}
& \hat{S}h(x_k^*, N - k) = \\
& = \sum_{S \subset I, i \in S} \frac{(|I| - |S|)! (|S| - 1)!}{|I|!} (\bar{V}(S; \tilde{x}_k^*, N - k) - \bar{V}(S \setminus i; \tilde{x}_k^*, N - k)) = \\
& = \sum_{S \subset I, i \in S} \frac{(|I| - |S|)! (|S| - 1)!}{|I|!} \left\{ \sum_{j=k}^{N-\bar{T}} [(V(S; x_{j,0}^*, j, j + \bar{T}) - \right. \\
& - V(S; x_{j,1}^*, j + 1, j + \bar{T})) - (V(S \setminus i; x_{j,0}^*, j, j + \bar{T}) - V(S \setminus i; x_{j,1}^*, j + 1, j + \bar{T}))] + \\
& + \sum_{j=y(k)}^N [(V(S; x_{N-\bar{T}, j-N+\bar{T}}^*, j, N) - V(S; x_{N-\bar{T}, j+1-N+\bar{T}}^*, j + 1, N)) - \\
& \left. - (V(S \setminus i; x_{N-\bar{T}, j-N+\bar{T}}^*, j, N) - V(S \setminus i; x_{N-\bar{T}, j+1-N+\bar{T}}^*, j + 1, N))] \right\}.
\end{aligned} \tag{2.31}$$

Сравнивая (2.30) и (2.31), получаем необходимый результат.

Теорема доказана. \square

Предположим, что в каждой усеченной подыгре $\Gamma_k^c(x_k^*, k, k + \bar{T})$ игроки выбирают в качестве кооперативного решения S -ядро $C(x_k^*, k, k + \bar{T})$. Тогда каждый результирующий дележ $\hat{\xi}(x^*, N - k)$ принадлежит S -ядру, определенному с помощью результирующей характеристической функции $\bar{V}(S, \tilde{x}_k^*, N - k)$.

Теорема 2.3.5 *Предположим, что в каждой усеченной подыгре $\Gamma_k^c(x_k^*, k, k + \bar{T})$:*

$$M_k(x^*, k, k + \bar{T}) = C(x_k^*, k, k + \bar{T}),$$

где $k \in \{1, \dots, N - \bar{T}\}$. Тогда для каждого $\xi^k(x_k^*, k, k + \bar{T}) \in C(x_k^*, k, k + \bar{T})$ выполнено следующее условие:

$$\hat{\xi}(\tilde{x}_k^*, N - k) \in \hat{C}(\tilde{x}_k^*, N - k), \quad k \in [k, N],$$

где $\hat{C}(\tilde{x}_k^*, N - k)$ — S -ядро, вычисленное с помощью результирующей характе-

ристической функции $\bar{V}(S, \tilde{x}_k^*, N - k)$, $S \subseteq I$, при условии

$$\begin{aligned} & \sum_{i \in S} \bar{\xi}_i^k(x_{k,0}^*, k, k + \bar{T}) - V(S; x^*, k, k + \bar{T}) \geq \\ & \geq \sum_{i \in S} \bar{\xi}_i^k(x_{k,1}^*, k + 1, k + \bar{T}) - V(S; x^*, k + 1, k + \bar{T}) \geq 0, \quad S \subseteq I. \end{aligned} \quad (2.32)$$

Доказательство Утверждение теоремы имеет место при справедливости следующих утверждений:

- Если в каждой усеченной подыгре игроки выбирают дележ $\xi^k(x_k^*, k, k + \bar{T}) \in C(x_k^*, k, k + \bar{T})$, вычисленный с помощью $V(S; x^*, k, k + \bar{T})$ при $k \in \{1, \dots, N - \bar{T}\}$, то результирующий дележ $\hat{\xi}(\tilde{x}_k^*, N - k)$ принадлежит S -ядру $\hat{C}(\tilde{x}_k^*, N - k)$, вычисленному с помощью результирующей характеристической функции $\bar{V}(S, \tilde{x}_k^*, N - k)$.
- S -ядро $\hat{C}(\tilde{x}_k^*, N - k)$ не должно содержать дележ $\hat{\xi}(\tilde{x}_k^*, N - k)$, для которого невозможно найти множество дележей в усеченной подыгре $\xi^k(x_k^*, k, k + \bar{T}) \in C(x_k^*, k, k + \bar{T})$.

Рассмотрим первое утверждение. Если множество дележей $\xi^k(x_k^*, k, k + \bar{T})$ удовлетворяет системе неравенств

$$\sum_{i \in S} \xi_i^k(x_{k,0}^*, k, k + \bar{T}) \geq V(S; x^*, k, k + \bar{T}), \quad S \subseteq I, \quad (2.33)$$

при $k \in \{1, \dots, N - \bar{T}\}$, $i \in I$, то результирующий дележ $\hat{\xi}(\tilde{x}_k^*, N - k)$ удовлетворяет системе неравенств

$$\sum_{i \in S} \hat{\xi}_i(\tilde{x}_k^*, N - k) \geq \bar{V}(S, \tilde{x}_k^*, N - k). \quad (2.34)$$

Подставим выражение для $\xi_i^k(x_{k,0}^*, k, k + \bar{T})$ в левую часть (2.34), а (2.22) — в правую часть (2.34). Поэтому для каждого $S \subseteq I$, $k \in \{1, \dots, N - \bar{T}\}$, из справедливости (2.32) следует справедливость (2.33); таким образом, (2.34) выполнено. Первое утверждение доказано.

Рассмотрим второе утверждение. Для любого $\hat{\xi}(\tilde{x}_k^*, N - k) \in \hat{C}(\tilde{x}_k^*, N - k)$, $k \in \{1, \dots, N - \bar{T}\}$, справедливы следующие неравенства на множестве дележей

$\xi_i^k(x_{k,0}^*, k, k + \bar{T})$, $i \in \{1, \dots, n\}$, $k \in \{1, \dots, N - \bar{T}\}$:

$$\begin{aligned}
& \sum_{i \in S} \left[\sum_{j=k}^{N-\bar{T}} (\xi_i^j(x_{j,0}^*, j, j + \bar{T}) - \xi_i^k(x_{j,1}^*, j + 1, j + \bar{T})) + \right. \\
& \left. + \sum_{j=y(k)}^N (\xi_i^{N-\bar{T}}(x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - \xi_i^{N-\bar{T}}(x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N)) \right] \geq \\
& \geq \sum_{j=k}^{N-\bar{T}} (V(S; x_{j,0}^*, j, j + \bar{T}) - V(S; x_{j,1}^*, j + 1, j + \bar{T})) + \\
& + \sum_{j=y(k)}^N (V(S; x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - V(S; x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N)).
\end{aligned} \tag{2.35}$$

Допустим, что существует одно множество дележей $\xi_i^k(x_{k,0}^*, k, k + \bar{T})$, где $i \in \{1, \dots, n\}$, $k \in \{1, \dots, N - \bar{T}\}$ в (2.35), для которого неравенства (2.33) нарушены. Тогда при $k \in \{1, \dots, N - \bar{T}\}$ получаем:

$$\sum_{i \in S} \xi_i^k(x_{k,0}^*, k, k + \bar{T}) - V(S; x^*, k, k + \bar{T}) < 0, \tag{2.36}$$

что противоречит (2.32), и второе утверждение также справедливо. Теорема доказана. □

2.4 Случайный информационный горизонт

В этом разделе рассмотрим случай, когда информационный горизонт не фиксирован. Это означает, что игроки могут иметь определенную информацию на усеченных временных интервалах $[k, \bar{T}_k]$, но длина интервалов является случайной величиной. Обозначим через \bar{T}_k информационный горизонт на шаге k . Параметр \bar{T}_k может принимать значения из интервала $[\hat{T}_{k-1} + 1, N]$, где \hat{T}_{k-1} — реализация случайного горизонта \bar{T}_{k-1} в предыдущей усеченной подыгре, в частности $\hat{T}_0 = 0$. Стоит отметить, что процедура обновления информации останавливается как только $\bar{T}_k = N$. Параметр \bar{T}_k — дискретная случайная величина. В начальной позиции x_0 определены величины $\{\gamma_1, \dots, \gamma_N\}$ — ве-

роятности реализации на интервале $[1, N]$. В позиции x_1 (вторая позиция) вероятности пересчитываются следующим образом:

$$\frac{\gamma_{\bar{T}_{k-1}+1}}{\sum_{\zeta=\bar{T}_{k-1}+1}^N \gamma_{\zeta}}, \dots, \frac{\gamma_N}{\sum_{\zeta=\bar{T}_{k-1}+1}^N \gamma_{\zeta}}.$$

Уравнения движения на временном интервале $[k, \bar{T}_k]$ имеют вид

$$\begin{aligned} x_{k+1} &= f_k(x_k, u_1^k, u_2^k, \dots, u_n^k), \\ x_{1,0} &= x_1 = x^0. \end{aligned} \quad (2.37)$$

Функция выигрыша игрока $i \in \{1, \dots, n\}$ в k -ой усеченной подыгре определяется как

$$E \left\{ \sum_{\zeta=k}^{\bar{T}_k} g_i^{\zeta}(x_{\zeta}, u_{\zeta}, x_{\zeta+1}) \right\} = \sum_{\bar{T}_k=\bar{T}_{k-1}+1}^N \frac{\gamma_{\bar{T}_k}}{\sum_{j=\bar{T}_{k-1}+1}^N \gamma_j} \sum_{\zeta=k}^{\bar{T}_k} g_i^{\zeta}(x_{\zeta}, u_{\zeta}, x_{\zeta+1}). \quad (2.38)$$

Рассмотрим некооперативную игру. Чтобы определить равновесие по Нэшу в позиционных стратегиях, зададим функцию Беллмана $V_i^k(x, k)$ для k -ой усеченной подыгры в этой модели следующим образом:

$$V_i^k(x, k) = \max_{u_i^k, \dots, u_i^{\bar{T}_k}} \left\{ \sum_{\bar{T}_k=\bar{T}_{k-1}+1}^N \frac{\gamma_{\bar{T}_k}}{\sum_{j=\bar{T}_{k-1}+1}^N \gamma_j} \sum_{\zeta=k}^{\bar{T}_k} g_i^{\zeta}(x_{\zeta}, u_{\zeta}, x_{\zeta+1}) \mid x_k = x \right\}$$

со множеством стратегий $\{u_i^j = \psi_i^j, j = \overline{k, \bar{T}_k}, i \in I\}$. Тогда:

$$\begin{aligned} V_i^k(x, k) &= \max_{u_i^k, \dots, u_i^{\bar{T}_k}} \left\{ \sum_{\bar{T}_k=\bar{T}_{k-1}+1}^N \frac{\gamma_{\bar{T}_k}}{\sum_{j=\bar{T}_{k-1}+1}^N \gamma_j} \sum_{\zeta=k}^{\bar{T}_k} g_i^{\zeta}(x_{\zeta}, u_{\zeta}, x_{\zeta+1}) \right\} = \\ &= \sum_{\bar{T}_k=\bar{T}_{k-1}+1}^N \frac{\gamma_{\bar{T}_k}}{\sum_{j=\bar{T}_{k-1}+1}^N \gamma_j} \sum_{\zeta=k}^{\bar{T}_k} g_i^{\zeta}(x_{\zeta}^*, \psi_{\zeta}, x_{\zeta+1}^*) = \\ &= \sum_{l=k}^{\bar{T}_{k-1}} g_i^l(x_l^*, \psi_l, x_{l+1}^*) + \sum_{\bar{T}_k=\bar{T}_{k-1}+1}^N \frac{\gamma_{\bar{T}_k}}{\sum_{j=\bar{T}_{k-1}+1}^N \gamma_j} \sum_{\zeta=\bar{T}_{k-1}+1}^{\bar{T}_k} g_i^{\zeta}(x_{\zeta}^*, \psi_{\zeta}, x_{\zeta+1}^*), \end{aligned}$$

где $x_{k+1}^* = f_k(x_k^*, \psi_k(x_k^*))$.

Как и в теореме в [178], равновесие по Нэшу в позиционных стратегиях в некооперативном случае каждой усеченной подыгры со случайным горизонтом можно охарактеризовать следующим образом.

Теорема 2.4.1 *Набор стратегий $\{\psi_i^j, j = \overline{k, N}, i \in I\}$ образует ситуацию равновесия по Нэшу в позиционных стратегиях в игре (2.37), (2.38), если существуют функции $V_i^k(x, j)$, где $j \in \{k, \dots, N\}$ и $i \in I$, такие что выполняются следующие рекурсивные соотношения:*

$$\begin{aligned} \hat{V}_i^k(x, N) &= \max_{u_i^N} \left\{ g_i^N(x, \psi_1^N, \dots, \psi_{i-1}^N, u_i^N, \psi_{i+1}^N, \dots, \psi_n^N, \tilde{f}_N(x, u_i^N)) \right\}, \\ \hat{V}_i^k(x, \tau) &= \max_{u_i^\tau} \left\{ g_i^\tau(x, \psi_1^\tau, \dots, \psi_{i-1}^\tau, u_i^\tau, \psi_{i+1}^\tau, \dots, \psi_n^\tau, \tilde{f}_\tau(x, u_i^\tau)) + \right. \\ &\quad \left. + \frac{\sum_{j=\tau+1}^N \gamma_j}{\sum_{j=\tau}^N \gamma_j} \hat{V}_i^k(\tilde{f}_\tau(x, u_i^\tau), \tau + 1) \right\}, \text{ при } \tau \in \{\bar{T}_{k-1} + 1, \dots, N - 1\}, \\ V_i^k(x, \nu) &= \max_{u_i^\nu} \left\{ g_i^\nu(x, \psi_1^\nu, \dots, \psi_{i-1}^\nu, u_i^\nu, \psi_{i+1}^\nu, \dots, \psi_n^\nu, \tilde{f}_\nu(x, u_i^\nu)) + \right. \\ &\quad \left. + V_i^k(\tilde{f}_\nu(x, u_i^\nu), \nu + 1) \right\}, \text{ при } \nu \in \overline{k, \bar{T}_{k-1}}, \\ V_i^k(x, \bar{T}_{k-1} + 1) &= \hat{V}_i^k(x, \bar{T}_{k-1} + 1), \end{aligned} \tag{2.39}$$

где $\tilde{f}_k(x, u_i^k) = f_k(x, \psi_1^k, \dots, \psi_{i-1}^k, u_i^k, \psi_{i+1}^k, \dots, \psi_n^k)$.

Рассмотрим случай, когда все игроки соглашаются сотрудничать в каждой усеченной подыгре. Чтобы максимизировать ожидаемый кооперативный выигрыш в k -ой усеченной подыгре, необходимо решить следующую задачу динамического программирования с дискретным временем:

$$E \left\{ \sum_{i \in I} \sum_{\zeta=k}^{\bar{T}_k} g_i^\zeta(x_\zeta, u_\zeta, x_{\zeta+1}) \right\} = \sum_{i \in I} \left[\sum_{\bar{T}_k=\bar{T}_{k-1}+1}^N \frac{\gamma_{\bar{T}_k}}{\sum_{j=\bar{T}_{k-1}+1}^N \gamma_j} \sum_{\zeta=k}^{\bar{T}_k} g_i^\zeta(x_\zeta, u_\zeta, x_{\zeta+1}) \right] \tag{2.40}$$

при условии (2.37).

В качестве некооперативного случая, можно охарактеризовать оптимальное решение для (2.37) и (2.40) следующим образом:

Следствие 2.4.0.1 Набор стратегий $\{\phi_i^j, j = \overline{k, N}, i \in I\}$ образует оптимальные кооперативные стратегии для игры (2.37), (2.40), если существуют функции $W_i^k(x, j)$, где $j \in \{k, \dots, N\}$ и $i \in I$, такие что выполняются следующие рекурсивные соотношения:

$$\begin{aligned} \hat{W}_i^k(x, N) &= \max_{u_1^N, \dots, u_n^N} \left\{ \sum_{i \in I} g_i^N(x, u_1^N, \dots, u_n^N, f_N(x, u_1^N, \dots, u_n^N)) \right\}, \\ \hat{W}_i^k(x, \tau) &= \max_{u_1^\tau, \dots, u_n^\tau} \left\{ \sum_{i \in I} [g_i^\tau(x, u_1^\tau, \dots, u_n^\tau, f_\tau(x, u_1^\tau, \dots, u_n^\tau))] + \right. \\ &\quad \left. + \frac{\sum_{j=\tau+1}^N \gamma_j}{\sum_{j=\tau}^N \gamma_j} \hat{W}_i^k(f_\tau(x, u_1^\tau, \dots, u_n^\tau), \tau + 1) \right\}, \text{ при } \tau \in \{\bar{T}_{k-1} + 1, \dots, N - 1\}, \\ W_i^k(x, \nu) &= \max_{u_1^\nu, \dots, u_n^\nu} \left\{ \sum_{i \in I} [g_i^\nu(x, u_1^\nu, \dots, u_n^\nu, f_\nu(x, u_1^\nu, \dots, u_n^\nu))] + \right. \\ &\quad \left. + W_i^k(f_\nu(x, u_1^\nu, \dots, u_n^\nu), \nu + 1) \right\}, \text{ при } \nu \in \overline{k, \bar{T}_{k-1}}, \end{aligned} \tag{2.41}$$

где $W_i^k(x, \bar{T}_{k-1} + 1) = \hat{W}_i^k(x, \bar{T}_{k-1} + 1)$.

С помощью множества оптимальных стратегий, можно сгенерировать $\{\phi_i^j, j = \overline{k, N}, i \in I\}$ x_k^* из (2.37) и получить соответствующую кооперативную траекторию.

2.5 Динамическая игровая модель олигополии с рекламой

2.5.1 Исходная игра

Рассмотрим динамическую игровую модель олигополии с рекламой и n фирмами, определенную на временном отрезке $[t_0, T]$. Решение для некооперативного случая этой модели в дифференциальной форме было представлено в [179]. В данном разделе рассматривается динамическое обобщение модели, которое изучается не только для некооперативного, но и (главным образом) для кооперативного случая [114]. Рекламные усилия (стратегии) используются

фирмами для конкуренции на олигополистическом рынке. Каждая фирма стремится увеличить свою долю рынка, в то время как конкуренты пытаются минимизировать ее, используя свои рекламные усилия. Обозначим через x_i^k рыночную долю фирмы $i \in I \equiv \{1, \dots, n\}$ на шаге k . Полный список обозначений представлен в таблице ниже.

Динамическая олигополистическая маркетинговая модель рекламы определяется следующим образом. Уравнение движения фирмы $i \in I$ (динамика ее рыночной доли) имеет вид

$$\begin{cases} x_i^{k+1} = h\left(\frac{n}{n-1}\rho_i u_i^k \sqrt{1-x_i^k} - \frac{1}{n-1} \sum_{j \in I} \rho_j u_j^k \sqrt{1-x_j^k}\right) - (h\delta - 1)x_i^k + h\delta \frac{1}{n}, \\ x_i^0 = z_i^0, \end{cases} \quad (2.42)$$

где z_i^0 — положительная константа.

Функция выигрыша фирмы $i \in I$ в динамической игре (прибыль фирмы $i \in I$)

$$K_i^0(x^0, u^0) = \sum_{l=0}^N m_i x_i^l - (u_i^l)^2. \quad (2.43)$$

В отличие от уравнений движения (2.2) в модельном примере, здесь используем коэффициент дисконтирования $r = 0$. Основная причина такого выбора заключается в следующем: не усложняя результат модели, сначала покажем различие между подходом с динамическим обновлением информации и исходной дифференциальной игрой. Кроме того, мы можем смоделировать рекламный процесс на коротком временном интервале, т.е. для небольшого значения параметра T .

Переменные и параметры модели представлены в таблице ниже.

Таблица 2.1: Переменные и параметры модели

Обозначение	Описание
$x_i^k \leq 1$	Рыночная доля фирмы $i \in I \equiv \{1, \dots, n\}$ на шаге k
$u_i^k \geq 0$	Коэффициент рекламного усилия фирмы $i \in I$ на шаге k
$\rho_i > 0$	Параметр рекламной эффективности фирмы $i \in I$
$\delta > 0$	Параметр оттока клиентов
$m_i > 0$	Объем продаж в отрасли, умноженный на удельную валовую прибыль фирмы $i \in I$
$C(u_i^k)$	Стоимость рекламы фирмы $i \in I$ на шаге k , параметризованная $(u_i^k)^2$

Модель исходной игры определяется на временном интервале $[0, N]$. Предположим, что игроки имеют полную информацию об уравнениях движения и функциях выигрыша на временном интервале $[k, k + \bar{T}]$. Чтобы построить решение в игре с динамическим обновлением информации, используем понятие усеченной подыгры $\Gamma_k(x^{k,0}, k, k + \bar{T})$.

Функция выигрыша фирмы $i \in I$ в усеченной подыгре $\Gamma_k(x_k^0, k, k + \bar{T})$ имеет вид

$$\bar{K}_i^k(x_i^{k,0}, u) = \sum_{j=k}^{k+\bar{T}} m_i x_i^j - (u_i^j)^2. \quad (2.44)$$

В выражении (2.42), доли рынка должны удовлетворять требованию логической согласованности:

$$\sum_{i \in I} x_i^\gamma = 1, \quad x_i^\gamma \geq 0, \quad i \in I, \quad \gamma = \overline{0, N}. \quad (2.45)$$

2.5.2 Некооперативные исходы в усеченной подыгре

Сначала рассмотрим некооперативный случай для каждой усеченной подыгры $\Gamma_k(x_k^0, k, k + \bar{T})$, $k \in \{0, \dots, N - \bar{T}\}$, с начальным условием x_k^0 . Согласно Теореме 2.2.1, набор стратегий $\{\bar{u}_i^{k,l}$, где $i \in \{1, \dots, n\}$, $l = \overline{0, \bar{T}}$ и $k \in \{0, \dots, N - \bar{T}\}\}$, образует ситуацию равновесия по Нэшу в позиционных стратегиях для игры (2.42), (2.44), если существуют функции $V_i^k(l, x_i^{k,l})$, где $i \in \{1, \dots, n\}$, $l = \overline{0, \bar{T}}$ и $k \in \{0, \dots, N - \bar{T}\}$, в каждой усеченной подыгре $\Gamma_k(x_k^0, k, k + \bar{T})$, такие что

выполняются следующие рекурсивные соотношения при каждом $i \in I$:

$$\begin{aligned} V_i^k(l, x_i^{k,l}) &= \max_{u_i^{k,l}} \left\{ m_i x_i^{k,l} - (u_i^{k,l})^2 + V_i^k(l+1, x_i^{k,l+1}) \right\} = \\ &= m_i x_i^{k,l} - (\bar{u}_i^{k,l})^2 + V_i^k(l+1, x_i^{k,l+1}). \end{aligned}$$

Теорема 2.5.1 В каждой усеченной подыгре $\Gamma_k(x_k^0, k, k + \bar{T})$ на временном интервале $[k, k + \bar{T}]$, выигрыш игроков в равновесии по Нэшу в позиционных стратегиях имеет следующий вид:

$$V_i^k(l, x_i^{k,l}) = \sum_{i \in I} \tilde{A}_i^{k,l} x_i^{k,l} + \tilde{B}^{k,l}(i), \quad i \in \{1, \dots, n\}, \quad k \in \{0, \dots, N - \bar{T}\}, \quad l = \overline{0, \bar{T}},$$

где величины $\tilde{A}_i^{k,l}$, $\tilde{B}^{k,l}(i)$ определяются из соотношений

$$\begin{aligned} \tilde{A}_i^{k,l} &= m_i - (\hat{G}_i^{k,l+1} Z_i)^2 - \tilde{A}_i^{k,l+1}(\delta h - 1), \\ \tilde{A}_j^{k,l} &= -2(\hat{G}_j^{k,l+1} Z_j)^2 - \tilde{A}_j^{k,l+1}(\delta h - 1), \quad j \neq i, \quad j \in I \setminus i, \\ \tilde{B}^{k,l}(i) &= \sum_{i \in I} [2(\hat{G}_i^{k,l+1} Z_i)^2 + \frac{\tilde{A}_i^{k,l+1} h \delta}{n}] + \tilde{B}^{k,l+1}(i) - (\hat{G}_i^{k,l+1} Z_i)^2, \end{aligned}$$

с начальными условиями $\tilde{A}_i^{k, \bar{T}+1} = \tilde{B}^{k, \bar{T}+1}(i) = 0$, $i \in \{1, \dots, n\}$, где $\hat{G}_i^{k,l+1} = n\tilde{A}_i^{k,l+1} - \sum_{j \in I} \tilde{A}_j^{k,l+1}$, $Z_i = \frac{h\rho_i}{2(n-1)}$, $u \hat{G}_i^{k,l+1} = \max \left\{ 0, \hat{G}_i^{k,l+1} \right\}$.

Доказательство Запишем уравнение Беллмана для фирмы $i \in I$:

$$V_i^k(l, x_i^{k,l}) = \max_{u_i^{k,l} \geq 0} \left\{ m_i x_i^{k,l} - (u_i^{k,l})^2 + V_i^k(l+1, x_i^{k,l+1}) \right\}.$$

С учетом (2.42) и $V_i^k(l, x_i^{k,l}) = \sum_{i \in I} \tilde{A}_i^{k,l} x_i^{k,l} + \tilde{B}^{k,l}(i)$, где $i \in I$, $l = \overline{0, \bar{T}}$ и $k \in$

$\{0, \dots, N\}$, данное уравнение Беллмана приводится к виду

$$\begin{aligned} \sum_{i \in I} \tilde{A}_i^{k,l} x_i^{k,l} + \tilde{B}^{k,l}(i) &= \max_{u_i^k \geq 0} \left\{ m_i x_i^{k,l} - (u_{i,l}^k)^2 + \sum_{i \in I} \tilde{A}_i^{k,l+1} x_i^{k,l+1} + \tilde{B}^{k,l+1}(i) \right\} = \\ &= \max_{u_i^k \geq 0} \left\{ \sum_{i \in I} \tilde{A}_i^{k,l+1} \left[h \left(\frac{n}{n-1} \rho_i u_i^{k,l} \sqrt{1-x_i^{k,l}} - \frac{1}{n-1} \sum_{j \in I} \rho_j u_j^{k,l} \sqrt{1-x_j^{k,l}} \right) - \right. \right. \\ &\quad \left. \left. - (h\delta - 1)x_i^{k,l} + h\delta \frac{1}{n} \right] + \tilde{B}^{k,l+1}(i) + m_i x_i^{k,l} - (u_i^{k,l})^2 \right\}. \end{aligned} \quad (2.46)$$

В результате максимизации получаем:

$$-2u_i^{k,l} + \frac{h\rho_i}{n-1} n \tilde{A}_i^{k,l+1} \sqrt{1-x_i^{k,l}} - \frac{h\rho_i}{n-1} \sqrt{1-x_i^{k,l}} \sum_{j \in I} \tilde{A}_j^{k,l+1} = 0,$$

при $i \in I$, $l = \overline{0, T}$ и $k \in \{0, \dots, N - \bar{T}\}$. Тогда равновесная по Нэшу позиционная стратегия фирмы $i \in I$ может быть найдена следующим образом:

$$\begin{aligned} \bar{u}_i^{k,l} &= \max \left\{ 0, [n\tilde{A}_i^{k,l+1} - \sum_{j \in I} \tilde{A}_j^{k,l+1}] \frac{h\rho_i}{2(n-1)} \sqrt{1-x_i^{k,l}} \right\} = \\ &= \frac{h\rho_i \sqrt{1-x_i^{k,l}}}{2(n-1)} \max \left\{ 0, n\tilde{A}_i^{k,l+1} - \sum_{j \in I} \tilde{A}_j^{k,l+1} \right\}. \end{aligned}$$

Ожидая, что управления должны быть положительными, подставим их в (2.46). После приведения подобных членов, параметры $\tilde{A}_i^{k,l}$ и $\tilde{B}^{k,l}(i)$ можно выразить в виде

$$\begin{aligned} \tilde{A}_i^{k,l} &= m_i - (\hat{G}_i^{k,l+1} Z_i)^2 - \tilde{A}_i^{k,l+1}(\delta h - 1), \\ \tilde{A}_j^{k,l} &= -2(\hat{G}_j^{k,l+1} Z_j)^2 - \tilde{A}_j^{k,l+1}(\delta h - 1), \quad j \neq i, \quad j \in I \setminus i, \\ \tilde{B}^{k,l}(i) &= \sum_{i \in I} [2(\hat{G}_i^{k,l+1} Z_i)^2 + \frac{\tilde{A}_i^{k,l+1} h\delta}{n}] + \tilde{B}^{k,l+1}(i) - (\hat{G}_i^{k,l+1} Z_i)^2, \end{aligned}$$

где $\hat{G}_i^{k,l+1} = n\tilde{A}_i^{k,l+1} - \sum_{j \in I} \tilde{A}_j^{k,l+1}$, $\hat{G}_i^{k,l+1} = \max \left\{ 0, \hat{G}_i^{k,l+1} \right\}$, $Z_i = \frac{h\rho_i}{2(n-1)}$. \square

Соответствующие равновесные по Нэшу позиционные стратегии вычисля-

ются по формуле

$$\bar{u}_i^{k,l} = \hat{G}_i^{k,l+1} Z_i \sqrt{1 - x_i^{k,l}}, k \in \{0, \dots, N - \bar{T}\}, i \in \{1, \dots, n\}. \quad (2.47)$$

Теперь рассмотрим случай случайного горизонта для усеченной подыгры. Предположим, что игроки владеют некоторой информацией о структуре игры, но при этом не знают временной интервал, в течение которого эта информация соответствует действительности (или не изменилась). Пусть вероятности симметричны на интервале $[1, N]$, $\gamma_1 = \dots = \gamma_N = \frac{1}{N}$. В нашей модели первая усеченная подыгра начинается на шаге 0. Применим Теорему 2.4.1 для k -ой усеченной подыгры: если существуют непрерывно дифференцируемые функции $V_i^k(x, f)$, $i \in \{1, \dots, n\}$, $f \in \{k, \dots, \bar{T}_{k-1}\}$, и $\hat{V}_i^k(x, \tau)$, где $\tau \in \{\bar{T}_{k-1} + 1, \dots, N\}$, такие, что выполняются рекурсивные соотношения

$$\begin{aligned} \hat{V}_i^k(x, N) &= \max_{u_i^N \geq 0} \{m_i x_i^N - (u_i^N)^2\}, \\ \hat{V}_i^k(x, \tau) &= \max_{u_i^\tau \geq 0} \left\{ m_i x_i^\tau - (u_i^\tau)^2 + \frac{N - \tau}{N - \tau + 1} \hat{V}_i^k(x_i^{\tau+1}, \tau + 1) \right\}, \\ &\text{при } \tau \in \{\bar{T}_{k-1} + 1, \dots, N - 1\}, \\ V_i^k(x, f) &= \max_{u_i^f \geq 0} \left\{ m_i x_i^f - (u_i^f)^2 + V_i^k(x_i^{f+1}, f + 1) \right\}, \text{ при } f \in \overline{k, \bar{T}_{k-1}}. \\ V_i^k(x, \bar{T}_{k-1}) &= \hat{V}_i^k(x, \bar{T}_{k-1}), \end{aligned} \quad (2.48)$$

то стратегии u_i , $i \in I$, максимизирующие правую часть системы (2.48), образуют равновесие по Нэшу в позиционных стратегиях. Соответствующие функции Беллмана могут быть получены следующим образом.

Теорема 2.5.2 *Функции Беллмана для равновесия по Нэшу в позиционных стратегиях в k -ой усеченной подыгре в (2.48) имеют вид*

1

$$\hat{V}_i^k(x, \tau) = \sum_{i \in I} A_i^{k,\tau} x_i^\tau + B^{k,\tau}(i), \tau \in \{\bar{T}_{k-1} + 1, \dots, N - 1\}, \quad (2.49)$$

$$\begin{aligned}
& \text{где } A_i^{k,\tau} = m_i - \left(\frac{N-\tau}{N-\tau+1} \bar{Y}_i^{k,\tau+1} Z_i \right)^2 - \frac{N-\tau}{N-\tau+1} A_i^{k,\tau+1} (\delta h - 1), \\
& A_j^{k,\tau} = -2 \left(\frac{N-\tau}{N-\tau+1} \bar{Y}_j^{k,\tau+1} Z_j \right)^2 - \frac{N-\tau}{N-\tau+1} A_j^{k,\tau+1} (\delta h - 1), \quad j \in I \setminus i, \\
& B^{k,\tau}(i) = \sum_{i \in I} \left[2 \left(\frac{N-\tau}{N-\tau+1} \bar{Y}_i^{k,\tau+1} Z_i \right)^2 + \frac{N-\tau}{N-\tau+1} \frac{A_i^{k,\tau+1} h \delta}{n} \right] + \\
& + \frac{N-\tau}{N-\tau+1} B^{k,\tau+1}(i) - \left(\frac{N-\tau}{N-\tau+1} \bar{Y}_i^{k,\tau+1} Z_i \right)^2.
\end{aligned}$$

2

$$V_i^k(x, f) = \sum_{i \in I} A_i^{k,f} x_i^f + B^{k,f}(i), \quad f \in \{k, \dots, \bar{T}_{k-1}\}, \quad (2.50)$$

$$\begin{aligned}
& \text{где } A_i^{k,f} = m_i - \left(\bar{Y}_i^{k,f+1} Z_i \right)^2 - A_i^{k,f+1} (\delta h - 1), \\
& A_j^{k,f} = -2 \left(\bar{Y}_j^{k,f+1} Z_j \right)^2 - A_j^{k,f+1} (\delta h - 1), \quad j \in I \setminus i, \\
& B^{k,f}(i) = \sum_{i \in I} \left[2 \left(\bar{Y}_i^{k,f+1} Z_i \right)^2 + \frac{A_i^{k,f+1} h \delta}{n} \right] + B^{k,f+1}(i) - \left(\bar{Y}_i^{k,f+1} Z_i \right)^2
\end{aligned}$$

с начальными условиями $A_i^{k,N} = m_i$, $B^{k,N+1}(i) = 0$, $i \in \{1, \dots, n\}$, $A_j^{k,N} = 0$, $j \in I \setminus i$, и $Y_i^{k,t+1} = nA_i^{k,t+1} - \sum_{j \in I} A_j^{k,t+1}$, $\bar{Y}_i^{k,t+1} = \max \left\{ 0, Y_i^{k,t+1} \right\}$ $t \in \{k, \dots, N-1\}$.

Доказательство Рассмотрим функцию Беллмана на (последнем) шаге N . В результате максимизации получаем:

$$\psi_i^N = 0, \quad \text{при } i \in I.$$

В силу (2.49) и $\bar{V}_i^k(x_i^N, N) = m_i x_i^N$, второе уравнение в (2.48) принимает вид

$$\sum_{i \in I} A_i^{k,\tau} x_i^\tau + B^{k,\tau}(i) = \max_{u_i^\tau} \left\{ m_i x_i^\tau - (u_i^\tau)^2 + \frac{N-\tau}{N-\tau-1} \left(\sum_{i \in I} A_i^{k,\tau+1} x_i^{\tau+1} + B^{k,\tau+1}(i) \right) \right\}.$$

Вновь вычисляя максимум, получаем:

$$-2(u_i^\tau) + \frac{N-\tau}{N-\tau-1} \left(\frac{h\rho_i}{n-1} n A_i^{k,\tau+1} \sqrt{1-x_i^{k,\tau}} - \frac{h\rho_i}{n-1} \sqrt{1-x_i^{k,\tau}} \sum_{j \in I} A_j^{k,\tau+1} \right) = 0$$

при $i \in I$ и $\tau \in \{\bar{T}_{k-1} + 1, \dots, N-1\}$. Тогда равновесная по Нэшу позиционная

стратегия фирмы $i \in I$ может быть найдена в виде:

$$\begin{aligned} \psi_i^\tau &= \max \left\{ 0, \frac{N - \tau}{N - \tau - 1} (nA_i^{k,\tau+1} - \sum_{j \in I} A_j^{k,\tau+1}) \frac{h\rho_i}{2(n-1)} \sqrt{1 - x_i^{k,\tau}} \right\} = \\ &= \frac{N - \tau}{N - \tau - 1} \frac{h\rho_i \sqrt{1 - x_i^{k,\tau}}}{2(n-1)} \max \left\{ 0, nA_i^{k,\tau+1} - \sum_{j \in I} A_j^{k,\tau+1} \right\}. \end{aligned}$$

Ожидая, что управления должны быть положительными, подставим их во второе уравнение системы (2.48). После приведения подобных членов, параметры $A_i^{k,\tau}$ и $B^{k,\tau}(i)$ можно выразить в виде

$$\begin{aligned} A_i^{k,\tau} &= m_i - \left(\frac{N - \tau}{N - \tau + 1} \bar{Y}_i^{k,\tau+1} Z_i \right)^2 - \frac{N - \tau}{N - \tau + 1} A_i^{k,\tau+1} (\delta h - 1), \\ A_j^{k,\tau} &= -2 \left(\frac{N - \tau}{N - \tau + 1} \bar{Y}_j^{k,\tau+1} Z_j \right)^2 - \frac{N - \tau}{N - \tau + 1} A_j^{k,\tau+1} (\delta h - 1), \quad j \in I \setminus i, \\ B^{k,\tau}(i) &= \sum_{i \in I} \left[2 \left(\frac{N - \tau}{N - \tau + 1} \bar{Y}_i^{k,\tau+1} Z_i \right)^2 + \frac{N - \tau}{N - \tau + 1} \frac{A_i^{k,\tau+1} h \delta}{n} \right] + \\ &+ \frac{N - \tau}{N - \tau + 1} B^{k,\tau+1}(i) - \left(\frac{N - \tau}{N - \tau + 1} \bar{Y}_i^{k,\tau+1} Z_i \right)^2. \end{aligned}$$

На шаге $f \in \{k, \dots, \bar{T}_{k-1}\}$, в результате максимизации в третьем уравнении системы (2.48) получаем

$$\begin{aligned} \psi_i^f &= \max \left\{ 0, (nA_i^{k,f+1} - \sum_{j \in I} A_j^{k,f+1}) \frac{h\rho_i}{2(n-1)} \sqrt{1 - x_i^{k,f}} \right\} = \\ &= \frac{h\rho_i \sqrt{1 - x_i^{k,f}}}{2(n-1)} \max \left\{ 0, nA_i^{k,f+1} - \sum_{j \in I} A_j^{k,f+1} \right\}. \end{aligned}$$

Ожидая, что управления должны быть положительными, подставим их в третье уравнение системы (2.48). После приведения подобных членов, параметры

$A_i^{k,f}$ и $B^{k,f}(i)$ можно выразить в виде

$$\begin{aligned} A_i^{k,f} &= m_i - (\bar{Y}_i^{k+1} Z_i)^2 - A_i^{k,f+1}(\delta h - 1), \\ A_j^{k,f} &= -2(\bar{Y}_j^{k,f+1} Z_j)^2 - A_j^{k,f+1}(\delta h - 1), \quad j \in I \setminus i, \\ B^{k,f}(i) &= \sum_{i \in I} [2(\bar{Y}_i^{k,f+1} Z_i)^2 + \frac{A_i^{k,f+1} h \delta}{n}] + B^{k,f+1}(i) - (\bar{Y}_i^{k,f+1} Z_i)^2, \end{aligned}$$

где $Y_i^{k,t+1} = nA_i^{k,t+1} - \sum_{j \in I} A_j^{k,t+1}$ и $\bar{Y}_i^{k,t+1} = \max \left\{ 0, nA_i^{k,t+1} - \sum_{j \in I} A_j^{k,t+1} \right\}$, $t \in \{k, \dots, N-1\}$. \square

2.5.3 Кооперативные исходы в усеченной подыгре

В этом подразделе рассматривается кооперативная постановка, когда все фирмы соглашаются сотрудничать в каждой усеченной подыгре. Обозначим через $\Gamma_k^c(x_k^0, k, k + \bar{T})$ усеченную кооперативную подыгру на временном интервале $[k, k + \bar{T}]$ с начальным условием x_k^0 . Все фирмы стремятся максимизировать кооперативный выигрыш:

$$\sum_{i \in I} \bar{K}_i^k(x_i^{k,0}, u) = \sum_{i \in I} \sum_{j=k}^{k+\bar{T}} m_i x_i^j - (u_i^j)^2. \quad (2.51)$$

Предположим, что набор стратегий $\{u_i^{*,k,l}, \text{ где } i \in \{1, \dots, n\}, l = \overline{0, \bar{T}} \text{ и } k \in \{0, \dots, N - \bar{T}\}\}$, образует оптимальное управление в игре (2.42), (2.51). Применим Теорему 2.2: если существуют функции $\{W^k(l, x_k^l), x_k^l = (x_1^{k,l}, \dots, x_n^{k,l})\}$, где $i \in \{1, \dots, n\}, l = \overline{0, \bar{T}}$ и $k \in \{0, \dots, N - \bar{T}\}$ в каждой усеченной подыгре $\Gamma_k^c(x_k^0, k, k + \bar{T})$, то выполняются следующие рекурсивные соотношения:

$$\begin{aligned} W^k(l, x_k^l) &= \max_{u_1^{k,l}, \dots, u_n^{k,l} \geq 0} \left\{ \sum_{i \in I} m_i x_i^{k,l} - (u_i^{k,l})^2 + W^k(l+1, x_k^{l+1}) \right\} = \\ &= \sum_{i \in I} m_i x_i^{k,l} - (u_i^{*,k,l})^2 + W^k(l+1, x_k^{l+1}). \end{aligned}$$

Теорема 2.5.3 В каждой усеченной подыгре $\Gamma_k^c(x_k^0, k, k + \bar{T})$ максимальный

кооперативный выигрыш игроков имеет следующий вид:

$$W^k(l, x_k^l) = \sum_{i \in I} \hat{C}_i^{k,l} x_i^{k,l} + \hat{D}_i^{k,l}, \quad (2.52)$$

где $\hat{C}_i^{k,l}$ и $\hat{D}_i^{k,l}$, $i \in \{1, \dots, n\}$, $k \in \{0, \dots, N - \bar{T}\}$, $l = \overline{0, \bar{T}}$, удовлетворяют соотношениям

$$\begin{aligned} \hat{C}_i^{k,l} &= m_i - (\bar{G}_i^{k,l+1} Z_i)^2 - \hat{C}_i^{k,l+1}(\delta h - 1), \\ \hat{D}_i^{k,l} &= (\bar{G}_i^{k,l+1} Z_i)^2 + \frac{\hat{C}_i^{k,l+1} \delta h}{n} + D_i^{k,l+1} \end{aligned}$$

с начальными условиями $\hat{C}_i^{k, \bar{T}+1} = 0$, $\hat{D}_i^{k, \bar{T}+1} = 0$, $\bar{G}_i^{k,l+1} = n\hat{C}_i^{k,l+1} - \sum_{j \in I} \hat{C}_j^{k,l+1}$ и $\bar{G}_i^{k,l+1} = \max \{0, \bar{G}_i^{k,l+1}\}$.

Доказательство Запишем уравнение Беллмана для кооперативной постановки:

$$W^k(l, x_i^{k,l}) = \max_{u_1^{k,l}, \dots, u_n^{k,l}} \sum_{i \in I} \left\{ m_i x_i^{k,l} - (u_i^{*,k,l})^2 + W^k(l+1, x_i^{k,l+1}) \right\}. \quad (2.53)$$

Подставляя (2.52) в уравнение (2.53), получаем:

$$\begin{aligned} W^k(l, x_k^l) &= \max_{u_1^{k,l}, \dots, u_n^{k,l}} \sum_{i \in I} \left\{ m_i x_i^{k,l} - (u_i^{*,k,l})^2 + W^k(l+1, x_k^{l+1}) \right\} \\ \Rightarrow \sum_{i \in I} \hat{C}_i^{k,l} x_i^{k,l} + \hat{D}_i^{k,l} &= \max_{u_1^k, \dots, u_n^k} \left\{ \sum_{i \in I} [m_i x_i^{k,l} - (u_i^{k,l})^2] + \sum_{i \in I} \hat{C}_i^{k,l+1} x_i^{k,l+1} + \hat{D}_i^{k,l+1} \right\}, \end{aligned} \quad (2.54)$$

$$W^k(\bar{T} + 1, x_k^{\bar{T}}) = 0.$$

Далее, подставим правую часть выражения (2.42) в (2.54). В результате максимизации в (2.54) получаем:

$$-2u_i^k + (n\hat{C}_i^{k,l+1} - \sum_{j \in I} \hat{C}_j^{k,l+1}) \frac{h\rho_i \sqrt{1 - x_i^{k,l}}}{n-1} = 0.$$

Оптимальные кооперативные стратегии могут быть найдены по формуле

$$u_i^{*,k,l} = \max \left\{ 0, \left(n\hat{C}_i^{k,l+1} - \sum_{j \in I} \hat{C}_j^{k,l+1} \right) \frac{h\rho_i \sqrt{1 - x_i^{k,l}}}{2(n-1)} \right\} =$$

$$= \frac{h\rho_i \sqrt{1 - x_i^{k,l}}}{2(n-1)} \max \left\{ 0, n\hat{C}_i^{k,l+1} - \sum_{j \in I} \hat{C}_j^{k,l+1} \right\}$$

при $i \in I$, $l = \overline{0, \bar{T}}$ и $k \in \{0, \dots, N\}$.

Ожидая, что кооперативные стратегии должны быть положительными, подставим эту формулу в (2.54). Следовательно, $\hat{C}_i^{k,l}$ и $\hat{D}_i^{k,l}$ выражаются в виде

$$\hat{C}_i^{k,l} = m_i - (\bar{G}_i^{k,l+1} Z_i)^2 - \hat{C}_i^{k,l+1}(\delta h - 1),$$

$$\hat{D}_i^{k,l} = (\bar{G}_i^{k,l+1} Z_i)^2 + \frac{\hat{C}_i^{k,l+1} \delta h}{n} + D_i^{k,l+1}.$$

□

В каждой усеченной подыгре оптимальные кооперативные стратегии имеют вид

$$u_i^{*,k,l} = \bar{G}_i^{k,l+1} Z_i \sqrt{1 - x_i^{k,l}}, \forall i \in I.$$

Обозначим через $x_k^{*,l} = (x_1^{*,k,l}, \dots, x_n^{*,k,l})$, $l = \overline{0, \bar{T}}$, решение (2.52) с оптимальными стратегиями u^* и кооперативной траекторией. В каждой усеченной подыгре с номером $k \in \{0, \dots, N - \bar{T}\}$ кооперативная траектория может быть рассчитана с помощью (2.42) следующим образом:

$$x_i^{*,k,l+1} = 2n\bar{G}_i^{k,l+1} Z_i^2 (1 - x_i^{*,k,l}) - \sum_{j \in I} 2\bar{G}_j^{k,l+1} Z_j^2 (1 - x_j^{*,k,l}) - (h\delta - 1)x_i^{*,k,l} + \frac{h\delta}{n}.$$

Таким образом, мы формализовали случай, когда все фирмы сотрудничают в модели игры с динамическим обновлением информации и фиксированным информационным горизонтом. Далее, изучим случай, когда информационный горизонт является случайным, а случайная величина \bar{T}_1 имеет симметричные вероятности $\gamma_1 = \dots = \gamma_N = \frac{1}{N}$ на интервале $[1, N]$. Рассмотрим k -ую

кооперативную усеченную подыгру. Если существуют функции $W^k(x, f)$, $i \in \{1, \dots, n\}$, $f \in \{k, \dots, \bar{T}_{k-1}\}$, и $\hat{W}^k(x, \tau)$, $\tau \in \{\bar{T}_{k-1} + 1, \dots, N\}$, такие что выполняются рекурсивные отношения

$$\begin{aligned} \hat{W}^k(x, N) &= \max_{u_1^N, \dots, u_n^N} \left\{ \sum_{i \in I} m_i x_i^N - (u_i^N)^2 \right\}, \\ \hat{W}^k(x, \tau) &= \max_{u_1^\tau, \dots, u_n^\tau} \left\{ \sum_{i \in I} m_i x_i^\tau - (u_i^\tau)^2 + \frac{N - \tau}{N - \tau + 1} \hat{W}^k(x_i^{\tau+1}, \tau + 1) \right\}, \\ &\text{при } \tau \in \{\bar{T}_{k-1} + 1, \dots, N - 1\}, \\ W^k(x, f) &= \max_{u_1^f, \dots, u_n^f} \left\{ \sum_{i \in I} m_i x_i^f - (u_i^f)^2 + W^k(x_i^{f+1}, f + 1) \right\}, \text{ при } f \in \overline{k, \bar{T}_{k-1}}, \\ W_i^k(x, \bar{T}_{k-1}) &= \hat{W}_i^k(x, \bar{T}_{k-1}), \end{aligned} \tag{2.55}$$

то стратегии $\{u_i^{*,k,l}, \text{ for } i \in \{1, \dots, n\}, l = \overline{0, \bar{T}} \text{ и } k \in \{0, \dots, N - \bar{T}\}\}$ являются оптимальными.

Функция значения может быть получена следующим образом.

Теорема 2.5.4 *Функции Беллмана в k -ой кооперативной усеченной подыгре в (2.55) имеют вид*

1

$$\hat{W}^k(x, \tau) = \sum_{i \in I} \tilde{C}_i^{k,\tau} x_i^\tau + \tilde{D}_i^{k,\tau}, \tau \in \{\bar{T}_{k-1} + 1, \dots, N - 1\}, \tag{2.56}$$

$$\begin{aligned} \text{где } \tilde{C}_i^{k,\tau} &= m_i - \left(\frac{N - \tau}{N - \tau + 1} \hat{G}_i^{k,\tau+1} Z_i \right)^2 - \frac{N - \tau}{N - \tau + 1} \tilde{C}_i^{k,\tau+1} (\delta h - 1), \\ \tilde{D}_i^{k,\tau} &= \left(\frac{N - \tau}{N - \tau + 1} \hat{G}_i^{k,\tau+1} Z_i \right)^2 + \frac{N - \tau}{N - \tau + 1} \frac{\tilde{C}_i^{k,\tau+1} h \delta}{n} + \frac{N - \tau}{N - \tau + 1} \tilde{D}_i^{k,\tau+1}; \end{aligned}$$

2

$$W_i^k(x, f) = \sum_{i \in I} \tilde{C}_i^{k,f} x_i^f + \tilde{D}_i^{k,f}, f \in \{k, \dots, \bar{T}_{k-1}\}, \tag{2.57}$$

$$\text{где } \tilde{C}_i^{k,l} = m_i - \left(\hat{G}_i^{k,l+1} Z_i \right)^2 - \tilde{C}_i^{k,l+1} (\delta h - 1),$$

$$\tilde{D}_i^{k,l} = \left(\hat{G}_i^{k,l+1} Z_i \right)^2 + \frac{\tilde{C}_i^{k,l+1} \delta h}{n} + \tilde{D}_i^{k,l+1},$$

с начальными условиями $\tilde{C}_i^{k,N} = m_i$, $\tilde{D}_i^{k,N} = 0$, $i \in \{1, \dots, n\}$, $\hat{G}_i^{k,t+1} =$

$$n\tilde{C}_i^{k,t+1} - \sum_{j \in I} \tilde{C}_j^{k,t+1}, \hat{G}_i^{k,t+1} = \max \left\{ 0, \hat{G}_i^{k,t+1} \right\}, t \in \{k, \dots, N-1\}.$$

Доказательство Рассмотрим функцию Беллмана на (последнем) шаге N . В результате максимизации получаем:

$$\phi_i^N = 0, \text{ for } i \in I.$$

В силу (2.56) и $\bar{W}_i^k(x_i^N, N) = \sum_{i \in I} m_i x_i^N$, второе уравнение в (2.55) принимает вид

$$\sum_{i \in I} \tilde{C}_i^{k,\tau} x_i^\tau + \tilde{D}_i^{k,\tau} = \max_{u_1^\tau, \dots, u_n^\tau} \left\{ \sum_{i \in I} m_i x_i^\tau - (u_i^\tau)^2 + \frac{N-\tau}{N-\tau+1} (\tilde{C}_i^{k,\tau+1} x_i^{\tau+1} + \tilde{D}_i^{k,\tau+1}) \right\}.$$

Вновь вычисляя максимум, получаем:

$$-2(u_i^\tau) + \frac{N-\tau}{N-\tau-1} \left(\frac{h\rho_i}{n-1} n\tilde{C}_i^{k,\tau+1} \sqrt{1-x_i^{k,\tau}} - \frac{h\rho_i}{n-1} \sqrt{1-x_i^{k,\tau}} \sum_{j \in I} \tilde{C}_j^{k,\tau+1} \right) = 0$$

при $i \in I$ и $\tau \in \{\bar{T}_{k-1} + 1, \dots, N-1\}$. Тогда равновесная по Нэшу позиционная стратегия фирмы $i \in I$ может быть найдена в виде

$$\begin{aligned} \phi_i^\tau &= \max \left\{ 0, \frac{N-\tau}{N-\tau-1} \left(n\tilde{C}_i^{k,\tau+1} - \sum_{j \in I} \tilde{C}_j^{k,\tau+1} \right) \frac{h\rho_i}{2(n-1)} \sqrt{1-x_i^{k,\tau}} \right\} = \\ &= \frac{N-\tau}{N-\tau-1} \frac{h\rho_i \sqrt{1-x_i^{k,\tau}}}{2(n-1)} \max \left\{ 0, n\tilde{C}_i^{k,\tau+1} - \sum_{j \in I} \tilde{C}_j^{k,\tau+1} \right\}. \end{aligned}$$

Ожидая, что управления должны быть положительными, подставим эту формулу во второе уравнение системы (2.55). После приведения подобных членов, параметры $\tilde{C}_i^{k,\tau}$ и $\tilde{D}_i^{k,\tau}$ можно выразить в виде

$$\begin{aligned} \tilde{C}_i^{k,\tau} &= m_i - \left(\frac{N-\tau}{N-\tau+1} \hat{G}_i^{k,\tau+1} Z_i \right)^2 - \frac{N-\tau}{N-\tau+1} \tilde{C}_i^{k,\tau+1} (\delta h - 1), \\ \tilde{D}_i^{k,\tau} &= \left(\frac{N-\tau}{N-\tau+1} \hat{G}_i^{k,\tau+1} Z_i \right)^2 + \frac{N-\tau}{N-\tau+1} \frac{\tilde{C}_i^{k,\tau+1} h\delta}{n} + \frac{N-\tau}{N-\tau+1} \tilde{D}_i^{k,\tau+1}. \end{aligned}$$

На шаге $f \in \{k, \dots, \bar{T}_{k-1}\}$, в результате максимизации в третьем уравнении

системы (2.55) получаем:

$$\begin{aligned}\phi_i^f &= \max \left\{ 0, (n\tilde{C}_i^{k,\tau+1} - \sum_{j \in I} \tilde{C}_j^{k,\tau+1}) \frac{h\rho_i}{2(n-1)} \sqrt{1 - x_i^{k,f}} \right\} = \\ &= \frac{h\rho_i \sqrt{1 - x_i^{k,f}}}{2(n-1)} \max \left\{ 0, n\tilde{C}_i^{k,\tau+1} - \sum_{j \in I} \tilde{C}_j^{k,\tau+1} \right\}.\end{aligned}$$

Ожидая, что управления должны быть положительными, подставим эту формулу в третье уравнение системы (2.55). После приведения подобных членов, параметры $\tilde{C}_i^{k,f}$ и $\tilde{D}_i^{k,f}$ можно выразить в виде

$$\begin{aligned}\tilde{C}_i^{k,l} &= m_i - (\hat{G}_i^{k,l+1} Z_i)^2 - \tilde{C}_i^{k,l+1}(\delta h - 1), \\ \tilde{D}_i^{k,l} &= (\hat{G}_i^{k,l+1} Z_i)^2 + \frac{\tilde{C}_i^{k,l+1} \delta h}{n} + D_i^{k,l+1},\end{aligned}$$

где $\hat{G}_i^{k,t+1} = n\tilde{C}_i^{k,t+1} - \sum_{j \in I} \tilde{C}_j^{k,t+1}$, $\hat{G}_i^{k,t+1} = \max \{0, \hat{G}_i^{k,t+1}\}$, $t \in \{k, \dots, N-1\}$.

□

2.5.4 Характеристическая функция в усеченной подыгре

Используя формулу характеристической функции (2.14), определим характеристическую функцию для каждой усеченной подыгры:

$$V^k(S; x_k^{*,0}) = \begin{cases} \sum_{i \in I} K_i^k(x_k^{*,0}; u), & S = I, \\ \tilde{V}_k(S, x_k^{*,0}), & S \subset I, \\ 0, & S = \emptyset, \end{cases} \quad (2.58)$$

где $\tilde{V}_k(S, x_k^{*,0})$ — общий выигрыш игроков из коалиции S в равновесии по Нэшу в позиционных стратегиях $u^{NE} = (\bar{u}_1, \dots, \bar{u}_n)$ в игре со следующим множеством игроков: коалицией S (действующей как один игрок) и игроками из множества $I \setminus S$, т.е. в игре с $|I \setminus S| + 1$ игроками.

Предположим, что функция Беллмана для коалиции $S \subseteq I$ и игроков $i \in I \setminus S$

в усеченной подыгре $\Gamma_k^c(x_k^0, k, k + \bar{T})$ имеет следующий вид:

$$\begin{aligned} V^k(S, x_k^0) &= \sum_{i \in I} \tilde{C}_i^{k,0} x_i^{k,0} + \bar{E}^{k,0}, \quad S \subseteq I \\ V^k(\{i\}, x_k^0) &= \sum_{i \in I} \hat{A}_i^{k,0} x_i^{k,0} + \hat{B}(i)^{k,0}, \quad i \in I \setminus S. \end{aligned}$$

Запишем соответствующую систему уравнений Беллмана:

$$\begin{cases} V^k(S, x_k^0) = \max_{u_i^k, \forall i \in S} \left\{ \sum_{i \in S} [m_i x_i^{k,0} - (u_i^k)^2] + V^k(S, x_k^1) \right\}, \\ V^k(\{i\}, x_k^0) = \max_{u_i^k} \left\{ m_i x_i^{k,0} - (u_i^k)^2 + V^k(\{i\}, x_k^1) \right\}, \quad \forall i \in I \setminus S, \\ V^k(S, x_k^{\bar{T}+1}) = 0, \\ V^k(\{i\}, x_k^{\bar{T}+1}) = 0. \end{cases} \quad (2.59)$$

2.5.5 Численное моделирование динамической игровой модели олигополии с рекламой и обновлением информации

Рассмотрим частный случай олигополии с тремя фирмами на временном интервале из $N = 8$ шагов, с фиксированным информационным горизонтом $\bar{T} = 2$. В качестве иллюстрации, выберем параметры модели для фирм следующим образом: $\rho = [0.4, 0.5, 0.3]$, $h = 0.4$, $\delta = 0.09$, $m = [0.6, 1, 1.2]$. Зададим начальные условия $z^0 = [0.3, 0.5, 0.2]$.

На рисунках 2.2 и 2.3 показаны соответственно равновесные по Нэшу позиционные стратегии фирм (некооперативный случай) и их кооперативные стратегии (кооперативный случай), как для исходной игровой модели, так и для игровой модели с динамическим обновлением информации.

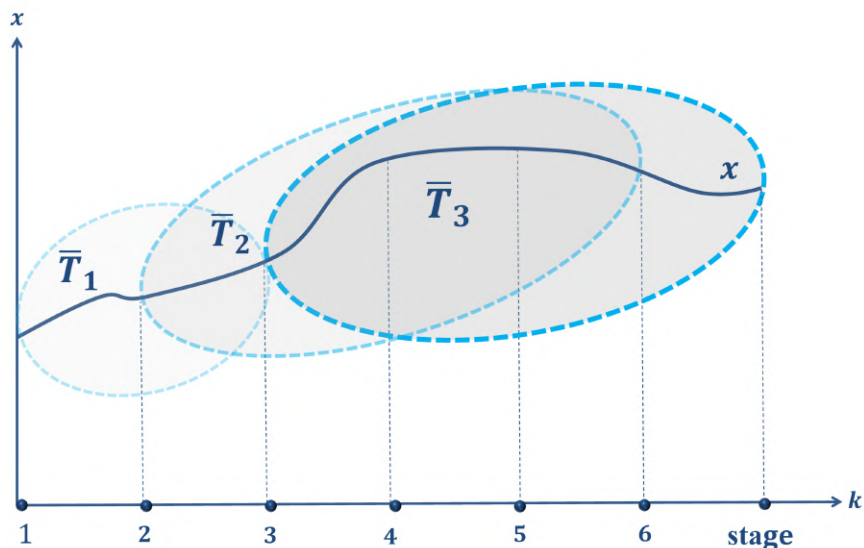


Рис. 2.1: Длина каждого голубого овала равна реализации бесконечного горизонта.

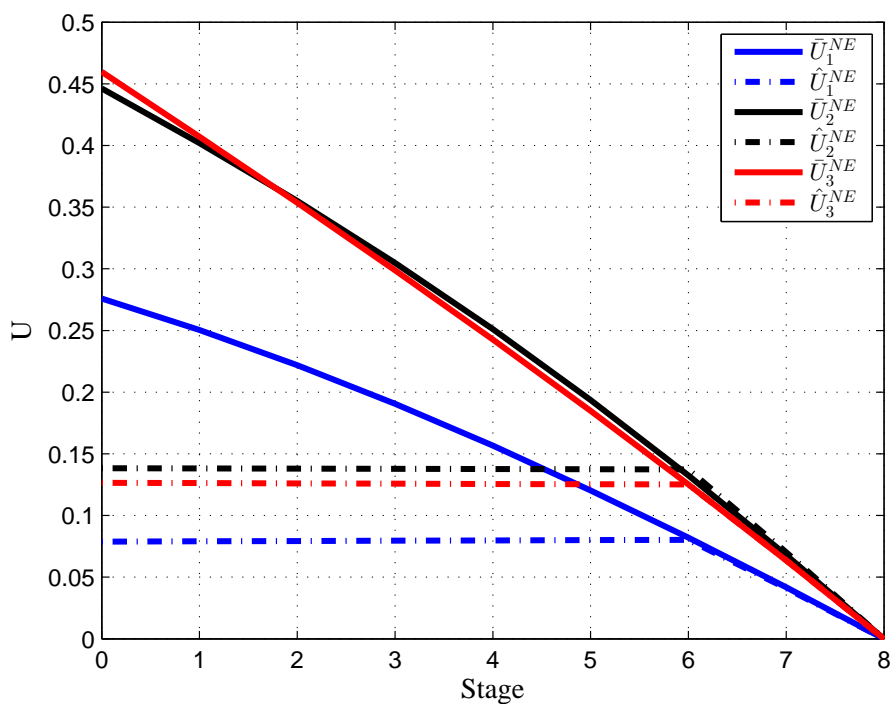


Рис. 2.2: Равновесие по Нэшу в позиционных стратегиях в некооперативной исходной игре (сплошная линия) и равновесие по Нэшу в позиционных стратегиях в некооперативной игре с динамическим обновлением информации (пунктирная линия).

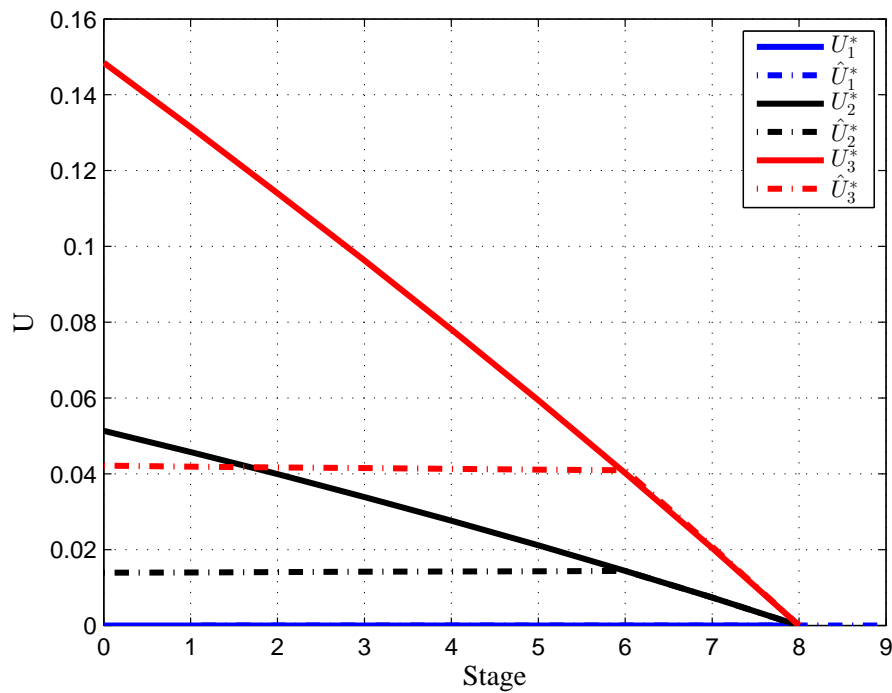


Рис. 2.3: Оптимальные кооперативные стратегии в кооперативной исходной игре (сплошная линия) и оптимальные кооперативные стратегии в кооперативной игре с динамическим обновлением информации (пунктирная линия).

На рисунках 2.4 и 2.5 показаны соответственно динамика рыночных долей игроков при равновесных по Нэшу позиционных стратегиях фирм в некооперативном случае и динамика их рыночных долей при оптимальных кооперативных стратегиях в кооперативном случае, как для исходной игровой модели, так и для игровой модели с динамическим обновлением информации.

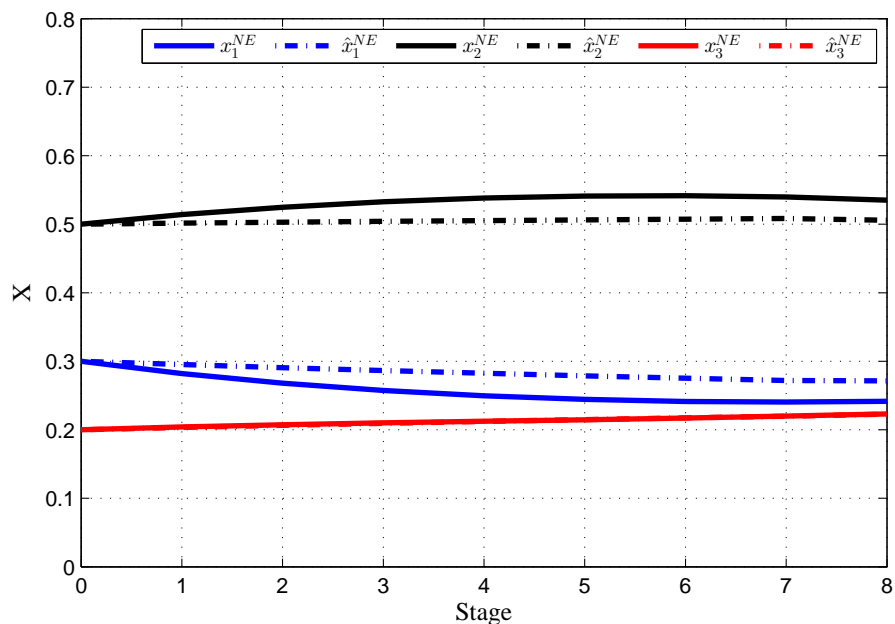


Рис. 2.4: Некооперативная траектория в исходной игре (сплошная линия) и результирующая некооперативная траектория в игре с динамическим обновлением информации (пунктирная линия).

На рисунках 2.6 и 2.7 показаны соответственно оптимальные выигрыши фирм в некооперативном случае и характеристическая функция для коалиций $\{1, 2\}$, $\{1, 3\}$ и $\{2, 3\}$ в кооперативном случае, как для исходной игровой модели, так и для игровой модели с динамическим обновлением информации.

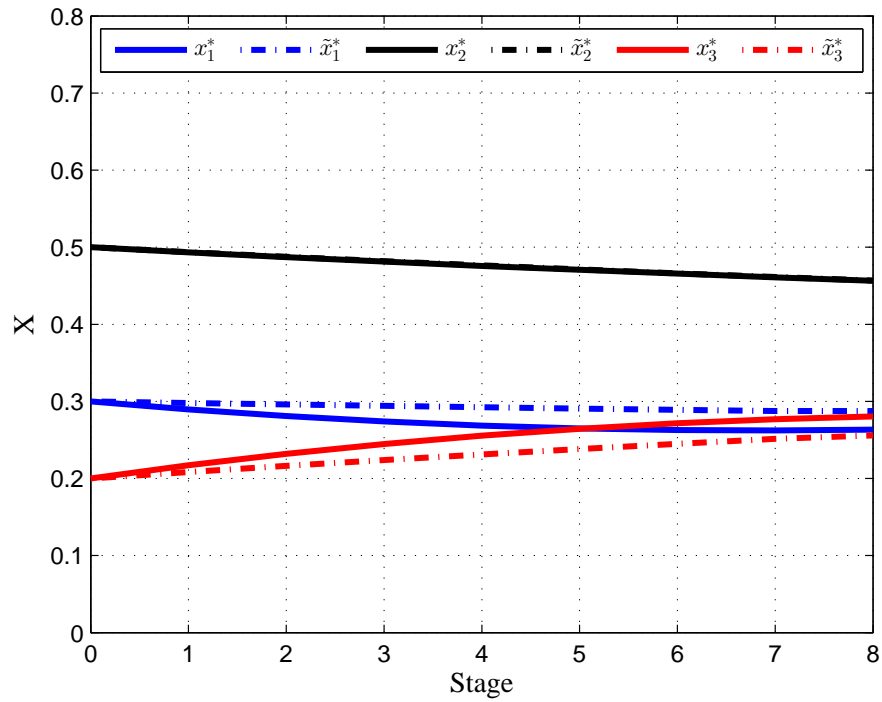


Рис. 2.5: Кооперативная траектория в исходной игре (сплошная линия) и результирующая кооперативная траектория в игре с динамическим обновлением информации (пунктирная линия).

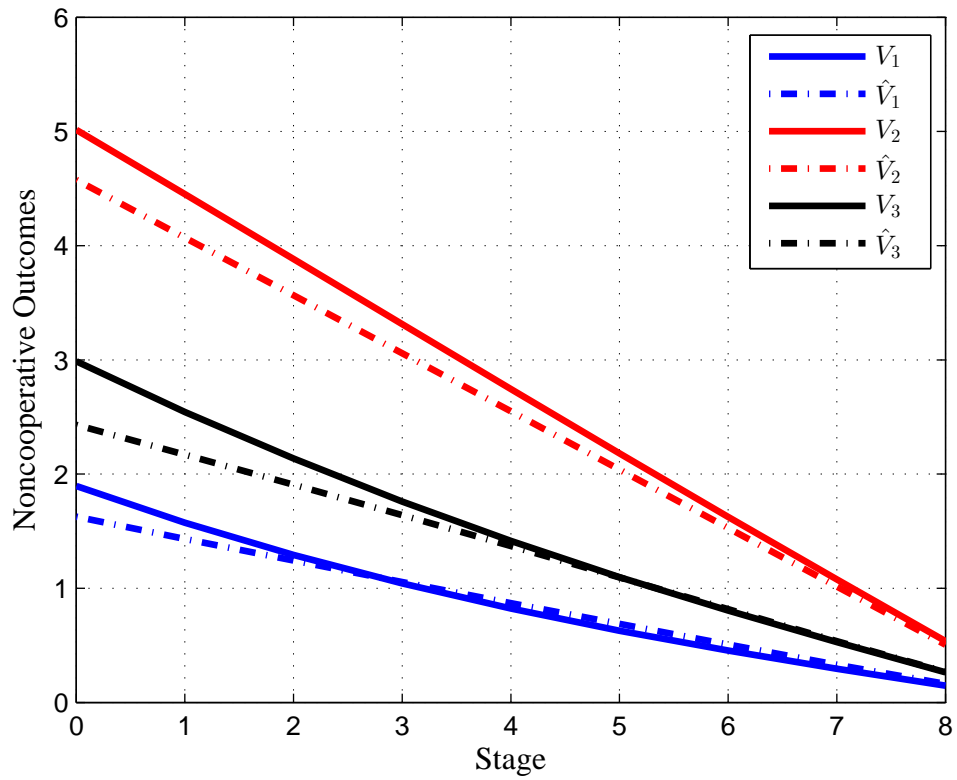


Рис. 2.6: Некооперативные исходы в исходной игре (сплошная линия) и результирующие некооперативные исходы в игре с динамическим обновлением информации (пунктирная линия).

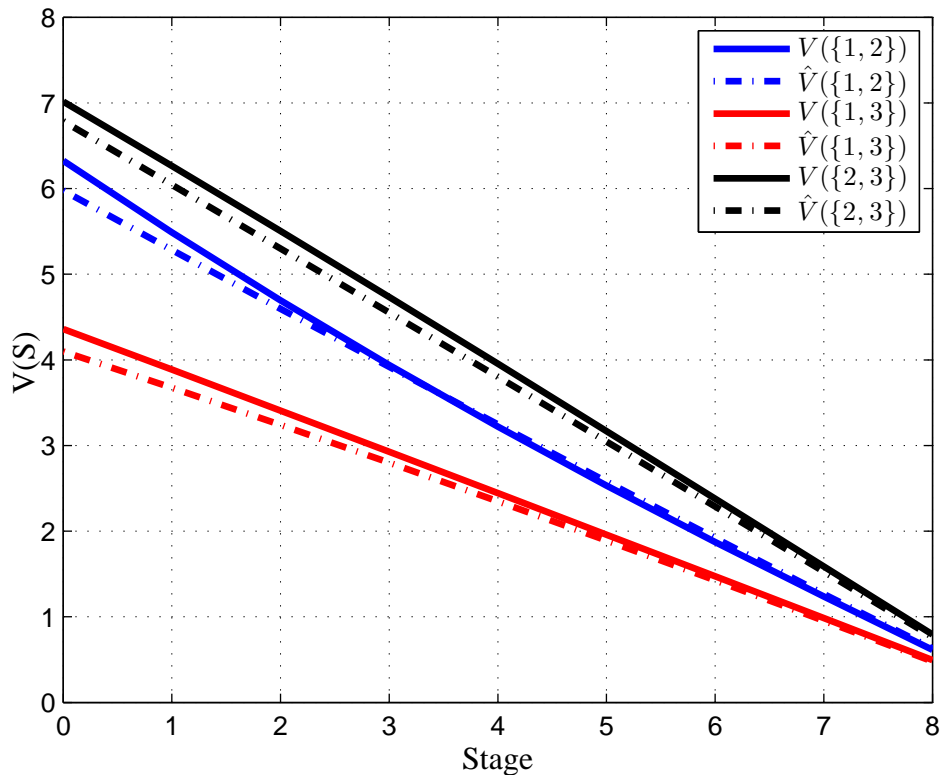


Рис. 2.7: Характеристическая функция в исходной игре (сплошная линия) и результирующая характеристическая функция в игре с динамическим обновлением информации (пунктирная линия).

На рисунке 2.8 показано различие между результирующим вектором Шепли и вектором Шепли в исходной игре. Как видно на рисунке, результирующий вектор Шепли для игры с плавающим информационным горизонтом изменяется более устойчивым образом от шага к шагу, чем вектор Шепли в исходной игре. Конкуренция между фирмами является более жесткой в исходной игре, чем в игре с плавающим информационным горизонтом, учитывая рыночные доли на рисунке 2.5 и затраты на рекламу на рисунке 2.6. Вектор Шепли каждой фирмы практически не изменяется на последнем шаге, а это означает, что и прибыль каждой фирмы не сильно изменяется (см. рисунок 2.8).

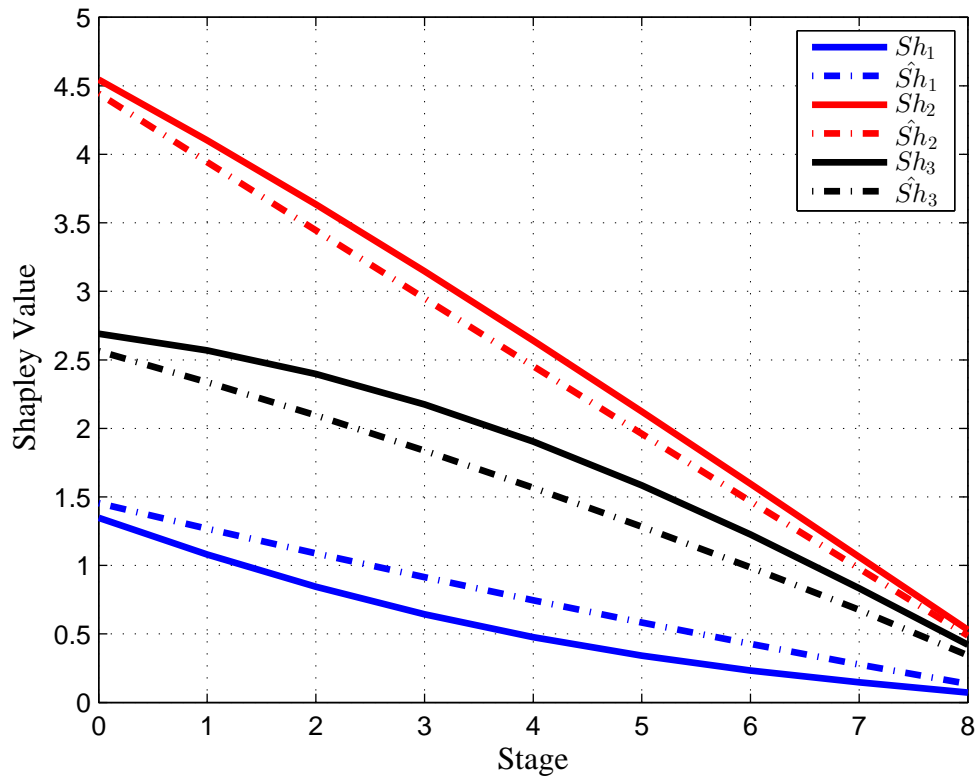


Рис. 2.8: Вектор Шепли в исходной игре (сплошная линия) и результирующий вектор Шепли в игре с динамическим обновлением информации (пунктирная линия).

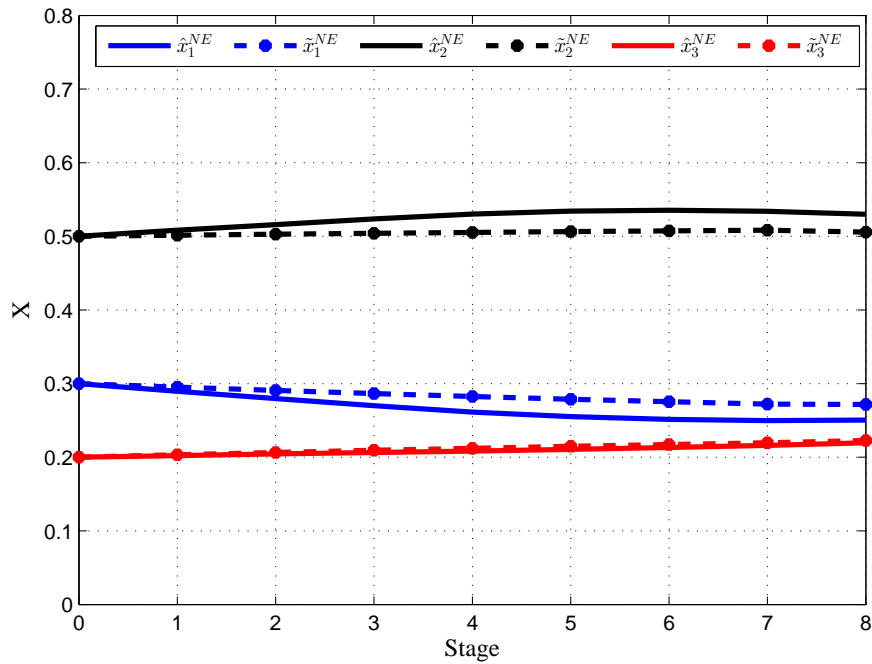


Рис. 2.9: Результирующая некооперативная траектория с фиксированным информационным горизонтом (пунктирная линия) и результирующая некооперативная траектория со случайным информационным горизонтом (сплошная линия).

Предположим, что информационный горизонт игры получил реализацию $\bar{T}_1 = 2$, $\bar{T}_2 = 5$, $\bar{T}_3 = 8$. На рисунке 2.9 показано различие между результирующей некооперативной траекторией с фиксированным информационным горизонтом и случайным информационным горизонтом.

На рисунке 2.10 показано различие между результирующими некооперативными исходами для фирм при фиксированном информационном горизонте и случайном информационном горизонте.

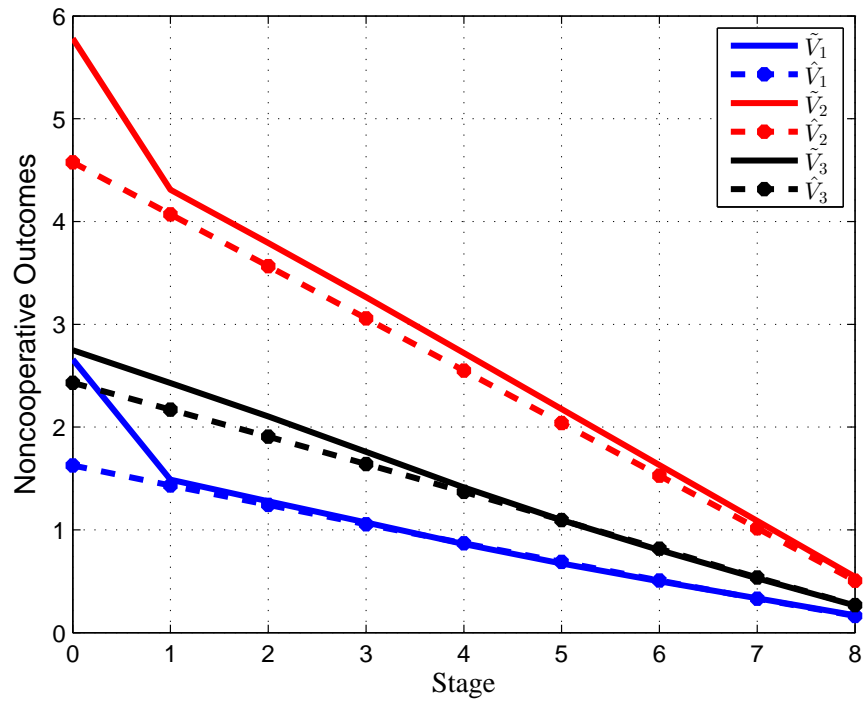


Рис. 2.10: Результирующие некооперативные исходы при фиксированном информационном горизонте (пунктирная линия) и результирующие некооперативные исходы при случайном информационном горизонте (сплошная линия).

На рисунке 2.11 показано различие между результирующими кооперативными траекториями при фиксированном и случайном информационном горизонте.

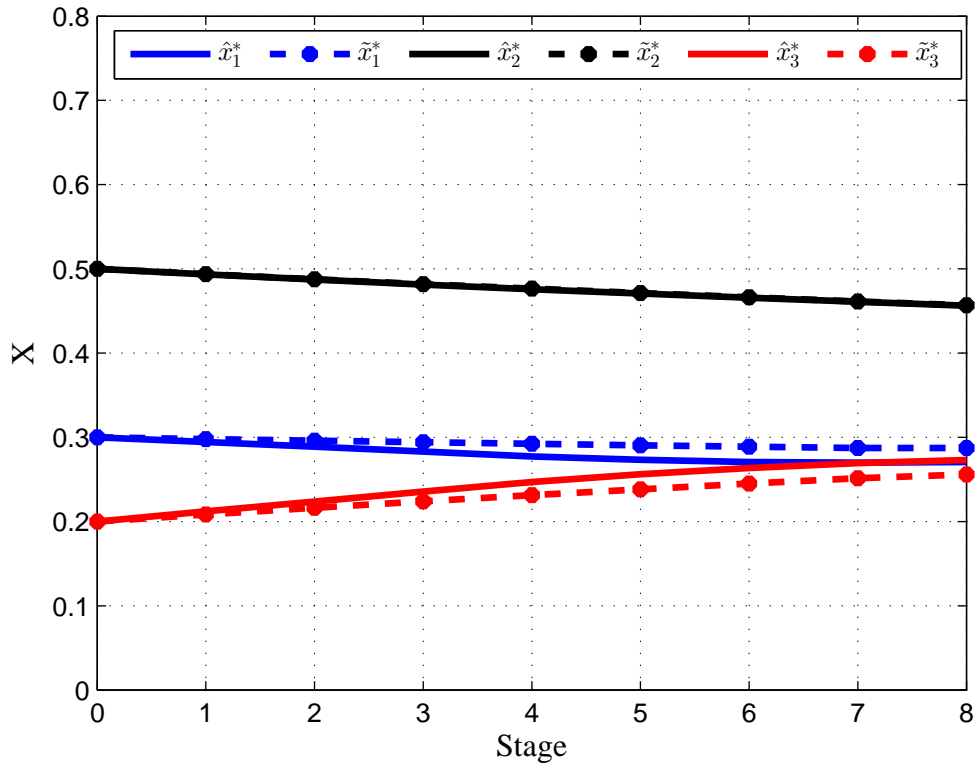


Рис. 2.11: Результирующая кооперативная траектория при фиксированном информационном горизонте (пунктирная линия) и результирующая кооперативная траектория при случайном информационном горизонте (сплошная линия).

Глава 3

Дифференциальные игры с динамическим обновлением информации

В этом разделе представлено описание и результаты для подхода с динамическим обновлением информации в классе кооперативных дифференциальных игр с трансферабельной полезностью. В отличие от предыдущего раздела, подход динамического обновления информации используется для моделей дифференциальных игр. Подход применяется для построения непрерывных теоретико-игровых моделей и определения решений для конфликтно-управляемых процессов, когда информация о процессе обновляется динамически. Изучается свойство динамической устойчивости кооперативных решений с динамическим обновлением информации, а также свойства, связанные с переходом к решению на усеченных временных интервалах, и свойства кооперативных решений, определенных на всем временном интервале игры. Рассматриваются конкретные модели обновления информации, такие как динамическое обновление информации со стохастическим прогнозом и динамическое обновление информации со случайным горизонтом. В конце раздела изучается прикладная модель нефтяного рынка, в которой реальное кооперативное соглашение описывается коалиционной кооперативной игрой с нетрансферабельной полезностью, а также проводится анализ исторических данных.

3.1 Общий класс кооперативных дифференциальных игр с динамическим обновлением информации

3.1.1 Исходная дифференциальная игра

Рассмотрим дифференциальную игру n лиц $\Gamma(x_0, T-t_0)$ с конечным горизонтом $T-t_0$ и начальным состоянием $x_0 \in X \subset R^m$, где X — пространство состояний, а t_0 — начальный момент времени. (t_0 и T являются фиксированными величинами.) Обозначим множество игроков через $N = \{1, \dots, n\}$. В каждый момент времени игрок $i \in N$ выбирает управление или стратегию $u_i \in U_i \subset \text{Comp}R^k$. Пространство всех стратегий (управлений) игроков имеет вид $U = U_1 \times \dots \times U_n$. Зададим функцию выигрыша игрока $i \in N$ следующим образом:

$$K_i(x_0, T-t_0; u) = \int_{t_0}^T h_i(\tau, x(\tau), u(\tau, x)) e^{-r(\tau-t_0)} d\tau \quad (3.1)$$

при условии

$$\dot{x} = g(t, x(t), u(t, x)), \quad x(t_0) = x_0, \quad (3.2)$$

где $x(t) \in X \subset R^m$ — траектория (решение) динамической системы (3.2) при управлении $u = (u_1, \dots, u_n) \in U$, $r \geq 0$, а функции $h_i(t, x, u) : [t_0, T] \times X \times U \rightarrow R$ и $g(t, x, u) : [t_0, T] \times X \times U \rightarrow R$ являются дифференцируемыми. Функция $h_i(t, x, u)$ описывает мгновенный выигрыш игрока $i \in N$ в состоянии $x(t) \in R^m$ при управлении $u = (u_1, \dots, u_n)$, а функция $e^{-r(t-t_0)}$ определяет коэффициент дисконтирования для выигрышей игроков. Решение системы (3.2) задает траекторию игры.

При использовании программных стратегий, требуется кусочная непрерывность с конечным числом разрывов. Для позиционных стратегий применяется подход из [180], т.е. считается, что любой набор стратегий $u(t, x) = (u_1(t, x), \dots, u_n(t, x))$ является допустимым в смысле условий в п.1.1.1. Это означает, что для любого n -элементного набора стратегий $u(t, x) = (u_1(t, x), \dots, u_n(t, x))$ будем считать, решение задачи Коши (3.2) существует, единственно и продолжимо на временном интервале $[t_0, T]$. Существенно более тонкий подход необходимый при определении стратегий в антагонистической игре предложен

и реализован в [181].

3.1.2 Усеченная подыгра

Как упоминалось выше, информация о функциях $g(t, x, u)$ и $h_i(t, x, u)$ обновляется в фиксированные моменты времени. Предположим, что интервал $[t_0, T] \subset R$ разделен на $l = \frac{T-t_0}{\Delta t} - 1$ подинтервалов $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, l$, равной длины $\Delta t \geq 0$. Тогда в течение временного интервала $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$ игроки имеют полную информацию о динамике игры (уравнениях движения) $g(t, x, u)$ и функции выигрыша $h_i(t, x, u)$ на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, где \bar{T} — фиксированная величина (информационный горизонт). Точнее говоря, информация об игре описывается следующим образом: на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ динамика игры имеет вид $g(t, x, u) = g_k(t, x, u)$, а выигрыш — вид $h_i(t, x, u) = h_i^k(t, x, u)$, где функции $g_k(t, x, u)$, $h_i^k(t, x, u) : [t_0 + k\Delta t, t_0 + (k+1)\Delta t] \times X \times R$, $k = j, \dots, j + \frac{\bar{T}}{\Delta t}$. В момент времени $t = t_0 + (j+1)\Delta t$ обновляется информация о динамике игры $g(t, x, u)$ и выигрыше $h_i(t, x, u)$. В течение следующего временного интервала $(t_0 + (j+1)\Delta t, t_0 + (j+2)\Delta t]$ игроки получают полную информацию о $g(t, x, u)$ и выигрыше $h_i(t, x, u)$ на временном интервале $(t_0 + (j+1)\Delta t, t_0 + (j+1)\Delta t + \bar{T}]$. В момент времени $t = t_0 + (j+1)\Delta t$ игроки получают дополнительную информацию о динамике игры $g(t, x, u) = g_k(t, x, u)$ и выигрыше $h_i(t, x, u) = h_i^k(t, x, u)$ на временном интервале $(t_0 + k\Delta t, t_0 + (k+1)\Delta t + \bar{T}]$, где $k = j+1 + \frac{\bar{T}}{\Delta t}$. Эта новая информация о функциях $g(t, x, u)$ и $h_i(t, x, u)$ играет роль динамического обновления, а новая информация — это само обновление.

Для моделирования такой информационной структуры вводится следующее определение. Обозначим вектор $x_{j,0} = x(t_0 + j\Delta t)$.

Определение 3.1.1 Пусть $j = 0, \dots, l$. Усеченная подыгра $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ определена на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$. Уравнения движения и функция выигрыша усеченной подыгры на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ совпадают с уравнениями движения и функцией выигрыша исходной игры $\Gamma(x_0, T - t_0)$ на том же временном интервале.

Функция выигрыша игрока $i \in N$ в усеченной игре j имеет вид

$$K_i^j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; u) = \int_{t_0 + j\Delta t}^{t_0 + j\Delta t + \bar{T}} h_i(\tau, x(\tau), u(\tau)) e^{-r(\tau - t_0)} d\tau \quad (3.3)$$

при условии

$$\dot{x} = g(t, x, u), \quad x(t_0 + j\Delta t) = x_{j,0}, \quad (3.4)$$

где $x_{j,0}$ — начальное состояние (позиция) усеченной подыгры $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$.

3.1.2.1 Решение кооперативной усеченной подыгры

Рассмотрим усеченную кооперативную подыгру $\bar{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ с начальным условием $x(t_0 + j\Delta t) = x_{j,0}$. В кооперативной постановке игрокам необходимо максимизировать суммарный выигрыш

$$\sum_{i \in N} K_i^j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; u) = \sum_{i \in N} \int_{t_0 + j\Delta t}^{t_0 + j\Delta t + \bar{T}} h_i(\tau, x(\tau), u(\tau)) e^{-r(\tau - t_0)} d\tau \quad (3.5)$$

при условии

$$\dot{x} = g(t, x, u), \quad x(t_0 + j\Delta t) = x_{j,0}. \quad (3.6)$$

Это задача оптимального управления. Необходимые условия для ее решения и соответствующие управления могут быть определены с помощью уравнения Гамильтона–Якоби–Беллмана [182]. Обозначим максимальное значение суммарного выигрыша игроков (3.5) через $W^{(j\Delta t)}(t, x)$:

$$W^{(j\Delta t)}(t, x) = \max_{u \in U} \left\{ \sum_{i \in N} K_i^j(x, t, t_0 + j\Delta t + \bar{T}; u) \right\}, \quad (3.7)$$

где x, t — начальные позиция и время подыгры усеченной игры и $U = U_1 \times \dots \times$

U_n .

Теорема 3.1.1 *Предположим, что существует непрерывно дифференцируемая функция $W^{(j\Delta t)}(t, x) : [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}] \times R^m \rightarrow R$, удовлетворяющая следующей системе уравнений в частных производных:*

$$-W_t^{(j\Delta t)}(t, x) = \max_{u \in U} \left\{ \sum_{i=1}^n h_i(t, x, u) e^{-r(t-t_0)} + W_x^{(j\Delta t)}(t, x) g(t, x, u) \right\}, \quad (3.8)$$

где $\lim_{t \rightarrow -t_0 + j\Delta t + \bar{T}} W^{(j\Delta t)}(t, x) = 0$, а максимум в (3.8) достигается при $u_j^*(t, x)$. Тогда $u_j^*(t, x)$ является оптимальным в задаче управления, определяемой (3.5), (3.6).

Теорема 3.1.1 требует, чтобы функция $W^{(j\Delta t)}$ принадлежала пространству C^1 . Однако можно предположить непрерывность, рассматривая только т.н. вязкостные (viscosity) решения с использованием подхода Субботина [183; 184]. В модельном примере ниже решение $W^{(j\Delta t)}$ получается из пространства C^1 .

3.1.2.2 Условно-кооперативная траектория

Во время игры $\Gamma(x_0, T - t_0)$ игрокам доступна ограниченная информация о её структуре. Этой информации недостаточно, чтобы определить кооперативное поведение для игроков во всей игре $\Gamma(x_0, T - t_0)$. Вместо кооперативной траектории в игре $\Gamma(x_0, T - t_0)$ будем строить условно-кооперативную траекторию:

Определение 3.1.2 *Условно-кооперативная траектория $\{\hat{x}^*(t)\}_{t=t_0}^T$ — это комбинация кооперативных траекторий $x_j^*(t)$ в усеченных кооперативных подыграх $\bar{\Gamma}_j^c(x_{j-1}^*(t_0 + j\Delta t), t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, определенных на последовательных временных интервалах $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$:*

$$\{\hat{x}^*(t)\}_{t_0}^T = \begin{cases} x_0^*(t), & t \in [t_0, t_0 + \Delta t), \\ \vdots \\ x_j^*(t), & t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t), \\ \vdots \\ x_l^*(t), & t \in [t_0 + l\Delta t, t_0 + (l+1)\Delta t), \end{cases} \quad (3.9)$$

На временном интервале $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$ условно-кооперативная траектория совпадает с кооперативной траекторией $x_j^*(t)$ в усеченной кооперативной подыгре $\bar{\Gamma}_j^c(x_{j-1}^*(t_0 + j\Delta t), t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, где $x_{j-1}^*(t_0 + j\Delta t)$ — позиция игры в момент времени $t = t_0 + j\Delta t$ вдоль кооперативной траектории предыдущей подыгры $\bar{\Gamma}_j^c(x_{j-2}^*(t_0 + (j-1)\Delta t), t_0 + (j-1)\Delta t, t_0 + (j-1)\Delta t + \bar{T})$. В момент времени $t = t_0 + (j+1)\Delta t$ обновляется информация об игровой структуре в позиции $x_j^*(t_0 + (j+1)\Delta t)$. На временном интервале $(t_0 + (j+1)\Delta t, t_0 + (j+2)\Delta t]$ траектория $\hat{x}^*(t)$ совпадает с кооперативной траекторией $x_{j+1}^*(t)$ в усеченной кооперативной подыгре $\bar{\Gamma}_{j+1}^c(x_j^*(t_0 + (j+1)\Delta t), t_0 + (j+1)\Delta t, t_0 + (j+1)\Delta t + \bar{T})$, начинающейся в момент времени $t = t_0 + (j+1)\Delta t$ в позиции $x_j^*(t_0 + (j+1)\Delta t)$. При $j = 0$, $x_{j-1}^*(t_0 + j\Delta t) = x_0$.

3.1.2.3 Характеристическая функция в усеченной подыгре

Для каждой коалиции $S \subseteq N$ и усеченной подыгры с номером $j = 0, \dots, l$ определим значения характеристической функции так, как это сделано в [185] или [186]:

$$V_j(S; x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) = \begin{cases} \sum_{i \in N} K_i^j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; u_j^*), & S = N, \\ \tilde{V}_j(S, x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}), & S \subset N, \\ 0, & S = \emptyset, \end{cases} \quad (3.10)$$

где $\tilde{V}_j(S, x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ — суммарный выигрыш игроков из коалиции S в равновесии по Нэшу $u_j^{NE} = (u_1^{NE,j}, \dots, u_n^{NE,j})$ в игре со следующим набором игроков: коалиция S (действующая как один игрок) и игроки из множества $N \setminus S$, т.е. в игре с $|N \setminus S| + 1$ игроками.

Любой дележ $\xi_j(x_{j,0}^*, t_0 + j\Delta t)$ в кооперативной усеченной подыгре $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ должен удовлетворять следующей системе неравенств $\forall i \in N$:

$$\begin{aligned} \xi_i^j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) &\geq V_j(\{i\}, x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}), \quad i \in N, \\ \sum_{i \in N} \xi_i^j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) &= V_j(N, x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}). \end{aligned}$$

Обозначим множество всевозможных дележей для усеченной подыгры через $E_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$. Предположим, что для каждой усеченной подыгры определено непустое решение:

$$W_j(x_{j,0}^*, t, t_0 + j\Delta t + \bar{T}) \subset E_j(x_{j,0}^*, t, t_0 + j\Delta t + \bar{T}) \quad (3.11)$$

Это может быть С-ядро, НМ-решение, N-ядро или вектор Шепли.

3.1.3 Концепция кооперативного решения с динамическим обновлением информации

Логично предположить, что распределение суммарного выигрыша между игроками в игре $\Gamma(x_0, T - t_0)$ вдоль условно-кооперативной траектории $\{\hat{x}^*(t)\}_{t=t_0}^T$ определено, как комбинация дележей на временных интервалах $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, $j = 0, \dots, l$. Эта конструкция является новой концепцией решения. Будем называть ее результирующим решением.

Комбинация семейства множеств $W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ не позволяет получить решение в игре $\Gamma(x_0, T - t_0)$ непосредственно. Для каждого $j = 0, \dots, l$ решение в усеченной подыгре $\hat{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ определено для временного интервала $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$. Но информация об игре обновляется с шагом Δt , а использование такого решения на временном интервале $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$ не представляется возможным. Необходимая часть решения может быть получена с помощью процедуры распределения дележа для каждой усеченной подыгры. ПРД также обеспечивает свойство динамической устойчивости новой концепции решения и возможность определять решения внутри временного интервала $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$.

Определение 3.1.0.1 *Решение $W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ ($\xi_j(x_{j,0}^*, t_0 + j\Delta t)$) называется динамически устойчивым, если для каждого дележа $\xi_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) \in W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ существует соответствующая ПРД $\beta_j(t, x_j^*)$, такая что при $\forall t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ справедливо*

включение

$$\left\{ \int_t^{t_0+j\Delta t+\bar{T}} \beta_i^j(\tau, x_j^*) d\tau \right\} \in W(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \quad (3.12)$$

$$\left(\left\{ \int_t^{t_0+j\Delta t+\bar{T}} \beta_i^j(\tau, x_j^*) d\tau \right\} = \xi_j(x_{j,0}^*, t_0 + j\Delta t) \right).$$

Для того, чтобы построить решение в игре $\Gamma(x_0, T - t_0)$ необходимо определить ПРД для всех усеченных подыгр $\hat{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, $j = 0, \dots, l$. Обозначим семейство подыгр вдоль кооперативной траектории $x_j^*(t)$ через $\hat{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, где $t \in (t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ — начальный момент времени подыгры. Характеристическая функция вдоль $x_j^*(t)$ в семействе подыгр $\hat{\Gamma}_j^c(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ определена также, как и в (3.10). Обозначим через $E_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ множество дележей в подыгре $\hat{\Gamma}_j^c(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$.

Предположим, что в каждой усеченной подыгре $\hat{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ решение $W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) \neq \emptyset$ вдоль кооперативной траектории $x_j^*(t)$ выбрано. Также предположим, что для любой усеченной подыгры $\hat{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ в начальной позиции $x_{j,0}^*$ выбран дележ

$$\xi_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) \in W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$$

и соответствующая ПРД

$$\beta_j(t, x_j^*) = [\beta_1^j(t, x_j^*), \dots, \beta_n^j(t, x_j^*)], \quad t \in (t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}],$$

что гарантирует динамическую устойчивость выбранного дележа [187]:

$$\xi_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) = \int_{t_0+j\Delta t}^{t_0+j\Delta t+\bar{T}} \beta_j(t, x_j^*) e^{-r(\tau-t_0)} dt. \quad (3.13)$$

ПРД $\beta_j(t, x_j^*)$ может быть получена путем дифференцирования дележа $\xi_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, соответствующая теорема представлена в [188]:

Теорема 3.1.2 Если функция $\xi_j(x_{j,0}^*, t, t_0 + j\Delta t + \bar{T})$ является непрерывно

дифференцируемой по t и x_j^* , то

$$\beta_j(t, x_j^*) = -\xi_t^j(x_j^*, t, t_0 + j\Delta t + \bar{T}) - \xi_{x_j^*}^j(x_j^*, t, t_0 + j\Delta t + \bar{T})g\left(\tau, x_j^*(\tau), u_1^{*j}(\tau, x_j^*), \dots, u_n^{*j}(\tau, x_j^*)\right). \quad (3.14)$$

Новая концепция решения в игре $\Gamma(x_0, T - t_0)$ состоит из комбинации решений $W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ (соответствующих ПРД) в усеченных подыграх $\hat{\Gamma}_c^j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, $j = 0, \dots, l$. Пусть для каждого дележа $\xi_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) \in W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ существует ПРД $\beta_j(t, x_j^*)$. Определим результирующую ПРД для всей игры $\Gamma(x_0, T - t_0)$ следующим образом:

Определение 3.1.3 Результирующая ПРД $\hat{\beta}(t, \hat{x}^*)$ определяется для каждого набора $\xi_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) \in W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ с соответствующими ПРД $\beta_j(t, x_j^*)$ следующим образом:

$$\hat{\beta}(t, \hat{x}^*) = \begin{cases} \beta_0(t, x_0^*), & t \in [t_0, t_0\Delta t], \\ \vdots \\ \beta_j(t, x_j^*), & t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t], \\ \vdots \\ \beta_l(t, x_l^*), & t \in [t_0 + l\Delta t, t_0 + (l+1)\Delta t]. \end{cases} \quad (3.15)$$

С помощью результирующей ПРД $\hat{\beta}(t, \hat{x}^*)$ определим следующий вектор:

Определение 3.1.4 Результирующий дележ $\hat{\xi}(\hat{x}^*(t), T - t)$ — это вектор, определенный с помощью результирующей ПРД $\hat{\beta}(t, \hat{x}^*)$ следующим образом

при $t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$:

$$\begin{aligned} \hat{\xi}(\hat{x}^*(t), T - t) &= \int_t^T \hat{\beta}(\tau, \hat{x}^*(\tau)) e^{-r(\tau-t_0)} d\tau = \int_t^{t_0+j\Delta t} \beta_j(\tau, x_j^*(\tau)) e^{-r(\tau-t_0)} d\tau \\ &+ \sum_{m=j+1}^l \left[\int_{t_0+m\Delta t}^{t_0+(m+1)\Delta t} \beta_m(\tau, x_m^*(\tau)) e^{-r(\tau-t_0)} d\tau \right]; \quad (3.16) \end{aligned}$$

в частности,

$$\hat{\xi}(x_0, T - t_0) = \int_{t_0}^T \hat{\beta}(\tau, \hat{x}^*(\tau)) e^{-r(\tau-t_0)} d\tau.$$

Введем понятие результирующего решения в игре $\Gamma(x_0, T - t_0)$ с динамическим обновлением информации.

Определение 3.1.5 Результирующее решение $\hat{W}(\hat{x}^*(t), T - t)$ — это множество результирующих дележей $\hat{\xi}(\hat{x}^*(t), T - t)$, построенных с помощью (3.16) для всевозможных результирующих ПРД $\hat{\beta}(t, \hat{x}^*)$ (3.15).

В статье [77] показано, что с помощью результирующего дележа $\hat{\xi}(\hat{x}^*(t), T - t)$ и соответственно результирующего решения $\hat{W}(\hat{x}^*(t), T - t)$ можно разделить фактический суммарный выигрыш между игроками:

Следствие 3.1.0.1 Любой результирующий дележ $\hat{\xi}(x_0, T - t_0) \in \hat{W}(x_0, T - t_0)$ и соответствующая результирующая ПРД $\hat{\beta}(t, \hat{x}^*(t))$ распределяет текущий суммарный выигрыш игроков (3.5) вдоль условно-кооперативной траектории $\hat{x}^*(t)$ в игре с предписанной продолжительностью $\Gamma(x_0, T - t_0)$, т.е. $\forall t \in [t_0, T]$ выполняется условие

$$\sum_{i=1}^n \int_{t_0}^t \hat{\beta}_i(\tau, \hat{x}^*(\tau)) e^{-r(\tau-t_0)} d\tau = \sum_{i=1}^n \int_{t_0}^t h_i(\tau, \hat{x}^*(\tau), \hat{u}^*(\tau)) e^{-r(\tau-t_0)} d\tau. \quad (3.17)$$

Результирующее решение $\hat{W}(x_0, T - t_0)$ является динамически устойчивым по построению. Также доказано [76], что оно обладает и свойством сильной динамической устойчивости:

Определение 3.1.6 Решение $W(x_0, T-t_0)$ называется сильно Δt -динамически устойчивым, если для каждого $j = 0, \dots, l$ и каждого $\xi(x_0, T-t_0) \in W(x_0, T-t_0)$ соответствующая ПРД $\beta(t, x^*)$ удовлетворяет условию

$$\int_{t_0}^{t_0+j\Delta t} \beta(\tau, x^*(\tau)) e^{-r(\tau-t_0)} d\tau \oplus W(x_{j,0}^*, T-t_0+j\Delta t) \subset W(x_0, T-t_0) \quad (3.18)$$

в котором $a \oplus A = \{a + a' : a' \in A\}$.

Теорема 3.1.3 Произвольное результирующее решение $\hat{W}(x_0, T-t_0)$ является сильно Δt -динамически устойчивым в игре $\Gamma(x_0, T-t_0)$ с предписанной продолжительностью.

При произвольном результирующем решении определяется любое результирующее решение, которое строится с использованием решений $W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, $j = 0, \dots, l$, как произвольное подмножество множества дележей $E_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$. Решения в каждой усеченной подыгре могут различаться, т.е. в первой усеченной подыгре игроки могут выбрать S -ядро, во второй усеченной подыгре — вектор Шепли и т.д.

3.1.4 Построение характеристической функции в игре с динамическим обновлением информации

В качестве характеристической функции в дифференциальной игре $\Gamma(x_0, T-t_0)$ с динамическим обновлением информации будем использовать понятие результирующей характеристической функции.

Определение 3.1.7 Результирующей характеристической функцией $\bar{V}(S; \hat{x}^*(t), T-t)$ в игре $\Gamma(\hat{x}^*(t), T-t)$ с динамическим обновлением информации будем называть функцию, которая вычисляется с помощью значений характеристических функций $V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ в каждой текущей усеченной подыгре $\hat{\Gamma}_c^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ вдоль условно-кооперативной траектории $\hat{x}^*(t)$ для $j = 0, \dots, l$, $\forall t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$,

тогда:

$$\begin{aligned} \bar{V}(S; \hat{x}^*(t), T - t) = & \sum_{m=j+1}^l \left[V_m(S; x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \right. \\ & \left. - V_m(S; x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \right] + \\ & + \left[V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(S; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) \right], \end{aligned} \quad (3.19)$$

где $x_{j,0}^* = \hat{x}^*(t_0 + j\Delta t)$ и $x_{j,1}^* = \hat{x}^*(t_0 + (j+1)\Delta t)$.

В следующей теореме показано, что результирующий дележ $\hat{\xi}(x_0, T - t_0)$ может быть использован в качестве дележа в игре $\Gamma(x_0, T - t_0)$ с определенной результирующей характеристической функцией $\bar{V}(S; x_0, T - t_0)$.

Теорема 3.1.4 *Результирующий дележ $\hat{\xi}(x_0, T - t_0)$ является дележом в игре $\Gamma(x_0, T - t_0)$ с динамическим обновлением информации, если для $\forall t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, l$ выполняется следующее условие:*

$$\begin{aligned} & \xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \geq \\ & \geq \xi_i^j(x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) - V_j(\{i\}; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}). \end{aligned} \quad (3.20)$$

Доказательство Прежде всего, необходимо показать, что для $\forall t \in [t_0, T]$ выполняются следующие условия:

$$\sum_{i=1}^n \hat{\xi}_i(\hat{x}^*(t), T - t) = \bar{V}(N; \hat{x}^*(t), T - t), \quad (3.21)$$

$$\hat{\xi}_i(\hat{x}^*(t), t) \geq \bar{V}(\{i\}; \hat{x}^*(t), T - t). \quad (3.22)$$

В соответствии с определением $\hat{\xi}(\hat{x}^*(t), T - t)$ и $\bar{V}(S; \hat{x}^*(t), T - t)$ левую часть

(3.21) можно переписать следующим образом:

$$\begin{aligned}
\sum_{i=1}^n \hat{\xi}_i(\hat{x}^*(t), T-t) &= \sum_{i=1}^n \int_t^T \hat{\beta}(\tau, \hat{x}^*(\tau)) d\tau = \\
&= \sum_{i=1}^n \left[\sum_{m=j+1}^l \left[\int_{m\Delta t}^{(m+1)\Delta t} \beta_m(\tau, x_m^*(\tau)) d\tau \right] + \left[\int_t^{j\Delta t} \beta_j(\tau, x_m^*(\tau)) d\tau \right] \right] = \\
&= \sum_{i=1}^n \left[\sum_{m=j+1}^l [\xi_i^m(x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \right. \\
&\quad \left. - \xi_i^m(x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T})] + \right. \\
&\quad \left. + [\xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - \xi_i^j(x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T})] \right]. \quad (3.23)
\end{aligned}$$

Поскольку условие (3.24) в выражении (3.23) выполнено, формула (3.21) справедлива.

$$\sum_{i=1}^n \xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = V_j(N; x_j^*(t), t, t_0 + j\Delta t + \bar{T}), \quad j = 0, \dots, l, \quad (3.24)$$

Перейдем к доказательству (3.22). Подставим выражение для $\hat{\xi}_i(\hat{x}^*(t), T-t)$ и $\bar{V}(\{i\}; \hat{x}^*(t), T-t)$ в левую часть (3.22). В правую часть (3.22) подставим (3.19)

$\bar{V}(\{i\}; x_0, T - t_0)$:

$$\begin{aligned}
& \sum_{m=j+1}^l \left[\xi_i^m(x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \right. \\
& \quad \left. - \xi_i^m(x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \right] + \\
& + \left[\xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - \xi_i^j(x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) \right] \geq \\
& \geq \sum_{m=j+1}^l \left[V_m(\{i\}; x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \right. \\
& \quad \left. - V_m(\{i\}; x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \right] + \\
& + \left[V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(\{i\}; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) \right]. \tag{3.25}
\end{aligned}$$

(3.25) выполняется для $\forall t \in [t_0, T]$, если для $\forall m = 0, \dots, l$ справедливо

$$\begin{aligned}
& \xi_i^m(x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \xi_i^m(x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \geq \\
& V_m(\{i\}; x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - V_m(\{i\}; x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}), \tag{3.26}
\end{aligned}$$

а для $\forall t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $m = 0, \dots, l$, справедливо

$$\begin{aligned}
& \xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - \xi_i^j(x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) \geq \\
& V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(\{i\}; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}). \tag{3.27}
\end{aligned}$$

Видно, что выполнение условия (3.27) для $\forall t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $m = 0, \dots, l$ влечет выполнение условия (3.26). Перепишем условие (3.27) в виде

$$\begin{aligned}
& \xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \geq \\
& \geq \xi_i^j(x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) - V_j(\{i\}; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}). \tag{3.28}
\end{aligned}$$

Условие (3.28) означает, что в каждой усеченной подыгре изменение значений характеристической функции и дележа в зависимости от времени происходит равномерно относительно друг друга. Теорема доказана. \square

В этом разделе было введено понятие характеристической функции $\bar{V}(S; x_0, T - t_0)$ в игре $\Gamma(x_0, T - t_0)$ с динамическим обновлением информации. Было показано, что результирующий дележ $\hat{\xi}(x_0, T - t_0)$ действительно является дележом в классическом понимании при выполнении дополнительных условий. Стоит заметить, что условие (3.20) выполняется не всегда.

3.1.5 Связь решения в усеченных подыграх и результирующего решения

В этом разделе показано, что если игроки в каждой усеченной подыгре будут выбирать дележ $\xi_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \in E_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ на основе $V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, $j = 0, \dots, l$ по одинаковому правилу, то предложенный результирующий дележ $\hat{\xi}(\hat{x}^*(t), T - t)$ соответствует дележу, выбранному по такому же правилу на основе предложенной характеристической функции $\bar{V}(S; \hat{x}^*(t), T - t)$. Докажем это для нескольких принципов оптимальности.

Сначала покажем, что если в каждой усеченной подыгре $\hat{\Gamma}_j(x_j^*, t, t_0 + j\Delta t + \bar{T})$ игроки будут выбирать вектор Шепли $Sh_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ в качестве дележа, то соответствующий результирующий вектор $\hat{\xi}(\hat{x}^*(t), T - t)$ (3.16) будет совпадать с вектором Шепли $\hat{S}h(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, рассчитанным на основе результирующей характеристической функции $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19).

Теорема 3.1.5 Пусть в каждой усеченной подыгре $\hat{\Gamma}_j(x_j^*, t, t_0 + j\Delta t + \bar{T})$:

$$\xi_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = Sh_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}),$$

где $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$. Тогда результирующий вектор $\hat{\xi}(\hat{x}^*(t), T - t)$ будет совпадать с $\hat{S}h(\hat{x}^*(t), T - t)$:

$$\hat{\xi}(\hat{x}^*(t), T - t) = \hat{S}h(\hat{x}^*(t), T - t), \quad \forall t \in [t, T],$$

где $\hat{S}h(\hat{x}^*(t), T-t)$ — вектор Шепли, рассчитанный на основе результирующей характеристической функции $\bar{V}(S; \hat{x}^*(t), T-t)$ (3.19).

Доказательство В данном случае результирующий дележ $\hat{\xi}(\hat{x}^*(t), T-t)$ рассчитывается с помощью формул (3.15), (3.16) на основе векторов Шепли $Sh_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ в каждой усеченной подыгре $\hat{\Gamma}_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$:

$$\begin{aligned} \hat{\xi}(\hat{x}^*(t), T-t) = & \sum_{m=j+1}^l \left[Sh_m(x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \right. \\ & \left. - Sh_m(x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \right] + \\ & + [Sh_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - Sh_j(x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T})], \quad (3.29) \end{aligned}$$

где $Sh_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$ рассчитывается по формуле

$$\begin{aligned} Sh_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = & \sum_{\substack{S \subset N \\ i \in S}} \frac{(|N| - |S|)! (|S| - 1)!}{|N|!} \\ & \cdot \left(V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(S \setminus \{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \right). \quad (3.30) \end{aligned}$$

$\hat{S}h(\hat{x}^*(t), T-t)$ рассчитывается с помощью результирующей характеристической функции $\bar{V}(S; \hat{x}^*(t)$,

$T-t$). Подставим выражение для $\bar{V}(S; \hat{x}^*(t), T-t)$ в формулу для $\hat{S}h(\hat{x}^*(t), T-t)$. Положив $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, получаем:

$$\begin{aligned}
\hat{S}h(\hat{x}^*(t), T - t) &= \sum_{\substack{S \subset N \\ i \in S}} \frac{(|N| - |S|)! (|S| - 1)!}{|N|!} \\
&\cdot \left(\sum_{m=j+1}^l \left[\left[V_m(S; x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \right. \right. \right. \\
&\quad \left. \left. \left. - V_m(S; x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \right] - \right. \right. \\
&\quad \left. \left. \left. - \left[V_m(S \setminus \{i\}; x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \right. \right. \right. \\
&\quad \left. \left. \left. - V_m(S \setminus \{i\}; x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \right] \right] + \right. \\
&\quad \left. \left[\left[V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - \right. \right. \right. \\
&\quad \left. \left. \left. - V_j(S; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) \right] - \left[V_j(S \setminus \{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - \right. \right. \right. \\
&\quad \left. \left. \left. - V_j(S \setminus \{i\}; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) \right] \right] \right). \quad (3.31)
\end{aligned}$$

При подстановке (3.30) в (3.29), после выноса знака суммы, мы получим правую часть (3.31). Теорема доказана. \square

Тот же результат можно получить для пропорционального решения. Предположим, что характеристическая функция $V(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ дифференцируема вдоль кооперативной траектории $x_j^*(t)$. Определим пропорциональное решение, используя его ПРД $\beta_j^{Prop}(t)$, следующим образом:

$$\begin{aligned}
\beta_i^{Prop,j}(t) &= \frac{-\frac{d}{dt} V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T})}{\sum_{i \in N} -\frac{d}{dt} V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T})} \\
&\quad \left(-\frac{d}{dt} V_j(N; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \right), \quad i \in N. \quad (3.32)
\end{aligned}$$

Соответствующий дележ, полученный прямым интегрированием $\beta_j^{Prop}(t)$ с помощью формулы (3.13), обозначим через $Prop_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$.

Покажем, что если в каждой усеченной подыгре $\hat{\Gamma}_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ игроки будут выбирать пропорциональное решение $Prop_j(x_j^*(t), t, t_0 + j\Delta t +$

\bar{T}) (3.32), то результирующий дележ, определенный по формуле $\hat{\xi}(\hat{x}^*(t), T - t)$ (3.16), будет совпадать с пропорциональным решением $\hat{Pr}op(\hat{x}^*(t), T - t)$ (3.32), рассчитанным на основе характеристической функции $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19).

Теорема 3.1.6 Пусть в каждой усеченной подыгре $\hat{\Gamma}_j(x_j^*, t, t_0 + j\Delta t + \bar{T})$:

$$\xi_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = Prop_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}),$$

где $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$. Тогда результирующий дележ $\hat{\xi}(\hat{x}^*(t), T - t)$ будет совпадать с $\hat{Pr}op(\hat{x}^*(t), T - t)$ (3.32):

$$\hat{\xi}(\hat{x}^*(t), T - t) = \hat{Pr}op(\hat{x}^*(t), T - t), \quad \forall t \in [t, T],$$

где $\hat{Pr}op(\hat{x}^*(t), T - t)$ — пропорциональное решение, рассчитанное на основе характеристической функции $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19).

Доказательство В данном случае результирующий дележ $\hat{\xi}(\hat{x}^*(t), T - t)$ рассчитывается с помощью формулы (3.15) для ПРД на основе комбинации значений ПРД для пропорционального решения (3.32) в каждой усеченной подыгре на временном интервале $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$.

Покажем, что формула для $\hat{Pr}op(\hat{x}^*(t), T - t)$, а именно для его ПРД $\hat{\beta}^{Prop}(t)$ (3.32) с характеристической функцией $\bar{V}(S; \hat{x}^*(t), T - t)$, сводится к правой части (3.32). Подставим в (3.32) выражение для $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19). Рассмотрим отдельно одно из слагаемых. Пусть $t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$:

$$\begin{aligned} -\frac{d}{dt}(\bar{V}(\{i\}; \hat{x}^*(t), T - t)) &= -\frac{d}{dt} \left(\sum_{k=j+1}^l \left[V_k(\{i\}; x_{k,0}^*, t_0 + k\Delta t, t_0 + k\Delta t + \bar{T}) - \right. \right. \\ &\quad \left. \left. - V_k(\{i\}; x_{k,0}^*, t_0 + (k + 1)\Delta t, t_0 + k\Delta t + \bar{T}) \right] + \right. \\ &\quad \left. + \left[V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(\{i\}; x_{j,0}^*, t_0 + (j + 1)\Delta t, t_0 + j\Delta t + \bar{T}) \right] \right). \end{aligned} \quad (3.33)$$

Из (3.33) видно, что для $t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, $j = 0, \dots, l$, под знаком

производной находится только одно слагаемое, зависящее от t , поэтому

$$-\frac{d}{dt}\left(\bar{V}(\{i\}; \hat{x}^*(t), T-t)\right) = -\frac{d}{dt}\left(V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T})\right). \quad (3.34)$$

Подставим (3.34) и аналогичную формулу для $\bar{V}(N; \hat{x}^*(t), T-t)$ в (3.32). Нетрудно убедиться, что в результате мы получим выражение эквивалентное правой части (3.32). Теорема доказана. \square

Покажем теперь, что если в каждой усеченной подыгре $\hat{\Gamma}_j(x_j^*, t, t_0 + j\Delta t + \bar{T})$ игроки будут выбирать в качестве принципа оптимальности S -ядро $C_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, то результирующее решение, каждый элемент которого $\hat{\xi}(\hat{x}^*(t), T-t)$ рассчитан по формуле (4.147), будет являться S -ядром, рассчитанным на основе результирующей характеристической функции $\bar{V}(S; \hat{x}^*(t), T-t)$ (3.19).

Теорема 3.1.7 Пусть в каждой усеченной подыгре $\hat{\Gamma}_j(x_j^*, t, t_0 + j\Delta t + \bar{T})$:

$$W_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = C_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}),$$

где $\forall t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$. Тогда для всех $\xi_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \in C_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, при которых выполняется условие

$$\begin{aligned} & \sum_{i \in S} \xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \geq \\ & \geq \sum_{i \in S} \xi_i^j(x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) - V_j(S; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}), \end{aligned} \quad (3.35)$$

справедливо

$$\hat{\xi}(\hat{x}^*(t), T-t) \in \hat{C}(\hat{x}^*(t), T-t), \quad \forall t \in [t, T],$$

где $\hat{C}(\hat{x}^*(t), T-t)$ — S -ядро, рассчитанное на основе результирующей характеристической функции $\bar{V}(S; \hat{x}^*(t), T-t)$ (3.19).

Доказательство Для доказательства этой теоремы необходимо будет доказать два утверждения:

- 1 Если игроки в каждой усеченной подыгре будут выбирать дележ $\xi_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \in C_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ на основе $V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, $j = 0, \dots, l$, то предложенный результирующий дележ $\hat{\xi}(\hat{x}^*(t), T - t)$ будет принадлежать S -ядру $\hat{C}(\hat{x}^*(t), T - t)$, рассчитанному на основе предложенной характеристической функции $\bar{V}(S; \hat{x}^*(t), T - t)$.
- 2 S -ядро $\hat{C}(\hat{x}^*(t), T - t)$ не должно содержать дележей $\hat{\xi}(\hat{x}^*(t), T - t)$, для которых нельзя найти набора дележей в усеченных подыграх $\xi_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \in C_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$.

Докажем первое утверждение: если набор дележей $\xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, удовлетворяет системе неравенств

$$\sum_{i \in S} \xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \geq V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}), \quad S \subset N,$$

для любых $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$, $i = 1, \dots, n$, то результирующий дележ $\hat{\xi}(\hat{x}^*(t), T - t)$ удовлетворяет системе неравенств

$$\sum_{i \in S} \hat{\xi}_i(\hat{x}^*(t), T - t) \geq \bar{V}(S; \hat{x}^*(t), T - t), \quad \forall t \in [t_0, T], \quad S \subset N. \quad (3.36)$$

Подставим выражение для $\hat{\xi}(\hat{x}^*(t), T - t)$ в левую часть (3.36). В правую часть (3.36) подставим (3.19). Поступим также, как и в теореме 3.1.6. Покажем, для любого $S \subset N$, $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ выполнение (3.36) сводится к выполнению условия (3.35). Первое утверждение доказано.

Перейдем к доказательству второго утверждения, а именно, что во множестве наборов дележей $\xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, $i = 1, \dots, n$, $j = 0, \dots, l$, удовлетворяющих системе неравенств $\forall j = 0, \dots, l, S \subset N$:

$$\begin{aligned}
& \sum_{m=j+1}^l \left[\xi_i^m(x_m^*(t), t, t_0 + m\Delta t + \bar{T}) - \right. \\
& \quad \left. - \xi_i^m(x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \right] + \\
& + \left[\xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - \xi_i^j(x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) \right] \geq \\
& \geq \sum_{m=j+1}^l \left[V_m(S; x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \right. \\
& \quad \left. - V_m(S; x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \right] + \\
& + \left[V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(S; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) \right], \\
& \tag{3.37}
\end{aligned}$$

существует хотя бы один набор $\xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, $i = 1, \dots, n$, $j = 0, \dots, l$ удовлетворяющий:

$$\begin{aligned}
& \sum_{i \in S} \xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \geq \\
& \geq V_j(S; x_j^*, t, t_0 + j\Delta t + \bar{T}), \quad \forall j = 0, \dots, l, \quad S \subset N. \tag{3.38}
\end{aligned}$$

Докажем это от противного. Допустим, что для дележей, удовлетворяющих (3.35) и (3.38), не выполняется (3.37). Покажем, что для $\forall j = 0, \dots, l$ выполняется:

$$\begin{aligned}
& \sum_{i \in S} \xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \geq \\
& \geq \sum_{i \in S} \xi_i^j(x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) - V_j(S; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}), \\
& \tag{3.39}
\end{aligned}$$

тогда из (3.38) получается, что знак правой и левой части всегда положительный, а из (3.35) следует, что в этом случае (3.39) выполняется. Теорема доказана. \square

Те же результаты можно получить для ПРД-ядра, нового кооперативного

решения, представленного в [100]. Введем следующее обозначение:

$$U_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = -\frac{d}{dt}V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}), \quad (3.40)$$

где $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ и $S \subseteq N$.

Определим $B_j(t)$ как набор интегрируемых вектор-функций $\beta_j(t)$, удовлетворяющих системе неравенств

$$B_j(t, x_j^*) = \left\{ \beta_j(t) = (\beta_1^j(t), \dots, \beta_n^j(t)) : \right. \\ \left. \sum_{i \in S} \beta_i^j(t) \geq U_j(S, x_j^*(t), t, t_0 + j\Delta t + \bar{T}), \right. \\ \left. \sum_{i \in N} \beta_i^j = U_j(N, x_j^*(t), t, t_0 + j\Delta t + \bar{T}), \forall S \subset N \right\}. \quad (3.41)$$

Предположим, что $B_j(t, x_j^*) \neq \emptyset$, $\forall t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$. Тогда с помощью множества $B_j(t, x_j^*)$ можно определить следующее множество векторов.

Определение 3.1.8 Множество всех $\xi_j(x_j^*(t), t)$ для некоторых интегрируемых селекторов $\beta_j(t, x_j^*) \in B_j(t, x_j^*)$ будем называть ПРД-ядром и обозначать через $\bar{C}_j(x_j^*(t), t_0 + j\Delta t + \bar{T})$, где

$$\bar{C}_j(x_j^*(t), t) = \{ \xi_j(x_j^*(t), t), t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}] \} \quad (3.42)$$

и при $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$:

$$\xi_j(x_j^*(t), t) = \int_t^T \beta_j(\tau, x_j^*) d\tau. \quad (3.43)$$

В статье [100] доказано, что ПРД-ядро является сильно динамически устойчивым.

Покажем теперь, что если в каждой усеченной подыгре $\hat{\Gamma}_j(x_j^*, t, t_0 + j\Delta t + \bar{T})$ игроки будут выбирать в качестве принципа оптимальности ПРД-ядро $\bar{C}_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ (3.42), то результирующее решение,

каждый элемент которого определен по формуле (3.16), будет являться ПРД-ядром, рассчитанным на основе результирующей характеристической функции $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19).

Теорема 3.1.8 Пусть в каждой усеченной подыгре $\hat{\Gamma}_j(x_j^*, t, t_0 + j\Delta t + \bar{T})$:

$$W_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = \bar{C}_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \neq \emptyset,$$

где $\forall t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$. Тогда

$$\hat{W}(\hat{x}^*(t), T - t) = \hat{C}(\hat{x}^*(t), T - t), \quad \forall t \in [t, T],$$

где $\hat{C}(\hat{x}^*(t), T - t)$ – ПРД-ядро, рассчитанное на основе характеристической функции $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19).

Доказательство Результирующее решение $\hat{W}(\hat{x}^*(t), T - t)$ состоит из векторов $\hat{\xi}(\hat{x}^*(t), T - t)$, каждый из которых определяется с помощью ПРД набора дележей $\xi_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \in \bar{C}_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, $j = 0, \dots, l$ по формуле (3.15). По построению ПРД каждый дележ из ПРД-ядра удовлетворяет следующей системе неравенств (3.41):

$$\sum_{i \in S} \beta_i^j(t, \hat{x}^*(t)) \geq -\frac{d}{dt} \left(V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \right),$$

$$\sum_{i \in N} \beta_i^j(t, \hat{x}^*(t)) = -\frac{d}{dt} \left(V_j(N; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \right), \quad \forall S \subset N.$$

Таким образом, результирующее решение $\hat{W}(\hat{x}^*(t), T - t)$ определяется с помощью (3.41) для $t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, $j = 0, \dots, l$.

Выпишем выражение для $\hat{C}(\hat{x}^*(t), T - t)$ с результирующей характеристической функцией $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19). Покажем, что эта система сводится к (3.41). Рассмотрим отдельно одно из ограничений в (3.41) при $t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$ и подставим туда выражение для $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19) при $t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$:

$$\begin{aligned}
-\frac{d}{dt}\left(\bar{V}(S; \hat{x}^*(t), T-t)\right) &= -\frac{d}{dt}\left(\sum_{k=j+1}^l \left[V_k(S; x_{k,0}^*, t_0 + k\Delta t, t_0 + k\Delta t + \bar{T}) - \right. \right. \\
&\quad \left. \left. - V_k(S; x_{k,0}^*, t_0 + (k+1)\Delta t, t_0 + k\Delta t + \bar{T})\right] + \right. \\
&\quad \left. + \left[V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(S; x_{j,0}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T})\right]\right). \tag{3.44}
\end{aligned}$$

Из (3.44) видно, что для $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, l$, под знаком производной находится только одно слагаемое, зависящее от t . Поэтому

$$-\frac{d}{dt}\left(\bar{V}(S; \hat{x}^*(t), T-t)\right) = -\frac{d}{dt}\left(V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T})\right). \tag{3.45}$$

Подставим (3.45) в формулу для $\hat{C}(\hat{x}^*(t), T-t)$:

$$\begin{aligned}
\sum_{i \in S} \hat{\beta}_i(t, \hat{x}^*(t)) &\geq -\frac{d}{dt}\left(V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T})\right), \\
\sum_{i \in N} \hat{\beta}_i(t, \hat{x}^*(t)) &= -\frac{d}{dt}\left(V_j(N; \hat{x}^*(t), t, t_0 + j\Delta t + \bar{T})\right), \forall S \subset N.
\end{aligned}$$

Таким образом, ПРД-ядро $\hat{C}(\hat{x}^*(t), T-t)$, рассчитанное на основе результирующей характеристической функции $\bar{V}(S; \hat{x}^*(t), T-t)$, совпадает с результирующим решением $\hat{W}(\hat{x}^*(t), T-t)$, рассчитанным с помощью комбинации решений $\bar{C}_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ в усеченных подыграх. Теорема доказана. \square

Таким образом, в этом разделе показано, что если игроки в каждой усеченной подыгре $\hat{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ в качестве принципа оптимальности будут выбирать пропорциональное решение, вектор Шепли, дележ из S -ядра или дележ из ПРД-ядра, то результирующий дележ также будет являться пропорциональным решением, вектором Шепли, дележом из S -ядра или дележом из ПРД-ядра в игре $\Gamma(x_0, T-t_0)$ с динамическим обновлением информации.

Доказанные в этом разделе теоремы дают подход для непосредственного вычисления результирующего решения.

3.1.6 Кооперативная игра добычи ограниченного ресурса с динамическим обновлением информации

Рассмотрим игру добычи ограниченного ресурса на ограниченном временном интервале. Решение игры двух лиц в классическом виде представлено в [69]. Проблема динамической устойчивости изучена Дэвидом Янгом [188]. В этом примере представлена игра добычи ограниченного ресурса с динамическим обновлением информации для трех лиц. В качестве принципа оптимальности используется С-ядро. Характеристическая функция рассчитывается также, как это сделано в [185]. В последней части примера показано свойство сильной динамической устойчивости.

3.1.6.1 Исходная игра

Уравнения движения, описывающие изменение запаса ресурса $x(t) \in X \subset R$, имеют вид

$$\dot{x} = a\sqrt{x(t)} - bx(t) - \sum_{i=1}^3 u_i, \quad x(t_0) = x_0,$$

где u_i — уровень добычи игрока $i = \overline{1, 3}$. Определим функцию выигрыша игрока i следующим образом:

$$K_i(x_0, t_0; u) = \int_{t_0}^T h_i(\tau, x(\tau), u(\tau)) d\tau,$$

где

$$h_i(\tau, x(\tau), u(\tau)) = \sqrt{u_i(\tau)} - \frac{c_i}{\sqrt{x(\tau)}} u_i(\tau), \quad i = \overline{1, 3},$$

а c_i — константа, $c_i \neq c_k, \forall i \neq k = \overline{1, 3}$.

3.1.6.2 Усеченная подыгра

Исходная игра $\Gamma(x_0, T - t_0)$ определена на временном интервале $[t_0, T]$. Предположим, что для любого $t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, $j = 0, \dots, l$, игроки имеют усеченную информацию об игре. Она включает в себя информацию

об уравнениях движения и функциях выигрыша на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$. Смоделируем это с помощью усеченной подыгры $\hat{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$. Уравнения движения и начальные условия имеют вид

$$\dot{x} = a\sqrt{x(t)} - bx(t) - \sum_{i=1}^3 u_i, \quad x(t_0 + j\Delta t) = x_{j,0}. \quad (3.46)$$

Определим функцию выигрыша игрока i следующим образом:

$$K_i^j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; u) = \int_{t_0 + j\Delta t}^{t_0 + j\Delta t + \bar{T}} h_i(\tau, x(\tau), u(\tau)) d\tau.$$

Рассмотрим случай, когда игроки соглашаются на кооперацию в усеченной подыгре $\hat{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$. Тогда игроки будут действовать исходя из максимизации их суммарного выигрыша.

3.1.6.3 Кооперативная траектория

Максимальный суммарный выигрыш в каждой усеченной подыгре $\hat{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ имеет следующий вид [69]:

$$W^j(t, x) = A^j(t)\sqrt{x} + C^j(t), \quad (3.47)$$

где функции $A^j(t)$ и $C^j(t)$ удовлетворяют системе дифференциальных уравнений

$$\begin{aligned} \dot{A}^j(t) &= \frac{b}{2}A^j(t) - \sum_{i=1}^3 \left[\frac{1}{4 \left[c_i + \frac{A^j(t)}{2} \right]} \right], \\ \dot{C}^j(t) &= -\frac{a}{2}A^j(t), \\ A^j(t_0 + j\Delta t + \bar{T}) &= 0, \quad C^j(t_0 + j\Delta t + \bar{T}) = 0. \end{aligned}$$

Кооперативная траектория $x_j^*(t)$ в каждой усеченной подыгре может быть

вычислена на временном интервале следующим образом [69]:

$$x_j^*(t) = \varpi_j^2(t_0 + j\Delta t, t) \left[\sqrt{x_{j,0}^*} + \frac{1}{2}a \cdot \int_{t_0+j\Delta t}^t \varpi_j(t_0 + j\Delta t, \tau)^{-1} d\tau \right]^2,$$

где $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ и

$$\varpi_j(t_0 + j\Delta t, t) = \exp \int_{t_0+j\Delta t}^t - \left[\frac{1}{2}b + \sum_{i=1}^3 \left[\frac{1}{4 \left[c_i + \frac{A^j(\tau)}{2} \right]^2} \right] \right] d\tau.$$

Начальное положение для кооперативной траектории в каждой усеченной подыгре устанавливается из предыдущей усеченной подыгры: $x_{0,0}^* = x_0$ и $x_{j,0}^* = x_{j-1}^*(t_0 + j\Delta t)$ для $1 \leq j \leq l$. Определим условно-кооперативную траекторию $\hat{x}^*(t)$ в игре $\Gamma(x_0, T - t_0)$:

$$\hat{x}_j^*(t) = x_j^*(t), \quad t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t], \quad j = 0, \dots, l.$$

3.1.6.4 Характеристическая функция

Для того, чтобы распределить максимальный суммарный выигрыш между игроками в каждой усеченной подыгре, необходимо определить значения характеристической функции $V_j(S; x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ для каждой коалиции $S \subset N$. В соответствии с формулой (3.10), максимальный суммарный выигрыш игроков $W_j(t_0 + j\Delta t, x_{j,0})$ (3.47) соответствует значению характеристической функции $V_j(N; x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ коалиции $S = N$ в усеченной подыгре $\hat{\Gamma}_c^j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$:

$$V_j(N; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = W_j(t, x_j^*(t)), \quad (3.48)$$

где $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$. Далее, нам необходимо определить значения характеристической функции для следующих коалиций:

$$\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}.$$

Для каждой коалиции вида $\{i\}$, $i = \overline{1, 3}$ нам необходимо определить равновесие по Нэшу в усеченной подыгре $\hat{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \overline{T})$ и, как результат, $V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \overline{T})$.

3.1.6.5 Коалиции, состоящие из одного игрока

Равновесие по Нэшу в усеченной подыгре $\hat{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \overline{T})$ определяется следующими стратегиями игроков:

$$u_i^j(t, x) = \frac{x}{4[c_i + A_i^j(t)/2]^2}, \quad i = \overline{1, 3},$$

где функции $A_i^j(t)$ для $i = \overline{1, 3}$ находятся из системы дифференциальных уравнений

$$\begin{aligned} \dot{A}_i^j(t) &= A_i^j(t) \left[\frac{b}{2} + \sum_{k \neq i} \frac{1}{8(c_k + A_k^j(t)/2)^2} \right] - \frac{1}{4(c_i + A_i^j(t)/2)}, \\ \dot{C}_i^j(t) &= -\frac{a}{2} A_i^j(t), \\ A_i^j(t_0 + j\Delta t + \overline{T}) &= 0, \quad C_i^j(t_0 + j\Delta t + \overline{T}) = 0. \end{aligned}$$

Выигрыш игрока $i = \overline{1, 3}$ в ситуации равновесия по Нэшу определяется функцией

$$V_i^j(t, x) = A_i^j(t)\sqrt{x} + C_i^j(t), \quad i = \overline{1, 3}.$$

Таким образом, значение характеристической функции для коалиций, состоящих из одного игрока $S = \{i\}$, $i \in N$, вычисляется следующим образом:

$$V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \overline{T}) = V_i^j(t, x_j^*(t)), \quad (3.49)$$

где $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \overline{T}]$, $j = 0, \dots, l$.

3.1.6.6 Коалиции, состоящие из двух игроков

В соответствии с формулой (3.10), значение характеристической функции $V_j(S; x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \overline{T})$ для коалиций, состоящих из двух игроков $S =$

$\{1, 2\}$, $\{1, 3\}$, $\{2, 3\}$, определяется как наилучший ответ коалиции против стратегий, входящих в ситуацию равновесия по Нэшу $u_j^{NE} = (u_1^{NE,j}, u_2^{NE,j}, u_3^{NE,j})$ в усеченной подыгре $\hat{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, используемых игроками, не входящими в коалицию. В нашем случае это означает, что игроки из коалиции S действуют как один игрок и максимизируют свой суммарный выигрыш. Используя этот подход, определим равновесие между двумя игроками в следующем виде: комбинированный игрок (коалиция S) и игрок, не входящий в коалицию S (коалиция N/S).

Рассмотрим формулу для $V_j(S; x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ в случае, когда $S = \{1, 2\}$. Формулы для вычисления остальных коалиций можно получить по такому же принципу.

$$\begin{aligned} V_{\{1,2\}}^j(t, x) &= A_{\{1,2\}}^j(t)\sqrt{x} + C_{\{1,2\}}^j(t), \\ V_3^j(t, x) &= A_3^j(t)\sqrt{x} + C_3^j(t), \end{aligned}$$

где функции $A_{\{1,2\}}^j(t)$, $A_3^j(t)$, $C_{\{1,2\}}^j(t)$, и $C_3^j(t)$ удовлетворяют системе дифференциальных уравнений

$$\begin{aligned} \dot{A}_{\{1,2\}}^j(t) &= A_{\{1,2\}}^j(t) \left[\frac{b}{2} + \frac{1}{8(c_3 + A_3^j(t)/2)^2} \right] - \sum_{k \in S} \frac{1}{4(c_k + A_{\{1,2\}}^j(t)/2)}, \\ \dot{A}_3^j(t) &= A_3^j(t) \left[\frac{b}{2} + \sum_{k \in S} \frac{1}{8(c_k + A_{\{1,2\}}^j(t)/2)^2} \right] - \frac{1}{4(c_3 + A_3^j(t)/2)}, \\ \dot{C}_{\{1,2\}}^j(t) &= -\frac{a}{2}A_{\{1,2\}}^j(t), \quad \dot{C}_3^j(t) = -\frac{a}{2}A_3^j(t) \end{aligned}$$

соответственно с начальными условиями $A_{\{1,2\}}^j(t_0 + j\Delta t + \bar{T}) = 0$, $A_3^j(t_0 + j\Delta t + \bar{T}) = 0$, $C_{\{1,2\}}^j(t_0 + j\Delta t + \bar{T}) = 0$, и $C_3^j(t_0 + j\Delta t + \bar{T}) = 0$.

Таким образом, значение характеристической функции коалиции $S = \{1, 2\}$ вычисляется следующим образом:

$$V_j(\{1, 2\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = V_{\{1,2\}}^j(t, x_j^*(t)), \quad (3.50)$$

где $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$.

3.1.6.7 Концепция решения

Пусть игроки в каждой кооперативной усеченной подыгре $\hat{\Gamma}_c^j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ используют в качестве принципа оптимальности C -ядро $C_j(x_j^*(t), T - t)$. Это означает, что игроки в каждой усеченной подыгре выбирают дележ $\xi_j(x_j^*, t, t_0 + j\Delta t + \bar{T}) \in C_j(x_j^*(t), T - t)$ по следующему правилу:

$$\sum_{i \in S} \xi_i^j(x_j^*, t, t_0 + j\Delta t + \bar{T}) \geq V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}), \quad S \subset N,$$

для любого $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$. Тогда результирующий дележ $\hat{\xi}(\hat{x}^*(t), T - t)$ для любого набора дележей в усеченных подыграх $\xi_j(x_j^*, t, t_0 + j\Delta t + \bar{T}) \in C_j(x_j^*(t), T - t)$, $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$, может быть вычислен по формуле (3.16). Через $\hat{C}(\hat{x}^*(t), T - t)$ обозначим множество дележей $\hat{\xi}(\hat{x}^*(t), T - t)$, построенных с помощью (3.15) и (3.16).

На основе результатов, полученных в разделах 3.1.4 и 3.1.5, решение $\hat{C}(\hat{x}^*(t), T - t)$ можно построить по следующему правилу:

$$\sum_{i \in S} \hat{\xi}_i(\hat{x}^*(t), T - t) \geq \bar{V}(S; \hat{x}^*(t), T - t), \quad S \subset N, \quad (3.51)$$

где $\bar{V}(S; \hat{x}^*(t), T - t)$ вычисляется по формуле (3.19).

Далее на примере конкретных дележей из $\hat{C}(\hat{x}^*(t), T - t)$ покажем, что построенное решение является сильно Δt -динамически устойчивым в игре $\Gamma(x_0, T - t_0)$.

3.1.6.8 Численный пример

Рассмотрим игру добычи ограниченного ресурса, заданную на временном интервале длины $T - t_0 = 4$. Информация об игре известна на временном интервале с продолжительностью $\bar{T} = 2$ и обновляется каждый $\Delta t = 1$ временной интервал. Зафиксируем следующие параметры: для уравнений движения, $a = 5$ и $b = 0.3$; для функции выигрыша, $c_1 = 0.15$, $c_2 = 0.65$ и $c_3 = 0.45$; для начальных условий, $t_0 = 0$ и $x_0 = 250$.

На рисунке 3.1 показаны оптимальные стратегии для первого игрока в игре с динамическим обновлением информации (сплошная линия) и оптимальные стратегии в исходной игре [69] (пунктирная линия).

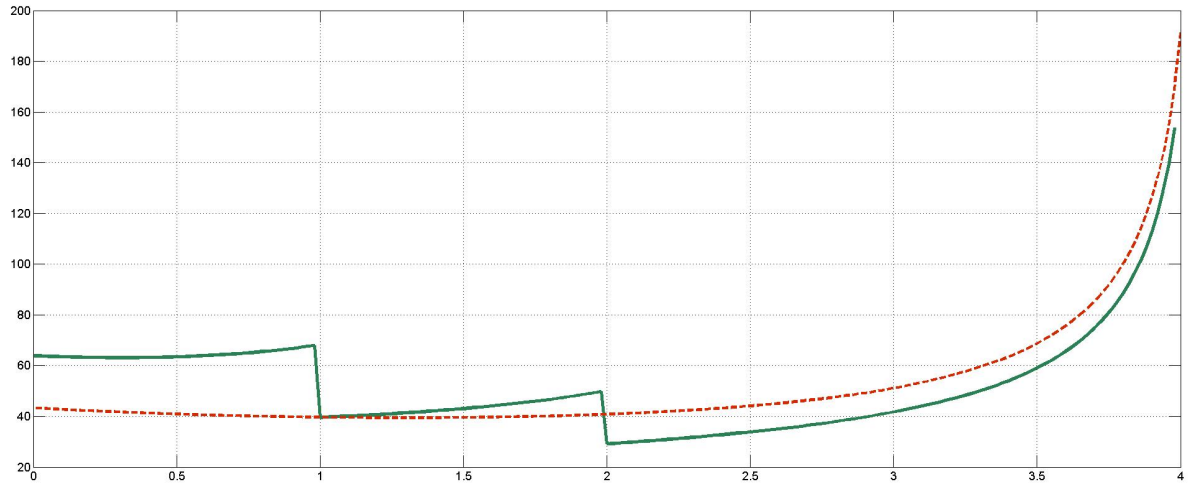


Рис. 3.1: Оптимальные стратегии игрока 1 в игре с динамическим обновлением информации (сплошная линия) и оптимальные стратегии в исходной игре [69] (пунктирная линия).

Условно-кооперативная траектория $\hat{x}^*(t)$ составлена из кооперативных траекторий в усеченных подыграх $\hat{\Gamma}_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ с уравнениями движения (3.46). На рисунке 3.2 представлено сравнение условно-кооперативной траектории $\hat{x}^*(t)$ (сплошная линия) в игре с динамическим обновлением информации и кооперативной траектории $x^*(t)$ (пунктирная линия) в исходной игре $\Gamma(x_0, T - t_0)$ [69]. Видно, что в случае ограниченной информации выработка ресурсов происходит быстрее, т.к. игроки ориентируются на урезанный временной интервал. Ось абсцисс на рисунке 3.2 соответствует времени t , а ось ординат — запасу ресурса x .

На основе значений характеристических функций

$$V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}), \quad t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t], \quad S \subset N, \quad i = 0, \dots, l,$$

вычисленных в (3.48), (3.49) и (3.50), получим выражение для результирующей характеристической функции $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19), $t \in [t_0, T]$. Далее с помощью (3.51) построим $\hat{C}(\hat{x}^*(t), T - t)$ в игре с динамическим обновлением информации $\Gamma(x_0, T - t_0)$ (см. рисунок 3.5).

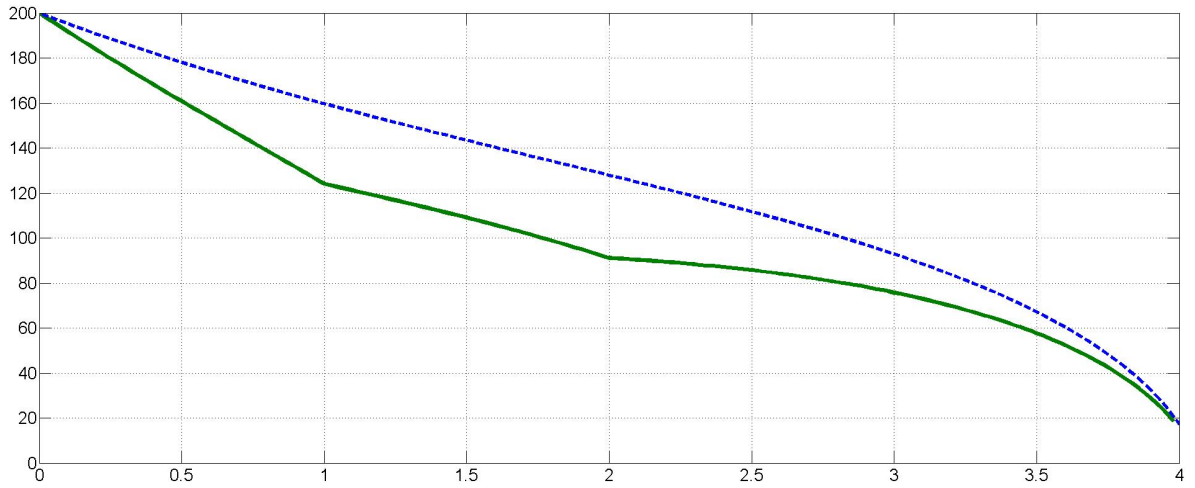


Рис. 3.2: Условно-кооперативная траектория $\hat{x}^*(t)$ (сплошная линия) в игре с динамическим обновлением информации и кооперативная траектория $x^*(t)$ (пунктирная линия) в исходной игре [69].

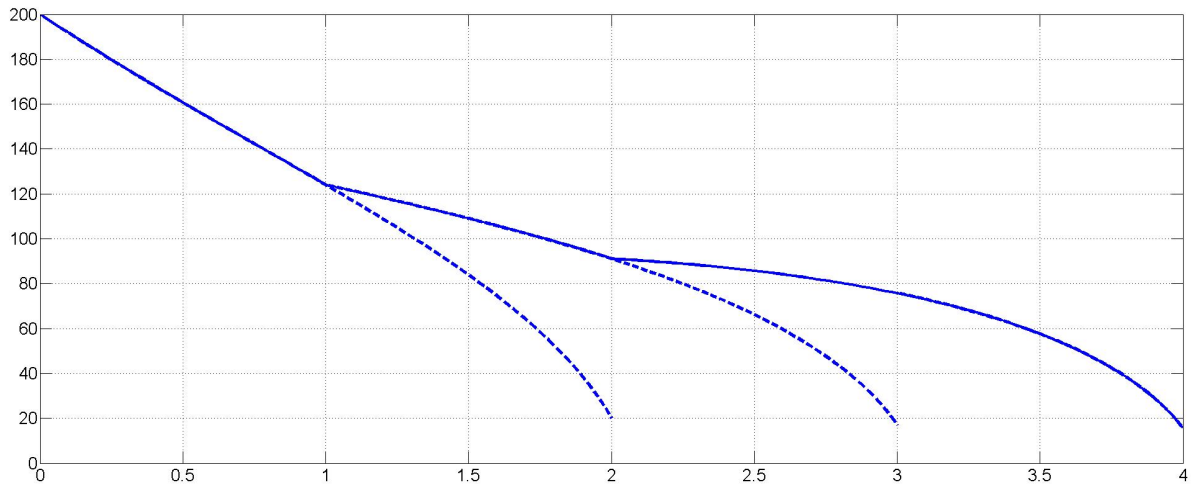


Рис. 3.3: Условно-кооперативная траектория $\hat{x}^*(t)$ (сплошная линия) в игре с динамическим обновлением информации и соответствующие кооперативные траектории в усеченных подыграх (пунктирные линии).

Продemonстрируем свойство сильной Δt динамической устойчивости решения $\hat{C}(\hat{x}^*(t), T - t)$. Предположим, что в начале игры $\Gamma(x_0, T - t_0)$ игроки договорились использовать пропорциональное решение $Prop(\hat{x}^*(t), T - t)$ (3.32) (далее покажем, что при заданных параметрах $Prop(\hat{x}^*(t), T - t) \in \hat{C}(\hat{x}^*(t), T - t)$). Теперь предположим, что в некоторый момент времени $t_{br} = t_0 + m\Delta t \in [t_0, T]$ игроки решили, что пропорциональное решение больше их не устраивает и выбрали другой дележ из $\hat{C}(\hat{x}^*(t_{br}), T - t_{br})$, например, вектор Шепли $\hat{Sh}(\hat{x}^*(t), T - t)$, $t \in [t_{br}, T]$ (3.30). Рассчитаем ПРД для пропорционального

решения и вектора Шепли по формулам (3.14).

Пусть $m = 2$, тогда ПРД для комбинированного решения (3.15) имеет следующий вид:

$$\hat{\beta}(t, \hat{x}^*) = \begin{cases} \hat{\beta}^{Prop}(t, \hat{x}^*), & t \in [t_0, t_{br}], \\ \hat{\beta}^{Sh}(t, \hat{x}^*), & t \in (t_{br}, T]. \end{cases} \quad (3.52)$$

На рисунке 3.4 изображены ПРД пропорционального решения, который выбрали игроки в начале игры $\hat{\beta}^{Prop}(t, \hat{x}^*)$ (сплошная линия) и ПРД $\hat{\beta}(t, \hat{x}^*)$ для комбинированного решения (3.52) (пунктирная линия).

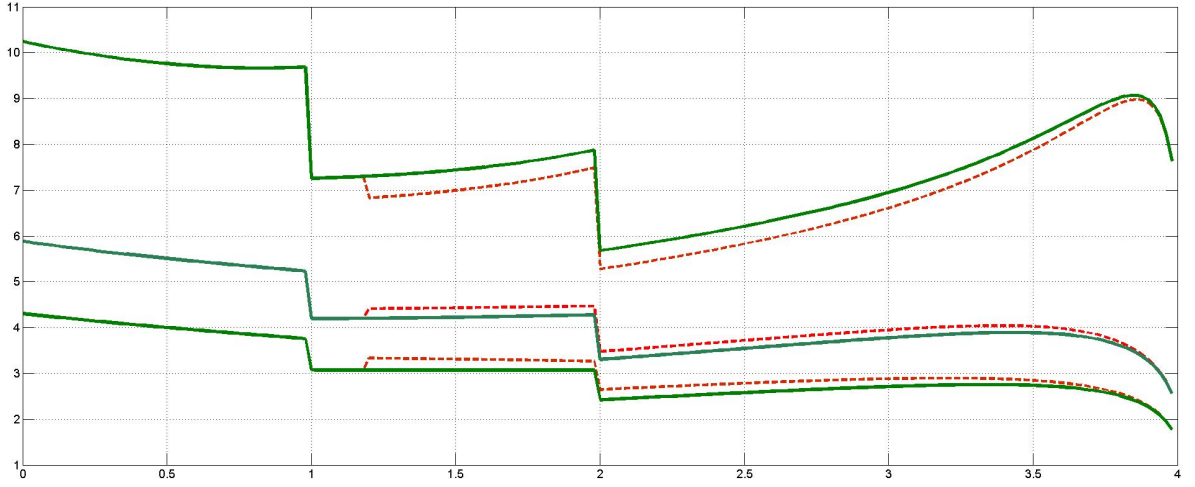


Рис. 3.4: ПРД $\hat{\beta}^{Prop}(t, \hat{x}^*)$ для пропорционального решения (сплошная линия), ПРД $\hat{\beta}(t, \hat{x}^*)$ для комбинированного решения (3.52) (пунктирная линия).

Проинтегрируем комбинированное решение $\hat{\beta}(t, \hat{x}^*)$ (3.52) по t и получим соответствующий дележ (3.16), обозначим его через $\hat{\xi}(\hat{x}^*(t), T - t)$. В соответствии с дележом $\hat{\xi}(\hat{x}^*(t), T - t)$ игроки разделят суммарный выигрыш в игре $\Gamma(x_0, T - t_0)$ с динамическим обновлением информации следующим образом:

$$\hat{\xi}(\hat{x}^*(t), T - t) = (12.3, 30.2, 16.8).$$

На рисунке 3.5 видно, что дележ, соответствующий комбинированному решению $\hat{\xi}(\hat{x}^*(t), T - t)$ (пунктирная линия) принадлежит $\hat{C}(\hat{x}^*(t), T - t)$ (выделенная область) для всех $t \in [t_0, T]$. Это показывает свойство сильной Δt -динамической устойчивости $\hat{C}(\hat{x}^*(t), T - t)$, т.к. дележ $\hat{\xi}(\hat{x}^*(t), T - t)$ был построен путем отклонения игроков от пропорционального решения $\hat{Pr}op(\hat{x}^*(t), T - t)$ (сплошная линия) в момент времени $t_{br} = t_0 + j\Delta t$ в пользу вектора Шепли $\hat{Sh}(\hat{x}^*(t_{br}), T -$

t_{br}).

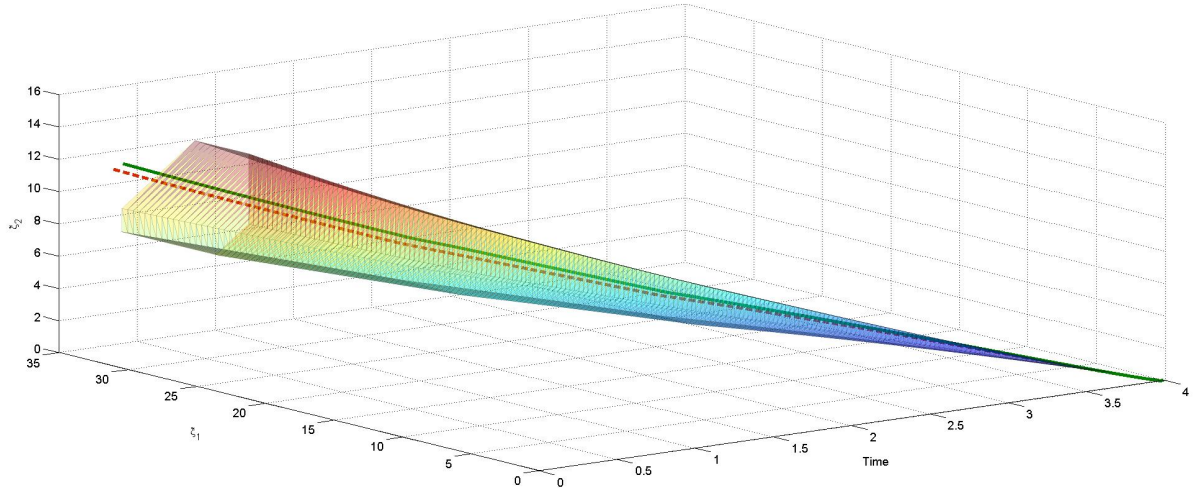


Рис. 3.5: Оси: ξ_1 , ξ_3 , t . ξ_2 можно вычислить, используя нормировочное условие.

3.2 Кооперативные дифференциальные игры с динамическим обновлением информации и стохастическим прогнозом

3.2.1 Исходная игра

Дифференциальная игра n лиц $\Gamma(x_0, 0, T)$ на временном интервале $[0, T]$ с начальным состоянием $x_0 \in R^n$ описывается уравнениями движения

$$\dot{x} = g(t, x, u), \quad x(0) = x_0, \quad (3.53)$$

где $x \in R^m$ и $u = (u_1 \dots u_n)$. Обозначим множество игроков через $N = \{1, 2, \dots, n\}$. Игрок i выбирает управление $u_i \in U_i \subset \text{Comp}R^k$ for $i = 1, \dots, n$.

Зададим функцию выигрыша игрока i следующим образом:

$$K_i(x_0, 0, T; u) := \int_0^T h_i(x(\tau), u(\tau)) d\tau + q_i(x(T)), \quad (3.54)$$

где $x(\tau)$ — траектория (решение) системы (3.53) на временном интервале $[0, T]$ при управлении u , а $q_i(x(T))$ — терминальный выигрыш.

Пусть функция $g(x, u)$ имеет вид

$$g(t, x, u) = \begin{cases} g_0(t, x, u), & t \in [0, \Delta t] \\ g_1(t, x, u), & t \in]\Delta t, 2\Delta t] \\ \dots \\ g_j(t, x, u), & t \in]j\Delta t, (j+1)\Delta t] \\ \dots \\ g_{l-1}(t, x, u), & t \in](l-1)\Delta t, l\Delta t]. \end{cases}$$

Необходимо выделить часть игровой динамики $g_j(t, x, u)$ на временном интервале $]j\Delta t, (j+1)\Delta t]$, $j = 0, \dots, l-1$, чтобы сформировать поведение игроков на этом временном интервале.

Предположим, что информация для игроков обновляется в фиксированные моменты времени $t = j\Delta t$, $j = 0, \dots, l-1$. В течение временного интервала $]j\Delta t, (j+1)\Delta t]$ игроки имеют определенную информацию о динамике игры, описываемой функцией $g_j(t, x, u)$. Однако им неизвестна игровая динамика на временном интервале $](j+1)\Delta t, T]$. Любой прогноз на этом интервале основан на текущей информации о процессе, описываемой в текущей модели с помощью функции динамики игры $g_j(t, x, u)$. Единственный вариант для игроков — действовать так, как будто структура игры не изменится, и использовать функцию $g_j(t, x, u)$ для прогноза динамики игры на временном интервале $[(j+1)\Delta t, T]$.

Для применения подхода с динамическим обновлением информации у игроков должна быть возможность использовать комбинированную усеченную информацию о структуре игры. Комбинированная усеченная информация в любой момент времени $t \in [j\Delta t, (j+1)\Delta t]$ включает уравнения движения на соответствующем временном интервале

$$\dot{x} = g_j(t, x, u), \quad (3.55)$$

а также прогноз движения на будущий временной интервал $[(j+1)\Delta t, T]$. Прогноз определяется стохастическим уравнением с той же детерминированной

частью:

$$dx := g_j(t, x, u)dt + \sigma(t, x)dz(t), \quad (3.56)$$

где $\sigma(t, x)$ — матрица размерности $m \times \theta$, а $z(t)$ — θ -мерный винеровский процесс. При обновлении информации о динамике игры в моменты времени $j\Delta t$, игроки пересчитывают свои решения, используя новую обновленную информацию.

Подобные задачи часто возникают в реальных жизненных ситуациях, поскольку структура конфликтного процесса на больших временных интервалах не определена и необходимо использовать прогноз.

3.2.2 Комбинированная усеченная подыгра

В течение первого временного интервала $[0, \Delta t]$ игроки имеют точную информацию об игре на интервале $[0, \Delta t]$ и прогноз на временном интервале $[\Delta t, T]$. В момент времени Δt информация об игре обновляется и на втором интервале $[\Delta t, 2\Delta t]$ игроки имеют точную информацию на $[\Delta t, 2\Delta t]$ и прогноз на $[2\Delta t, T]$. Для учета этого факта в модели, введем следующее определение.

Определение 3.2.1 Пусть $0 \leq j \leq l-1$. Комбинированная усеченная подыгра $\bar{\Gamma}_j(x, j\Delta, T)$ определена на временном интервале $[j\Delta t, T]$ следующим образом. На временном интервале $[j\Delta t, (j+1)\Delta t]$ уравнения движения, функция выигрыша в усеченной подыгре игре и исходной игре $\Gamma(x_0, 0, T)$ совпадают. Но на интервале $[(j+1)\Delta t, T]$ усеченная подыгра $\bar{\Gamma}_j(x, j\Delta, T)$ является стохастической дифференциальной игрой. Уравнения движения и начальное условие в комбинированной усеченной подыгре $\bar{\Gamma}_j(x, j\Delta t, T)$ имеют следующий вид:

$$dx := \begin{cases} g_j(t, x, u)dt, & t \in [j\Delta t, (j+1)\Delta t], \\ g_j(t, x, u)dt + \sigma(t, x)dz(t), & t \in [(j+1)\Delta t, T], \end{cases} \quad x(j\Delta t) = x. \quad (3.57)$$

Функция выигрыша игрока i равна

$$K_i^j(x, j\Delta t, T; u) := \int_{j\Delta t}^{(j+1)\Delta t} h_i(x(\tau), u(\tau))d\tau + E \left\{ \int_{(j+1)\Delta t}^T h_i(x(\tau), u(\tau))d\tau + q_i(x(T)) \right\}. \quad (3.58)$$

Очевидно, что комбинированная усеченная подыгра $\bar{\Gamma}_j(x, j\Delta t, T)$, $j = 0, \dots, l-1$, может быть определена как стохастическая дифференциальная игра с уравнениями движениями и начальными условиями вида

$$dx = g_j(t, x, u)dt + I(j, t) \cdot \sigma(t, x)dz(t), \quad x(j\Delta t) = x, \quad (3.59)$$

где

$$I(j, t) = \begin{cases} 0, & t \in [j\Delta t, (j+1)\Delta t], \\ 1, & t \in](j+1)\Delta t, T], \end{cases} \quad (3.60)$$

и функцией выигрыша игрока i , равной

$$K_i^j(x, j\Delta t, T; u) = E \left\{ \int_{j\Delta t}^T h_i(x(\tau), u(\tau))d\tau + q_i(x(T)) \right\}. \quad (3.61)$$

3.2.2.1 Решение комбинированной кооперативной усеченной подыгры

Рассмотрим комбинированную усеченную кооперативную подыгру $\bar{\Gamma}_j^c(x, j\Delta t, T)$ на временном интервале $[j\Delta t, T]$ с начальным условием $x(j\Delta t) = x$. В кооперативной постановке игрокам необходимо максимизировать суммарный выигрыш

$$\sum_{i \in N} K_i^j(x, j\Delta t, T; u^j) = E \left\{ \sum_{i \in N} \int_{j\Delta t}^T h_i(x(\tau), u(\tau))d\tau + q_i(x(T)) \right\}, \quad (3.62)$$

при условии

$$\begin{aligned} dx &= g_j(t, x, u)dt + I(j, t) \cdot \sigma(t, x)dz(t), \\ x(j\Delta t) &= x. \end{aligned} \quad (3.63)$$

Это задача стохастического управления. Достаточные условия для существования решения и оптимальной обратной связи даются следующим утверждением, которое является частным случаем [189, Theorem A.5].

Теорема 3.2.1 *Предположим, что существует дважды непрерывно дифференцируемая функция $W^{(j\Delta t)}(t, x) : [j\Delta t, T] \times R^m \rightarrow R$, удовлетворяющая следующей системе уравнений в частных производных:*

$$\begin{aligned} -W_t^{(j\Delta t)}(t, x) - \frac{1}{2} \sum_{h, \zeta=1}^m I(j, t) \sigma^{h, \cdot}(t, x(t)) \sigma^{\zeta, \cdot}(t, x(t))^T W_{x^h, x^\zeta}^{(j\Delta t)}(t, x) = \\ = \max_u \left\{ \sum_{i=1}^n h_i(t, x, u) + W_x^{(j\Delta t)}(t, x) g_j(t, x, u) \right\} \end{aligned} \quad (3.64)$$

с граничным условием

$$W^{(j\Delta t)}(T, x) = \sum_{i=1}^n q_i(x).$$

Предположим, что максимум в (3.64) достигается при $u = \psi^{*j}(t, x)$. Тогда позиционная стратегия $u(t, x) = \psi^{*j}(t, x)$ является оптимальной в игре $\bar{\Gamma}_j^c(x, j\Delta t, T)$, если замкнутая стохастическая система имеет решение.

3.2.2.2 Условно-кооперативная траектория

В соответствии с динамическим подходом, комбинированная усеченная информация о структуре игры доступна только игрокам в игре $\Gamma(x, 0, T)$. Этой информации недостаточно для построения оптимального кооперативного управления. Вместо оптимальной кооперативной траектории в игре $\Gamma(x, 0, T)$, строится

рекурсивная условно-оптимальная кооперативная траектория $\{\hat{x}^*(t)\}_{t=0}^T$ в виде

$$\{\hat{x}^*(t)\}_{t=0}^T := \begin{cases} x_0^*(t) & t \in [0, \Delta t] \\ \dots \\ x_j^*(t) & t \in]j\Delta t, (j+1)\Delta t] \\ \dots \\ x_{l-1}^*(t) & t \in](l-1)\Delta t, l\Delta t]. \end{cases} \quad (3.65)$$

На временном интервале $[0, \Delta t]$ траектория $x_0^*(t)$ является оптимальной в комбинированной усеченной кооперативной подыгре $\bar{\Gamma}_0^c(x_0, 0, T)$. В момент времени Δt обновляется информация о структуре игры, а положение игроков равно $x_0^*(\Delta t)$. На временном интервале $] \Delta t, 2\Delta t]$ функция $x_1^*(t)$ является частью оптимальной кооперативной траектории в комбинированной усеченной кооперативной подыгре $\bar{\Gamma}_1^c(x_0^*(\Delta t), \Delta t, T)$, начинающейся в момент времени Δt в позиции $x_0^*(\Delta t)$. В момент времени $j\Delta t$ обновляется информация о структуре игры и позиции системы: $x_{j-1}^*(j\Delta t)$. Функция $\hat{x}^*(t)$ на временном интервале $]j\Delta t, (j+1)\Delta t]$ определяется как часть оптимальной кооперативной траектории $x_j^*(t)$ в комбинированной усеченной кооперативной подыгре $\bar{\Gamma}_j^c(x_{j-1}^*(j\Delta t), j\Delta t, T)$, начинающейся в момент времени $j\Delta t$ в позиции $x_{j-1}^*(j\Delta t)$. Следовательно, условно-кооперативная траектория $\{\hat{x}^*(t)\}_{t=0}^T$ определяется как композиция кооперативных траекторий $x_j^*(t)$ комбинированных усеченных кооперативных подыгр $\bar{\Gamma}_j^c(x_{j-1}^*(j\Delta t), j\Delta t, T)$, определенных на последовательных временных интервалах $]j\Delta t, (j+1)\Delta t]$. Обозначим граничный вектор через $x_{j,0}^* := x_{j-1}^*(j\Delta t) = x_j^*(j\Delta t)$. Тогда комбинированная усеченная кооперативная подыгра — это $\bar{\Gamma}_j^c(x_{j,0}^*, j\Delta t, T)$.

3.2.2.3 Характеристическая функция для комбинированной усеченной подыгры

Для каждой коалиции $S \subset N$ и $j = 0, \dots, l - 1$, определим значения характеристической функции, как это было сделано в [190]:

$$V_j(S, x_{j,0}^*, j\Delta t, T) := \begin{cases} \sum_{i=1}^n K_i^j(x_{j,0}^*, j\Delta t, T; u_j^*) & S = N \\ \tilde{V}_j(S) & S \subset N \\ 0 & S = \emptyset, \end{cases} \quad (3.66)$$

где $\tilde{V}_j(S)$ — значение некооперативной комбинированной усеченной подыгры $\bar{\Gamma}_j(x_{j,0}^*, j\Delta t, T)$, определяемой как максимальная сумма выигрышей игроков из коалиции S относительно равновесных по Нэшу стратегий для проигравших игроков. Дележ $\xi^j(x_{j,0}^*, j\Delta t, T)$ для каждой комбинированной усеченной кооперативной подыгры $\bar{\Gamma}_j^c(x_{j,0}^*, j\Delta t, T)$ определяется как произвольная вектор-функция, удовлетворяющая условиям

$$\begin{aligned} \xi_i^j(x_{j,0}^*, j\Delta t, T) &\geq V_j(\{i\}, x_{j,0}^*, j\Delta t, T), \quad i \in N \\ \sum_{i \in N} \xi_i^j(x_{j,0}^*, j\Delta t, T) &= V_j(N, x_{j,0}^*, j\Delta t, T). \end{aligned} \quad (3.67)$$

Обозначим множество всех возможных дележей для комбинированной усеченной подыгры через $E_j(x_{j,0}^*, j\Delta t, T)$. Предположим, что непустое и оптимальное решение $W_j(x_{j,0}^*, j\Delta t, T) \subset E_j(x_{j,0}^*, j\Delta t, T)$ выбрано для каждой комбинированной усеченной подыгры $\bar{\Gamma}_j^c(x_{j,0}^*, j\Delta t, T)$. Это может быть С-ядро, НМ-решение, N-ядро или вектор Шепли. Легко предположить, что распределение суммарного выигрыша игроков в игре $\Gamma(x_0, 0, T)$ вдоль условно-кооперативной траектории $\{\hat{x}^*(t)\}_{t=0}^T$ организовано как композиция дележей для каждого временного интервала $[j\Delta t, (j+1)\Delta t]$, $j = 0, \dots, l - 1$, в соответствии со структурой игры $\Gamma(x_0, 0, T)$. Ниже это предположение будет формализовано как новая концепция решения. Семейство множеств $W_j(x_{j,0}^*, j\Delta t, T)$ не составляет непосредственно решение игры $\Gamma(x_0, 0, T)$. При любом $j = 0, \dots, l - 1$ оптимальное решение для усеченной подыгры $\bar{\Gamma}_j^c(x_{j,0}^*, j\Delta t, T)$ определено на временном интервале $[j\Delta t, T]$. Это конкретное решение имеет смысл только на интервале $[j\Delta t, (j+1)\Delta t]$ потому, что информация об игровой структуре обновляется

после каждого временного интервала Δt , и неразумно использовать решение, основанное на устаревшей информации. Необходимая информация может быть извлечена с помощью процедуры распределения дележа (ПРД) [191] для каждой усеченной подыгры. Кроме того, ПРД обеспечивает динамическую устойчивость нового решения и возможность определить совместное решение в любой момент времени на интервале $[j\Delta t, T]$.

3.2.3 Концепция комбинированного кооперативного решения

Для того, чтобы построить решение в исходной игре $\Gamma(x_0, 0, T)$ с динамическим обновлением информации необходимо определить ПРД для всех усеченных подыгр $\bar{\Gamma}_j^c(x_{j,0}^*, j\Delta t, T)$, $j = 0, \dots, l - 1$.

Пусть $0 \leq j \leq l - 1$. Обозначим семейство подыгр для $\bar{\Gamma}_j^c(x_{j,0}^*, j\Delta t, T)$ вдоль кооперативной траектории $x_j^*(t)$ через $\bar{\Gamma}_j^c(x_j^*(t), t, T)$, где $t \in [j\Delta t, T]$ — начальный момент времени комбинированной подыгры.

Характеристическая функция вдоль $x_j^*(t)$ в семействе подыгр $\bar{\Gamma}_j^c(x_j^*(t), t, T)$ на временном интервале $]t, T]$ определена также, как и в (4.142). Обозначим через $E_j(x_j^*, t, T)$ множество дележей в подыгре $\bar{\Gamma}_j^c(x_j^*(t), t, T)$ вдоль x_j^* .

Предположим, что для любой подыгры ранее определенной комбинированной усеченной кооперативной подыгры $\bar{\Gamma}_j^c(x_j^*(t), t, T)$ выбрано оптимальное решение $W_j(x_j^*, t, T) \neq \emptyset$ вдоль кооперативной траектории x_j^* .

Предположим, что для любой комбинированной усеченной подыгры $\bar{\Gamma}_j^c(x_{j,0}^*, j\Delta t, T)$ в начальных позициях $x_{j,0}^*$ игроки согласились выбрать дележи

$$\xi^j(x_{j,0}^*, j\Delta t, T) \in W_j(x_{j,0}^*, j\Delta t, T) \quad (3.68)$$

и соответствующие ПРД

$$B_j(t, x_j^*) = [B_1^j(t, x_j^*) \dots B_n^j(t, x_j^*)], \quad (3.69)$$

где $t \in]j\Delta t, T]$, что гарантирует динамическую устойчивость выбранного

дележа [192]:

$$\xi^j(x_{j,0}^*, j\Delta t, T) = E \left\{ \int_{j\Delta t}^T B_j(t, x_j^*) dt + q_i(x^*(T)) \right\}. \quad (3.70)$$

ПРД $B_j(t, x_j^*)$ может быть получена путем дифференцирования дележа $\xi_t^j(x_j^*, t, T)$ [189, Theorem 8.1].

Теорема 3.2.2 *Если дележи $\xi^j(x_j^*, t, T)$ являются дважды непрерывно дифференцируемыми по t и x_j^* , то*

$$\begin{aligned} B_j(\tau, x_j^*) = & - \left[\xi_t^j(x_j^*, t, T) \Big|_{t=\tau} \right] \\ & - \left[\xi_{x_j^*}^j(x_j^*, t, T) \Big|_{t=\tau} \right] g_j \left[\tau, x_j^*, \psi_1^{*j}(\tau, x) \dots \psi_n^{*j}(\tau, x) \right] - \\ & - \frac{1}{2} \sum_{h,\zeta=1}^m \Omega^{h,\zeta}(j, \tau, x_j^*) \left[\xi_{x^h, x^\zeta}^{(j\Delta t)}(x_j^*, t, T) \Big|_{t=\tau} \right], \end{aligned} \quad (3.71)$$

где $\Omega(j, t, x) = I(j, t)\sigma(t, x)\sigma(t, x)^T$ — ковариационная матрица, а $I(j, t)$ — индикаторная функция, определенная в (3.60).

Новая концепция решения в комбинированной дифференциальной игре $\Gamma(x_0, 0, T)$ состоит из рекурсивного слияния семейства динамически устойчивых решений $W_j(x_{j,0}^*, j\Delta t, T)$ комбинированных усеченных подыгр $\bar{\Gamma}_c(x_{j,0}^*(j\Delta t), j\Delta t, T)$, $j = 0, \dots, l-1$, полученное с помощью подхода с динамическим обновлением информации. Для любого дележа $\xi_j(x_{j,0}^*, j\Delta t, T) \in W_j(x_{j,0}^*, j\Delta t, T)$ существует ПРД $B_j(t, x_j^*)$. Определим комбинированное усеченное распределение

$$\hat{B}(t, \hat{x}^*) := B_j(t, x_j^*), \quad t \in [j\Delta t, (j+1)\Delta t], \quad 0 \leq j \leq l-1. \quad (3.72)$$

Функция $\hat{B}(t, \hat{x}^*)$ определяет комбинированный вектор

$$\hat{\xi}^j(x_{j,0}, j\Delta t, T) := \int_{j\Delta t}^T \hat{B}(\tau, \hat{x}^*(\tau)) d\tau + q(\hat{x}^*(T)) \quad (3.73)$$

$$= \sum_{m=j}^{l-1} \left[\int_{m\Delta t}^{(m+1)\Delta t} B_m(\tau, x_m^*(\tau)) d\tau \right] + q(x_l^*(T)) \quad (3.74)$$

при $j = 0, \dots, l-1$. Обозначим через $\hat{W}_j(x_{j,0}^*, j\Delta t, T)$ множество всех векторов $\hat{\xi}^j(x_{j,0}, j\Delta t, T)$, построенных по (3.72), (3.73).

В соответствие с новой концепцией, решение игры $\Gamma(x_0, 0, T)$ определим в виде $\hat{W} := (\hat{W}_j(x_{j,0}^*, j\Delta t, T))_{j=0}^{l-1}$.

Построение $\hat{B}(t, \hat{x}^*)$ не требует определения функции $B_j(t, x_j^*)$ на всем временном интервале, на котором определен соответствующий дележ $\xi_j(x_{j,0}^*, j\Delta t, T)$. Нам же необходимо, чтобы $B_j(t, x_j^*)$ была определена только на интервале $]j\Delta t, (j+1)\Delta t]$. Последнее слагаемое в (3.71) на этом интервале равняется нулю:

$$-\frac{1}{2} \sum_{h,\zeta=1}^m \Omega^{h,\zeta}(j, t, x_j^*) \left[\xi_{x^h, x^\zeta}^{(j\Delta t)}(t, x_j^*) \right] = 0, \quad t \in]j\Delta t, (j+1)\Delta t],$$

потому что $\Omega(j, t, x) = I(j, t)\sigma(t, x)\sigma(t, x)^T$, а индикаторная функция $I(j, t) = 0$ для $t \in [j\Delta t, (j+1)\Delta t]$. Таким образом, формула для расчета $\hat{B}(t, \hat{x}^*)$ является верной в случае, если дележ $\xi_j(x_j^*, t, T)$ только один раз непрерывно дифференцируем по t и x_j^* .

Легко видеть, что решение \hat{W} динамически устойчиво. Однако есть еще одно удивительное свойство \hat{W} .

Теорема 3.2.3 *Решение \hat{W} является сильно Δt -динамически устойчивым в игре $\Gamma(x_0, 0, T)$ с динамическим обновлением информации и стохастическим прогнозом.*

Доказательство Пусть $0 \leq j \leq l-1$ и дележ $\hat{\xi}^0(x_0, 0, T) \in \hat{W}_0(x_0, 0, T)$ порождает ПРД $\hat{B}(t, \hat{x}^*)$. Тогда для любого $0 \leq k < j$ существует $\xi^k(x_{k,0}^*, k\Delta t, T) \in$

$W_k(x_{k,0}^*, k\Delta t, T)$ с ПРД $B_k(t, x_k^*)$ таким, что

$$\hat{B}(t, \hat{x}^*) = B_k(t, x_k^*), \quad t \in [k\Delta t, (k+1)\Delta t[, \quad 0 \leq k \leq j-1.$$

Следовательно,

$$\int_0^{j\Delta t} \hat{B}(\tau, \hat{x}^*(\tau)) d\tau = \sum_{k=0}^{j-1} \int_{k\Delta t}^{(k+1)\Delta t} B_k(t, x_k^*(t)) dt.$$

Предположим, что $\xi'' \in W_j(x_{j,0}^*, j\Delta t, T)$. Тогда для любого $j \leq k \leq l-1$ существует $\xi^k(x_{k,0}^*, k\Delta t, T) \in W_k(x_{k,0}^*, k\Delta t, T)$ с ПРД $B_k(t, x_k^*)$ такой, что $\hat{B}(t, \hat{x}^*) = B_k(t, x_k^*)$ для $t \in [k\Delta t, (k+1)\Delta t[$ и

$$\xi'' = \sum_{m=j}^{l-1} \left[\int_{m\Delta t}^{(m+1)\Delta t} B_m(\tau, x_m^*(\tau)) d\tau \right] + q(x_l^*(T)).$$

Таким образом,

$$\int_0^{j\Delta t} \hat{B}(\tau, \hat{x}^*(\tau)) d\tau + \xi'' = \sum_{m=0}^{l-1} \left[\int_{m\Delta t}^{(m+1)\Delta t} B_m(\tau, x_m^*(\tau)) d\tau \right] + q(x_l^*(T)) \in \hat{W}_0(x_0, 0, T).$$

Теорема доказана. □

3.2.4 Кооперативная дифференциальная игра добычи ограниченного ресурса с динамическим обновлением информации и стохастическим прогнозом

Следующая игра добычи ресурса была рассмотрена Дэвидом Янгом и Штеффеном Йоргенсеном [193]. Проблема динамической устойчивости в этой игре была изучена Дэвидом Янгом [189]. В этом разделе подход с динамическим обновлением информации применяется к данной игре с новым элементом в решении — индикаторной функцией $I(j, t)$. Полученные подыгры образуют комбинированное решение.

Рассмотрим дифференциальную игру добычи ограниченного ресурса с двумя асимметричными игроками. Уравнения движения для запаса ресурса $x(t) \in X \subset R$ имеют следующий вид:

$$\begin{aligned} \dot{x} &:= a\sqrt{x(t)} - bx(t) - u_1 - u_2 \\ x(0) &:= x_0, \end{aligned} \quad (3.75)$$

где u_i — скорость добычи игрока $i = 1, 2$. Предположим, что динамика игры может изменяться, а параметры уравнений движения — переключаться на последовательных временных интервалах: $a = a_j$ и $b = b_j$ при $t \in [j\Delta t, (j+1)\Delta t)$, где $j = 0, \dots, l-1$ и $T = l\Delta t$.

Определим функцию выигрыша игрока i следующим образом:

$$K_i(x, 0, T; u_1, u_2) := \int_0^T h_i(x(\tau), u(\tau)) d\tau + q_i \sqrt{x(T)}, \quad i = 1, 2, \quad (3.76)$$

где

$$h_i(x(\tau), u(\tau)) := \sqrt{u_i(\tau)} - \frac{c_i}{\sqrt{x(\tau)}} u_i(\tau), \quad i = 1, 2, \quad (3.77)$$

а c_1 и c_2 — константы, $c_1 \neq c_2$.

Исходная игра $\Gamma(x_0, 0, T)$ определена на временном интервале $[0, T]$. Предположим, что для любого $t \in [j\Delta t, (j+1)\Delta t]$, $j = 0, \dots, l-1$, игроки имеют комбинированную усеченную информацию об игре. Она включает в себя информацию об уравнениях движения, функциях выигрыша на временном интервале $[j\Delta t, (j+1)\Delta t]$ и прогноз на временном интервале $[(j+1)\Delta t, T]$. Смоделируем это с помощью комбинированной усеченной подыгры $\bar{\Gamma}_j(x, j\Delta t, T)$. Уравнения движения и начальные условия имеют вид

$$\begin{aligned} dx &:= \left[a_j \sqrt{x(t)} - b_j x(t) - u_1 - u_2 \right] dt + I(j, t) \cdot \sigma x(t) dz(t), \\ x(j\Delta t) &:= x, \end{aligned} \quad (3.78)$$

где

$$I(j, t) = \begin{cases} 0, & t \in [j\Delta t, (j+1)\Delta t], \\ 1, & t \in [(j+1)\Delta t, T]. \end{cases} \quad (3.79)$$

Зададим функцию выигрыша игрока i в стохастической игре $\bar{\Gamma}_j(x, j\Delta t, T)$ следующим образом:

$$K_i^j(x, j\Delta t, T; u_1, u_2) := E \left\{ \int_{j\Delta t}^T h_i(x(\tau), u(\tau)) d\tau + q_i \sqrt{x(T)} \right\}. \quad (3.80)$$

Игра добычи ресурсов подробно изучена в [189; 193]. Комбинированная усеченная подыгра $\bar{\Gamma}_j(x, j\Delta t, T)$ имеет точку равновесия по Нэшу, определяемую позиционными стратегиями

$$u_i^j(t, x) := \frac{x}{4[c_i + A_i^j(t)/2]^2}, \quad i = 1, 2,$$

где функции $A_i^j(t)$ удовлетворяют уравнениям

$$\begin{aligned} \dot{A}_i^j(t) &:= A_i^j(t) \left[\frac{b_j}{2} + \frac{1}{8}\sigma^2 + \frac{1}{8(c_j + A_{3-i}^j(t)/2)^2} \right] - \frac{1}{4(c_i + A_i^j(t)/2)}, \\ \dot{C}_i^j(t) &:= -\frac{a_j}{2} A_i^j(t) \end{aligned}$$

при $i = 1, 2$ с граничными условиями $A_i^j(T) := q_i$ и $C_i^j(T) := 0$.

Согласно [193], функция цены игрока $i = 1, 2$ в ситуации равновесия по Нэшу равна

$$V_i^j(t, x) := A_i^j(t)\sqrt{x} + C_i^j(t), \quad i = 1, 2. \quad (3.81)$$

Рассмотрим случай, когда игроки соглашаются на кооперацию в комбинированной усеченной подыгре $\bar{\Gamma}_j^c(x, j\Delta t, T)$. Следуя принципу оптимальности, они максимизируют общий суммарный ожидаемый выигрыш и распределяют его между собой.

Максимальный ожидаемый суммарный выигрыш в каждой усеченной подыгре $\bar{\Gamma}_j^c(x, j\Delta t, T)$ можно найти по аналогии с [189; 193]:

$$W^j(t, x) := A^j(t)\sqrt{x} + C^j(t), \quad (3.82)$$

где функции $A^j(t)$ и $C^j(t)$ удовлетворяют системе дифференциальных уравне-

ний

$$\begin{aligned} \dot{A}^j(t) &:= \left[\frac{1}{8} I(j, t) \cdot \sigma^2 + \frac{b_j}{2} \right] A^j(t) - \frac{1}{4 \left[c_1 + \frac{A^j(t)}{2} \right]} - \frac{1}{4 \left[c_2 + \frac{A^j(t)}{2} \right]}, \\ \dot{C}^j(t) &:= -\frac{a_j}{2} A^j(t), \quad A^j(T) := \sum_{i=1}^2 q_i, \quad C^j(T) := 0. \end{aligned} \quad (3.83)$$

Оптимальная кооперативная траектория $x_j^*(t)$ стохастической игры $\bar{\Gamma}_j^c(x, j\Delta t, T)$ может быть представлена в явном виде [189] на полном временном интервале $[j\Delta t, T]$. Формула для начального отрезка траектории $x_j^*(t)$ на временном интервале $[j\Delta t, (j+1)\Delta t]$ проще, поскольку уравнения движения на этом интервале не являются стохастическими. Траектория с начальными условиями $x = x_{j,0}^*$ равна

$$x_j^*(t) = \varpi_j^2(j\Delta t, t) \left[\sqrt{x_{j,0}^*} + \frac{1}{2} a_j \int_{j\Delta t}^t \varpi_j(j\Delta t, \tau)^{-1} d\tau \right]^2, \quad t \in]j\Delta t, (j+1)\Delta t], \quad (3.84)$$

где

$$\varpi_j(j\Delta t, t) = \exp \int_{j\Delta t}^t - \left[\frac{1}{2} b_j + \frac{1}{8 \left[c_1 + \frac{A^j(\tau)}{2} \right]^2} + \frac{A^j(\tau)}{8 \left[c_2 + \frac{A^j(\tau)}{2} \right]^2} \right] d\tau. \quad (3.85)$$

Начальные условия определяются рекурсивно оптимальной траекторией предыдущей игры: $x_{0,0}^* := x_0$ и $x_{j,0}^* := x_{j-1}^*(j\Delta t)$ при $1 \leq j \leq l-1$. В соответствии с подходом с динамическим обновлением информации, условно-кооперативная траектория $\hat{x}^*(t)$ определяется как

$$\hat{x}_j^*(t) := x_j^*(t), \quad t \in [j\Delta t, (j+1)\Delta t], \quad 0 \leq j \leq l-1.$$

В качестве численного примера рассмотрим игру добычи ограниченного ресурса с $T = 4$ временными интервалами, в которой информация обновляется каждый $\Delta t = 1$ временной интервал. Зафиксируем следующие параметры

модели:

$$\begin{aligned} a_0 &:= 10, & a_1 &:= 9, & a_2 &:= 12, & a_3 &:= 8, \\ b_0 &:= 0.5, & b_1 &:= 0.8, & b_2 &:= 0.5, & b_3 &:= 1.6. \end{aligned} \quad (3.86)$$

Параметры для функции выигрыша положим равными $c_1 = 0.05$, $c_2 = 0.1$, $q_1 = 1.5$, $q_2 = 1$, а начальные условия — равными $x_0 = 1$.

Условно-кооперативная траектория $\hat{x}^*(t)$ составлена из кооперативных траекторий в усеченных подыграх $\bar{\Gamma}_j^c(x_{j,0}^*, j\Delta t, T)$ с уравнениями движения (3.78). Если параметр $\sigma = 0$, то прогноз является неслучайным и предполагается, что игроки считают, что информация об игре на временном интервале $[j\Delta t, T]$ не изменится. Если $\sigma \neq 0$, то прогноз становится стохастическим. Ниже сравниваются неслучайный прогноз, стохастический прогноз и решение без использования прогноза. Решение с использованием прогноза рассчитано для $\sigma = 2$.

Траектория без прогноза рассчитывается на каждом временном интервале $[j\Delta t, (j+1)\Delta t]$ как оптимальная траектория на этом интервале в игре с функцией выигрыша

$$\kappa^j(x, u) := \sum_{i=1}^2 \left[\int_{j\Delta t}^{(j+1)\Delta t} h_i(x(\tau), u(\tau)) d\tau + q_i \sqrt{x((j+1)\Delta t)} \right].$$

Начальными условиями на этом интервале является значение траектории на предыдущем временном интервале в граничной точке $j\Delta t$. На рисунке 3.6 представлено сравнение условно-кооперативной траектории $\hat{x}^*(t)$ без прогноза (красная линия), с неслучайным прогнозом (синяя линия) и со стохастическим прогнозом (зеленая линия). Уровень запасов ресурсов $\hat{x}^*(t)$ без прогноза растет быстрее всего. В случае неслучайного прогноза рост происходит медленно, а для стохастического прогноза еще медленнее. Но это не означает, что уменьшаются выигрыши.

Функции мгновенного выигрыша

$$h_{1,2}(\hat{x}^*(t), \hat{u}^*(t)) := \sum_{i=1}^2 h_i(\hat{x}^*(t), \hat{u}^*(t))$$

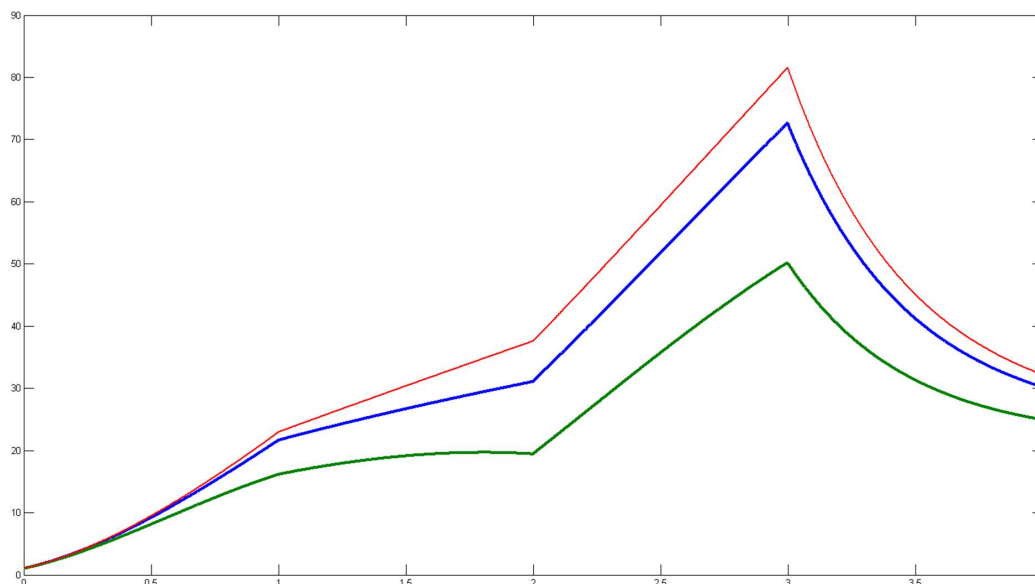


Рис. 3.6: Траектория уровня запаса ресурсов: без прогноза (красная линия), с неслучайным прогнозом (синяя линия), со стохастическим прогнозом (зеленая линия).

были рассчитаны по условно-кооперативным траекториям $\hat{x}^*(t)$ для систем с неслучайным прогнозом, со стохастическим прогнозом и без прогноза. Они показаны на рисунке 3.7.

На начальном отрезке временного интервала игры порядок функций выигрыша противоположен порядку запаса ресурса: наибольшее значение соответствует стохастическому прогнозу, а наименьшее значение получается без прогноза.

Функция кооперативного выигрыша

$$K^c(t, T) := \int_t^T h_{1,2}(\hat{x}^*(\tau), \hat{u}^*(\tau)) d\tau + (q_1 + q_2) \sqrt{\hat{x}^*(T)}, \quad t \in [0, T], \quad (3.87)$$

вдоль траекторий $\hat{x}^*(t)$ показаны на рисунке 3.8 сплошными линиями. Суммарный выигрыш составляет $K^c(0, T)$. Это значение равно 34,04 для стохастического прогноза, 34,0 для неслучайного прогноза и 33,54 без прогноза.

Пунктирными линиями на рисунке 3.8 показаны ожидаемые значения функций остаточного выигрыша в текущей усеченной подыгре $\bar{\Gamma}_j^c(x_{j,0}, j\Delta t, T)$. Эти подыгры изменяются в точках переключения $t = j = 1, 2, 3$, вызывая скачки значений ожидаемого кооперативного выигрыша.

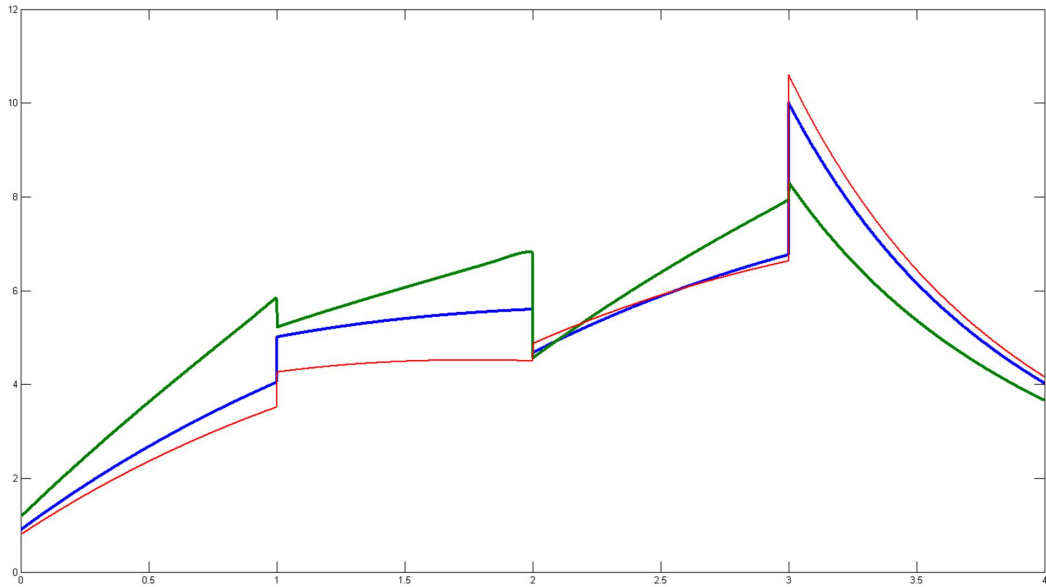


Рис. 3.7: Функции мгновенного кооперативного выигрыша $h_{1,2}(\hat{x}^*(t), \hat{u}^*(t))$: без прогноза (красная линия), с неслучайным прогнозом (синяя линия) и со стохастическим прогнозом (зеленая линия).

Анализ рисунков 3.7 и 3.8 (и подобных примеров численного моделирования игры добычи ресурса) позволяет сделать следующие выводы. Функция суммарного кооперативного выигрыша $K^c(0, T)$ почти одинакова для трех рассмотренных алгоритмов. Основное отличие состоит в распределении выигрыша во времени (рис. 3.7) и в росте запаса ресурса (рисунок 3.6). Неслучайный прогноз «осторожен» по сравнению с алгоритмом без прогноза: игроки предпочитают увеличивать выигрыш и отказываться от запаса ресурса в начале игры. Стохастический прогноз еще более «осторожен» по той же причине.

Функция дележа с непрерывным временем $\xi_i(t, T)$ должна удовлетворять принципу групповой и индивидуальной рациональности и должна быть динамически устойчивой [189; 193]. Достаточные условия получены с помощью точки равновесия по Нэшу, которая дает пропорциональный дележ для усеченной кооперативной подыгры $\bar{\Gamma}_j^c(x_{j,0}, j\Delta t, T)$, $0 \leq j \leq l - 1$,

$$\begin{aligned} \xi^{(j)i}(x, t, T) &:= \frac{V_i^j(t, x)}{\sum_{k=1}^2 V_k^j(t, x)} W^j(t, x) = \\ &= \frac{\left[A_i^j(t) \sqrt{x} + C_i^j(t) \right]}{\sum_{k=1}^2 \left[A_k^j(t) \sqrt{x} + C_k^j(t) \right]} \left[A^j(t) \sqrt{x} + C^j(t) \right]. \end{aligned}$$

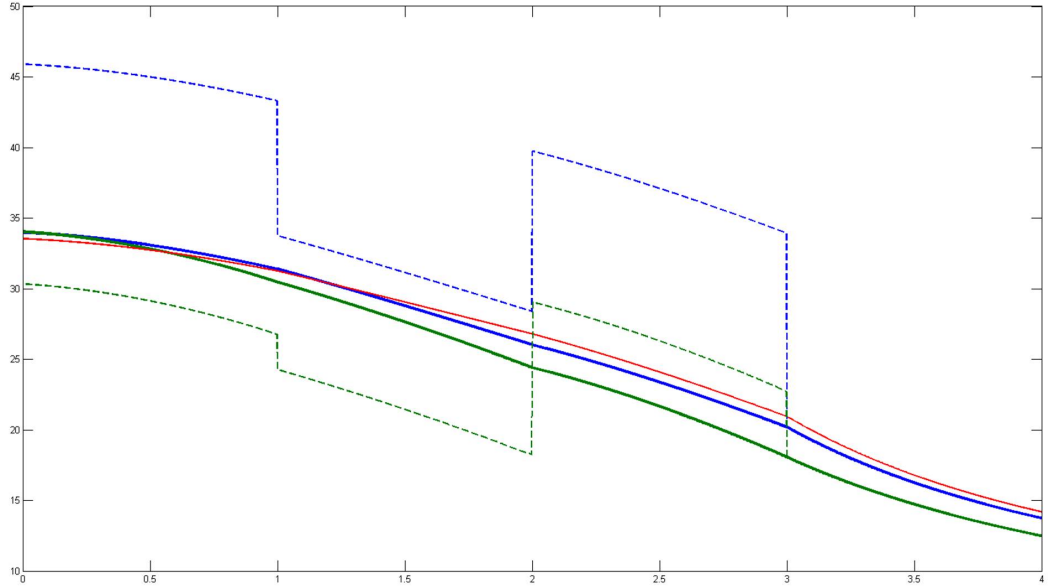


Рис. 3.8: Функции остаточного выигрыша $K^c(t, T)$: без прогноза (красная линия), с неслучайным прогнозом (синяя линия), со стохастическим прогнозом (зеленая линия). Ожидаемые значения функции остаточного выигрыша в текущей усеченной подыгре $\bar{\Gamma}_j^c(x_{j,0}, j\Delta t, T)$ (пунктирные линии).

Обозначим соответствующую ПРД через

$$B_j(t, x_j^*) := \left[B_1^j(t, x_j^*), B_2^j(t, x_j^*) \right]. \quad (3.88)$$

Общая формула для ПРД $B_i^j(t, x_j^*)$ в стохастической игре добычи ресурса была получена Янгом и Петросяном [189]:

$$\begin{aligned} B_i^j(\tau, x_j^*) = & - \left[\xi_t^{(j)i}(x_j^*, t, T)|_{t=\tau} \right] \\ & - \left[\xi_{x_j^*}^{(j)i}(x_j^*, t, T)|_{t=\tau} \right] \left[a_j \sqrt{x_j^*} - b_j x_j^* - \frac{x_j^*}{4[c_1 + A^j(t)]^2} - \frac{x_j^*}{4[c_2 + A^j(t)]^2} \right] - \\ & - I(j, t) \cdot \frac{1}{2} \sigma^2 (x_j^*)^2 \left[\xi_{x_j^*, x_j^*}^{(j)i}(t, x_j^*)|_{t=\tau} \right], \quad i = 1, 2. \end{aligned} \quad (3.89)$$

Комбинированное усеченное решение $\hat{B}_i(t, \hat{x}^*)$ составлено из $B_i^j(\tau, x_j^*)$ (3.72) в соответствии с подходом с динамическим обновлением информации, $i = 1, 2$. Последний член в (3.89) равен нулю на временном интервале $]j\Delta t, (j+1)\Delta t]$, поскольку $I(j, t) = 0$. Частные производные $\xi_t^{(j)i}$ и $\xi_x^{(j)i}$ представлены в явном виде с помощью функций \hat{x}^* , A^j , C^j , A_1^j , C_1^j , A_2^j и C_2^j , которые получены интегрированием дифференциальных уравнений.

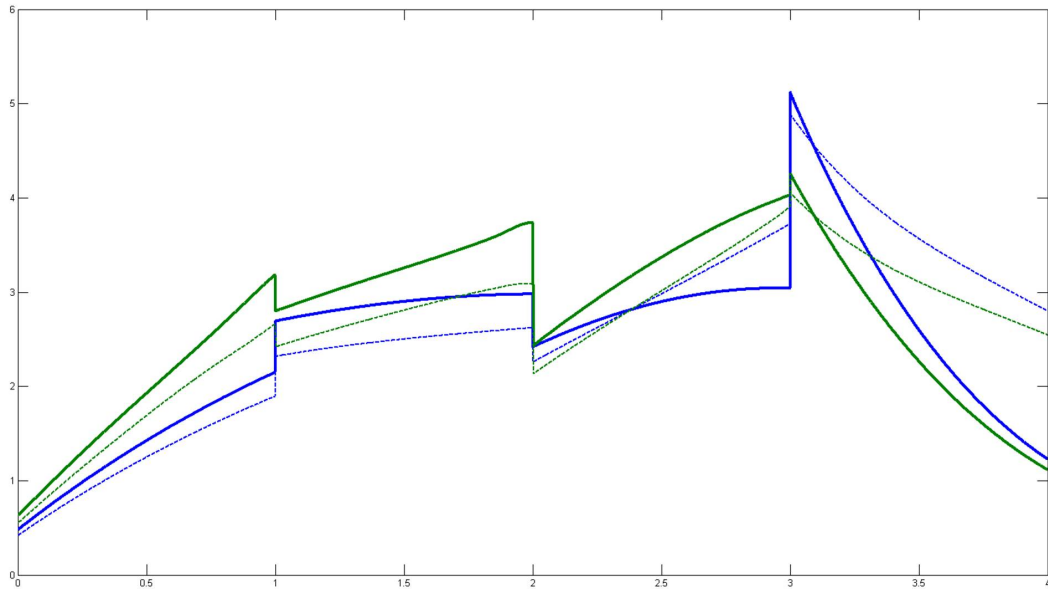


Рис. 3.9: ПРД для неслучайного прогноза (синяя линия) и для стохастического прогноза (зеленая линия). $\hat{B}_1(t, \hat{x}^*(t))$ — ПРД первого игрока (сплошная линия). $\hat{B}_2(t, \hat{x}^*(t))$ — ПРД второго игрока (пунктирная линия).

Результирующие ПРД показаны на рисунке 3.9 для решений с неслучайным и стохастическим прогнозами. Это — распределение функции мгновенного выигрыша h_{12} (рисунок 3.7) между игроками. Сумма сплошной и пунктирной линий на рисунке 3.9 совпадает с линией того же цвета на рисунке 3.7. За один раз каждый игрок может получить больше или меньше половины полной выплаты, и эта доля зависит как от момента времени, так и от алгоритма прогноза.

3.3 Кооперативные дифференциальные игры с динамическим обновлением информации и случайным горизонтом

3.3.1 Усеченная подыгра со случайной продолжительностью

Предположим, что информация об игре обновляется в фиксированные моменты времени $t = t_0 + j\Delta t$, $j = 0, \dots, l$, где $t_0 < \Delta t < T$ и $l = \frac{T-t_0}{\Delta t} - 1$. На протяжении временного интервала $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$ игроки обладают полной информацией о динамике игры (функции $g(t, x, u)$) и выигрышах

(функции $h_i(x(t), u(t))$). Однако, игроки точно не знают длительность этого интервала, т.к. \bar{T}_j является случайной величиной, но ее распределение известно. Реализацию случайной величины \bar{T}_j обозначим через \bar{t}_j , $j = 0, \dots, l$.

Как упомянуто выше, в течение временного интервала $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$ игроки имеют полную информацию о динамике игры и функции выигрыша на временном интервале $[t_0 + j\Delta t, \bar{T}_j]$, где \bar{T}_j — случайная величина, принимающая значения из интервала $[\max(t_0 + (j+1)\Delta t, \bar{t}_{j-1}), T]$, а \bar{t}_{j-1} — реализация случайной величины \bar{T}_{j-1} (\bar{T}_{j-1} реализуется в момент времени $t = t_0 + j\Delta t$). В момент времени $t = t_0 + (j+1)\Delta t$ информация об игре обновляется и реализуется случайная величина \bar{T}_j , т.е. \bar{t}_j становится известной игрокам. На следующем временном интервале $(t_0 + (j+1)\Delta t, t_0 + (j+2)\Delta t]$ игроки имеют полную информацию о структуре игры на временном интервале $(t_0 + (j+1)\Delta t, \bar{T}_{j+1}]$, где \bar{T}_{j+1} — случайная величина, принимающая значения из интервала $[\max(t_0 + (j+2)\Delta t, \bar{t}_j), T]$. Для $j = 0$ положим $\bar{t}_{j-1} = 0$.

Возможно, остается неясным вопрос, почему случайная величина \bar{T}_j реализуется в моменты времени $t = t_0 + (j+1)\Delta t$, но ее значение \bar{t}_j превышает $t = t_0 + (j+1)\Delta t$. Одна из интерпретаций заключается в следующем. Предположим, что в моменты времени $t = t_0 + j\Delta t$ игроки получают информацию об игре, но для точной оценки информационного горизонта (случайной величины \bar{T}_j) им требуется время Δt (для расчетов и т. д.). В момент времени $t = t_0 + (j+1)\Delta t$ выполняются вычисления и игрокам становится известно определенное значение информационного горизонта \bar{T}_j , т.е. \bar{t}_j . Другая интерпретация: в течение текущего Δt -временного интервала $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$ игроки получают дополнительную информацию, помогающую им оценить временной интервал, на котором информация о процессе является достоверной (т.е. определить значение информационного горизонта \bar{t}_j). В момент времени $t = t_0 + (j+1)\Delta t$, после выполнения оценки, игроки получают новую информацию о структуре игры со случайным информационным горизонтом \bar{T}_{j+1} , и та же процедура продолжается.

Определение 3.3.1 Пусть $j = 0, \dots, l$. Случайная усеченная подыгра $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t)$ определена на временном интервале $[t_0 + j\Delta t, \bar{T}_j]$, где \bar{T}_j — случайная величина, принимающая значения из $[\max(t_0 + (j+1)\Delta t, \bar{T}_{j-1}), T]$, а \bar{T}_{j-1} — реализация случайного информационного горизонта в предыдущей усеченной

подыгре $\bar{\Gamma}_{j-1}(x_{j-1,0}, t_0 + (j-1)\Delta t)$. Величина \bar{T}_{j-1} реализуется в момент времени $t = t_0 + j\Delta t$. Уравнения движения и функция выигрыша в случайной усеченной подыгре и исходной игре $\Gamma(x_0, T - t_0)$ на временном интервале $[t_0 + j\Delta t, \bar{T}_j]$ совпадают. Уравнения движения и начальные условия в усеченной подыгре $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t)$ имеют вид

$$\dot{x} = g(t, x, u), \quad x(t_0 + j\Delta t) = x_{j,0}. \quad (3.90)$$

Функция выигрыша игрока i в случайной усеченной подыгре j имеет следующий вид:

$$K_i^j(x_{j,0}, t_0 + j\Delta t; u) = \int_{t_0 + j\Delta t}^T \int_{t_0 + j\Delta t}^t h_i(x(\tau), u(\tau)) d\tau dF_j(t), \quad (3.91)$$

где $F_j(t)$ — функция распределения случайной величины \bar{T}_j :

$$\int_{t_0 + j\Delta t}^T dF_j(t) = \int_{\max(t_0 + (j+1)\Delta t, \bar{t}_{j-1})}^T dF_j(t) = 1. \quad (3.92)$$

По определению $F_j(t)$ является функцией условного распределения, т.е. $F_j(t) = F_j(t | \bar{T}_{j-1} = \bar{t}_{j-1})$. Далее для функции $F_j(t | \bar{T}_{j-1} = \bar{t}_{j-1})$ будет использоваться компактное обозначение $F_j(t)$.

Предположим, что реализация случайной величины \bar{T}_{j-1} в игре $\bar{\Gamma}_{j-1}(x_{j-1,0}, t_0 + (j-1)\Delta t)$ превышает время $t = t_0 + (j+1)\Delta t$:

$$\bar{t}_{j-1} > t_0 + (j+1)\Delta t. \quad (3.93)$$

Тогда значение случайного информационного горизонта \bar{T}_j должно превышать реализацию \bar{T}_{j-1} , т.к. информация об игре уже известна на временном интервале $[t_0 + j\Delta t, \bar{T}_{j-1}]$. Поэтому в формуле (3.92) вероятность того, что случайная

величина \bar{T}_j примет значение из интервала $[t_0 + j\Delta t, \bar{T}_{j-1}]$, равна нулю:

$$\int_{t_0 + j\Delta t}^{\max(t_0 + (j+1)\Delta t, \bar{T}_{j-1})} dF_j(t) = 0. \quad (3.94)$$

В большинстве статей, посвященных изучению кооперативных дифференциальных игр со случайной продолжительностью [194; 195; 196; 197; 198], функция распределения случайной величины \bar{T}_j определяется на бесконечном временном интервале. В данной постановке \bar{T}_j принимает значения на конечном временном интервале, т.к. исходная игра определена на конечном интервале $[t_0, T]$.

В [199] порядок интегрирования в двойном интеграле (3.91) изменен согласно теореме Тонелли:

$$K_i^j(x_{j,0}, t_0 + j\Delta t; u) = \int_{t_0 + j\Delta t}^T (1 - F_j(\tau)) h_i(x(\tau), u(\tau)) d\tau. \quad (3.95)$$

3.3.1.1 Решение кооперативной случайной усеченной подыгры

Рассмотрим усеченную случайную кооперативную подыгру $\bar{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t)$ на временном интервале $[t_0 + j\Delta t, \bar{T}_j]$ с начальным условием $x(t_0 + j\Delta t) = x_{j,0}$, где \bar{T}_j — случайная величина с функцией распределения (3.92). При классическом подходе на первом шаге кооперативных дифференциальных игр определяются кооперативные стратегии и соответствующая кооперативная траектория. На втором шаге определяется правило распределения кооперативного выигрыша между игроками вдоль кооперативной траектории. Для этого определяются характеристическая функция и соответствующее кооперативное решение. В этой игре необходимо максимизировать суммарный выигрыш игроков

$$\sum_{i \in N} K_i^j(x_{j,0}, t_0 + j\Delta t; u) = \sum_{i \in N} \int_{t_0 + j\Delta t}^T \int_{t_0 + j\Delta t}^t h_i(x(\tau), u(\tau)) d\tau dF_j(t) \quad (3.96)$$

при условии

$$\dot{x} = g(t, x, u), \quad x(t_0 + j\Delta t) = x_{j,0}. \quad (3.97)$$

Это задача оптимального управления. Достаточные условия для ее решения и соответствующие управления с обратной дают Теорема 3.3.1, впервые представленной в [197]. Обозначим через $W^{(j\Delta t)}(t, x)$ максимальное значение суммарного выигрыша игроков (3.96):

$$W^{(j\Delta t)}(t, x) = \max_{u \in U} \left\{ \sum_{i \in N} K_i^j(x, t; u) \right\}, \quad (3.98)$$

где x и t — соответственно начальная позиция и время подыгры усеченной игры, а $U = U_1 \times \dots \times U_n$.

Теорема 3.3.1 *Предположим, что существует непрерывно дифференцируемая функция $W^{(j\Delta t)}(t, x) : [t_0 + j\Delta t, \bar{T}_j] \times R^m \rightarrow R$, удовлетворяющая следующей системе уравнений в частных производных:*

$$\frac{f_j(t)}{1 - F_j(t)} W^{(j\Delta t)}(t, x) = W_t^{(j\Delta t)}(t, x) + \max_{u \in U} \left\{ \sum_{i=1}^n h_i(t, x, u) + W_x^{(j\Delta t)}(t, x) g(t, x, u) \right\}, \quad (3.99)$$

где $\lim_{t \rightarrow T^-} W^{(j\Delta t)}(t, x) = 0$, а $f_j(t)$ — функция плотности для случайной величины \bar{T}_j (3.92). Предположим, что максимум в (3.99) достигается при $u = u_j^*(t, x)$. Тогда $u = u_j^*(t, x)$ является оптимальным управлением в задаче (3.96), (3.97).

Теорема 3.3.1 (см. [197]) требует, чтобы функция $W^{(j\Delta t)}$ принадлежала пространству C^1 . Однако можно предположить непрерывность, рассматривая только т.н. вязкостные (viscosity) решения с использованием подхода Субботина [183; 184]. В модельном примере ниже решение $W^{(j\Delta t)}$ получается из пространства C^1 .

3.3.1.2 Условно-кооперативная траектория

Во время игры $\Gamma(x_0, T - t_0)$ игрокам доступна ограниченная информация о её структуре. Этой информации недостаточно, чтобы определить кооперативное

поведение для игроков во всей игре $\Gamma(x_0, T - t_0)$. Вместо кооперативной траектории в игре $\Gamma(x_0, T - t_0)$ будем строить условно-кооперативную траекторию.

Определение 3.3.2 Условно-кооперативная траектория $\{\hat{x}^*(t)\}_{t=t_0}^T$ — это комбинация кооперативных траекторий $x_j^*(t)$ в усеченных кооперативных подыграх $\bar{\Gamma}_j^c(x_{j-1}^*(t_0 + j\Delta t), t_0 + j\Delta t)$, определенных на последовательных временных интервалах $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$:

$$\{\hat{x}^*(t)\}_{t_0}^T = \begin{cases} x_0^*(t), & t \in [t_0, t_0 + \Delta t), \\ \dots, \\ x_j^*(t), & t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t), \\ \dots, \\ x_l^*(t), & t \in [t_0 + l\Delta t, t_0 + (l + 1)\Delta t], \end{cases} \quad (3.100)$$

На временном интервале $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$ условно-кооперативная траектория совпадает с кооперативной траекторией $x_j^*(t)$ в усеченной кооперативной подыгре $\bar{\Gamma}_j^c(x_{j-1}^*(t_0 + j\Delta t), t_0 + j\Delta t)$. В момент времени $t = t_0 + (j + 1)\Delta t$ обновляется информация об игровой структуре в позиции $x_j^*(t_0 + (j + 1)\Delta t)$. На временном интервале $(t_0 + (j + 1)\Delta t, t_0 + (j + 2)\Delta t]$ траектория $\hat{x}^*(t)$ совпадает с кооперативной траекторией $x_{j+1}^*(t)$ в усеченной кооперативной подыгре $\bar{\Gamma}_{j+1}^c(x_j^*(t_0 + (j + 1)\Delta t), t_0 + (j + 1)\Delta t)$, начинающейся в момент времени $t = t_0 + (j + 1)\Delta t$ в позиции $x_j^*(t_0 + (j + 1)\Delta t)$. При $j = 0$, $x_{j-1}^*(t_0 + j\Delta t) = x_0$.

3.3.1.3 Характеристическая функция

Для каждой коалиции $S \subset N$ и усеченной подыгры с номером $j = 0, \dots, l$ определим значения характеристической функции по аналогии с [185]:

$$V_j(S; x_{j,0}^*, t_0 + j\Delta t) = \begin{cases} \sum_{i \in N} K_i^j(x_{j,0}^*, t_0 + j\Delta t; u_j^*), & S = N, \\ \tilde{V}_j(S, x_{j,0}^*, t_0 + j\Delta t), & S \subset N, \\ 0, & S = \emptyset, \end{cases} \quad (3.101)$$

где $\tilde{V}_j(S, x_{j,0}^*, t_0 + j\Delta t)$ определяется как суммарный выигрыш игроков из коалиции S в равновесии по Нэшу $u_j^{NE} = (u_1^{NE,j}, \dots, u_n^{NE,j})$ в игре со

следующим набором игроков: коалиция S (действующая как один игрок) и игроки из множества $N \setminus S$, т.е. в игре с $|N \setminus S| + 1$ игроками.

Дележ $\xi_j(x_{j,0}^*, t_0 + j\Delta t)$ для каждой случайной усеченной кооперативной подыгры $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t)$ определяется как произвольный вектор, удовлетворяющий условиям

$$\begin{aligned} \xi_i^j(x_{j,0}^*, t_0 + j\Delta t) &\geq V_j(\{i\}, x_{j,0}^*, t_0 + j\Delta t), \quad i \in N, \\ \sum_{i \in N} \xi_i^j(x_{j,0}^*, t_0 + j\Delta t) &= V_j(N, x_{j,0}^*, t_0 + j\Delta t). \end{aligned} \quad (3.102)$$

Обозначим множество всевозможных дележей в случайной усеченной игре через

$E_j(x_{j,0}^*, t_0 + j\Delta t)$. В качестве принципа оптимальности (решения)

$$W_j(x_{j,0}^*, t_0 + j\Delta t) \subset E_j(x_{j,0}^*, t_0 + j\Delta t) \quad (3.103)$$

в каждой случайной усеченной подыгре $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t)$ будем использовать аналог ПРД-ядра, введенного в [100]. Построение этого решения основано на специальном классе ПРД [187].

Определение 3.3.3 *Функция $\beta_j(t, x_j^*)$, $t \in [t_0 + j\Delta t, \bar{T}_j]$, называется процедурой распределения дележа $\xi_j(x_{j,0}^*, t_0 + j\Delta t) \in E_j(x_{j,0}^*, t_0 + j\Delta t)$, если выполняется*

$$\xi_j(x_{j,0}^*, t_0 + j\Delta t) = \int_{t_0 + j\Delta t}^T (1 - F_j(\tau)) \beta_j(t, x_j^*(t)) d\tau. \quad (3.104)$$

Используя ПРД $\beta_j(t, x_j^*)$, можно определить правило распределения дележа $\xi_j(x_{j,0}^*, t_0 + j\Delta t)$ на временном интервале $[t_0 + j\Delta t, \bar{T}_j]$, где \bar{T}_j — случайная величина. Очевидно, что количество функций $\beta_j(t, x_j^*(t))$, удовлетворяющих уравнению (3.104), бесконечно, т.е. существует бесконечно много способов распределения кооперативного выигрыша между игроками. Однако, применяя известную формулу для класса игр со случайным горизонтом [194; 195],

$$\beta_j(t, x_j^*(t)) = \frac{f_j(t)}{1 - F_j(t)} \xi_j(x_j^*(t), t) - \frac{d}{d\tau} \xi_j(x_j^*(t), t), \quad (3.105)$$

можно определить единственную ПРД $\beta_j(t, x_j^*(t))$, обеспечивающую динамиче-

скую устойчивость [187] дележа $\xi_j(x_{j,0}^*, t_0 + j\Delta t)$ или кооперативное решение $W_j(x_{j,0}^*, t_0 + j\Delta t)$ (в случае нескольких принципов оптимальности).

Определение 3.3.4 *Решение $W_j(x_{j,0}^*, t_0 + j\Delta t)$ ($\xi_j(x_{j,0}^*, t_0 + j\Delta t)$) называется динамически устойчивым, если для каждого дележа $\xi_j(x_{j,0}^*, t_0 + j\Delta t) \in W_j(x_{j,0}^*, t_0 + j\Delta t)$ существует соответствующая ПРД $\beta_j(t, x_j^*)$, такая что при $\forall t \in [t_0 + j\Delta t, T]$ справедливо включение*

$$\left\{ \int_t^T (1 - F_j(t)) \beta_i^j(\tau, x_j^*) d\tau \right\} \in W(x_j^*(t), t) \left(\left\{ \int_t^T (1 - F_j(t)) \beta_i^j(\tau, x_j^*) d\tau \right\} = \xi_j(x_{j,0}^*, t_0 + j\Delta t) \right).$$

3.3.1.4 ПРД-ядро

Применим подход, предложенный в [100]: построим функции, которые могут использоваться в качестве ПРД для некоторых дележей, а затем составим соответствующие решения. Предположим, что характеристическая функция $V_j(S; x_j^*(t), t)$, $S \subset N$, непрерывно дифференцируема по t , $t \in [t_0 + j\Delta t, T]$, вдоль кооперативной траектории $x_j^*(t)$. Введем следующее обозначение:

$$U_j(S; x_j^*(t), t) = -\frac{d}{dt} V_j(S; x_j^*(t), t), \quad (3.106)$$

где $t \in [t_0 + j\Delta t, T]$ и $S \subseteq N$.

Определим $B_j(t, x_j^*)$ как набор интегрируемых векторных функций $\beta_j(t, x_j^*)$, удовлетворяющих следующим неравенствам:

$$B_j(t, x_j^*) = \left\{ \beta_j(t, x_j^*) = (\beta_1^j(t, x_j^*), \dots, \beta_n^j(t, x_j^*)) : \right. \\ \left. \sum_{i \in S} (1 - F_j(t)) \beta_i^j(t, x_j^*) \geq U_j(S, x_j^*(t), t), \right. \\ \left. \sum_{i \in N} \beta_i^j(t, x_j^*) = U_j(N, x_j^*(t), t), \forall S \subset N \right\}. \quad (3.107)$$

Предположим, что $B_j(t, x_j^*) \neq \emptyset$, $\forall t \in [t_0 + j\Delta t, T]$ и $j = 0, \dots, l$. Тогда, используя множество $B_j(t, x_j^*)$, можно определить следующее множество

векторов:

Определение 3.3.5 Множество всевозможных $\xi_j(x_j^*(t), t)$ для некоторых интегрируемых селекторов $\beta_j(t, x_j^*) \in B_j(t, x_j^*)$ будем называть ПРД-ядром и обозначать через $\bar{C}_j(x_j^*(t), t)$, где

$$\bar{C}_j(x_j^*(t), t) = \{\xi_j(x_j^*(t), t), t \in [t_0 + j\Delta t, T]\} \quad (3.108)$$

и, при $t \in [t_0 + j\Delta t, T]$,

$$\xi_j(x_j^*(t), t) = \int_t^T (1 - F_j(\tau)) \beta_j(\tau, x_j^*) d\tau. \quad (3.109)$$

Легко предположить, что распределение суммарного выигрыша игроков в игре $\Gamma(x_0, T - t_0)$ вдоль условно-кооперативной траектории $\{\hat{x}^*(t)\}_{t=t_0}^T$ может быть организовано как композиция ПРД для каждого временного интервала $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, l$, в соответствии со структурой игры $\Gamma(x_0, T - t_0)$. Ниже это предположение будет формализовано как новая концепция решения.

Семейство множеств $W_j(x_{j,0}^*, t_0 + j\Delta t) = \bar{C}_j(x_{j,0}^*, t_0 + j\Delta t)$ не составляет непосредственно решение игры $\Gamma(x_0, T - t_0)$. При любом $j = 0, \dots, l$ оптимальное решение для усеченной подыгры $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t)$ определено на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}_j]$. Это конкретное решение имеет смысл только на интервале $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$ потому, что информация об игровой структуре обновляется после каждого временного интервала Δt , и неразумно использовать решение, основанное на устаревшей информации. Необходимая информация может быть извлечена с помощью ПРД для каждой усеченной подыгры. Поэтому для построения оптимального решения для всей игры $\Gamma(x_0, T - t_0)$ используем набор ПРД $B_j(t)$ вместо набора дележей $\bar{C}_j(x_{j,0}^*, t_0 + j\Delta t)$.

3.3.2 Концепция кооперативного решения для игры с динамическим обновлением и случайным горизонтом

Для того, чтобы построить решение для дифференциальной игры $\Gamma(x_0, T - t_0)$ с динамическим обновлением информации, будем использовать семейство множеств $B_j(t, x_j^*)$, $j = 0, \dots, l$. Сначала построим результирующее множество ПРД для всей игры $\Gamma(x_0, T - t_0)$ следующим образом: для каждого набора $\beta_j(t, x_j^*) \in B_j(t, x_j^*)$, $j = 0, \dots, l$ определим функцию $\hat{\beta}(t, \hat{x}^*)$.

Определение 3.3.6 *Результирующая ПРД $\hat{\beta}(t, \hat{x}^*)$ — это функция, определенная как комбинация процедур распределения дележа $\beta_j(t, x_j^*) \in B_j(t, x_j^*)$ во всех усеченных кооперативных подыграх $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t)$, $j = 0, \dots, l$:*

$$\hat{\beta}(t, \hat{x}^*) = \begin{cases} (1 - F_0(t))\beta_0(t, x_0^*), t \in [t_0, t_0 + \Delta t], \\ \dots \\ (1 - F_j(t))\beta_j(t, x_j^*), t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t], \\ \dots \\ (1 - F_l(t))\beta_l(t, x_l^*), t \in [t_0 + l\Delta t, t_0 + (l + 1)\Delta t]. \end{cases} \quad (3.110)$$

Множество всевозможных функций $\hat{\beta}(t, \hat{x}^*)$ (3.110) для разных $\beta_j(t, x_j^*) \in B_j(t, x_j^*)$, $j = 0, \dots, l$, обозначим через $\hat{B}(t, \hat{x}^*)$.

Используя результирующую ПРД $\hat{\beta}(t, \hat{x}^*) \in \hat{B}(t)$, можно определить результирующий дележ, который может использоваться в качестве дележа в игре $\Gamma(x_0, T - t_0)$ с динамическим обновлением информации. Но остается вопрос: действительно ли результирующий дележ будет распределять суммарный кооперативный выигрыш вдоль условно-кооперативной траектории $\hat{x}^*(t)$? Этот факт доказан в Теореме 3.3.2.

Определение 3.3.7 *Результирующий дележ $\hat{\xi}(x_0, T - t_0)$ — это вектор, определенный следующим образом:*

$$\hat{\xi}(x_0, T - t_0) = \int_{t_0}^T \hat{\beta}(\tau, \hat{x}^*(\tau)) d\tau = \sum_{j=0}^l \left[\int_{j\Delta t}^{(j+1)\Delta t} (1 - F_j(\tau))\beta_j(\tau, x_j^*(\tau)) d\tau \right] \quad (3.111)$$

Обозначим через $\hat{W}(x_0, T - t_0)$ результирующее решение — множество всех результирующих дележей $\hat{\xi}(x_0, T - t_0)$, построенных по формулам (3.110), (3.111). В играх с динамическим обновлением информации множество \hat{W} используется в качестве решения.

Теорема 3.3.2 С помощью $\hat{\xi}(x_0, T - t_0) \in \hat{W}(x_0, T - t_0)$ можно распределить суммарный выигрыш (3.96) игроков вдоль условно-кооперативной траектории $\hat{x}^*(t)$ во время игры $\Gamma(x_0, T - t_0)$ и для $\forall t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, l$, справедливо соотношение

$$\sum_{i=1}^n \int_{t_0}^t \hat{\beta}_i(\tau, \hat{x}^*(\tau)) d\tau = \sum_{i=1}^n \left[\sum_{k=0}^{j-1} \left[\int_{k\Delta t}^{(k+1)\Delta t} (1 - F_k(\tau)) h_i(\hat{x}^*(\tau), \hat{u}^*(\tau)) d\tau \right] + \int_{t_0+j\Delta t}^t (1 - F_j(\tau)) h_i(\hat{x}^*(\tau), \hat{u}^*(\tau)) d\tau \right]. \quad (3.112)$$

Доказательство Для доказательства данного результата сначала рассмотрим последнюю случайную усеченную подыгру $\bar{\Gamma}_l(x_{l,0}^*, t_0 + l\Delta t)$. А именно, покажем, что для $\forall t \in [t_0 + l\Delta t, T]$ верно равенство

$$\sum_{i=1}^n \int_{t_0+l\Delta t}^t \hat{\beta}_i(\tau, \hat{x}^*(\tau)) d\tau = \sum_{i=1}^n \int_{t_0+l\Delta t}^t (1 - F_l(\tau)) h_i(\hat{x}^*(\tau), \hat{u}^*(\tau)) d\tau. \quad (3.113)$$

Действительно, максимальный суммарный выигрыш в этой игре определяется функцией $W^{(l\Delta t)}(t_0 + l\Delta t, x_{l,0}^*)$ (3.98). Согласно определению этой функции для $\forall t \in [t_0 + l\Delta t, T]$ имеем:

$$\begin{aligned} W^{(l\Delta t)}(t, \hat{x}^*(t)) = \\ \max_{u \in U} \left\{ \sum_{i \in N} K_i^l(\hat{x}^*(t), t; u) \right\} = \sum_{i=1}^n \int_t^T (1 - F_l(\tau)) h_i(\hat{x}^*(\tau), \hat{u}^*(\tau)) d\tau = \\ \sum_{i=1}^n \int_t^T \xi_i^l(\hat{x}^*(\tau), \tau) d\tau = \sum_{i=1}^n \int_t^T \beta_i^l(\tau, \hat{x}^*(\tau)) d\tau = \sum_{i=1}^n \int_t^T \hat{\beta}_i(\tau, \hat{x}^*(\tau)) d\tau. \quad (3.114) \end{aligned}$$

Однако,

$$W^{(l\Delta t)}(t_0 + l\Delta t, x_{l,0}^*) - W^{(l\Delta t)}(t, \hat{x}^*(t)) = \sum_{i=1}^n \int_{t_0+l\Delta t}^t (1 - F_l(\tau)) h_i(\hat{x}^*(\tau), \hat{u}^*(\tau)) d\tau, \quad (3.115)$$

где $\forall t \in [t_0 + l\Delta t, T]$. Из (3.114) и (3.115) следует, что соотношение (3.113) имеет место при $\forall t \in [t_0 + l\Delta t, T]$. Используя этот результат, докажем, что он также справедлив и для случайной усеченной подыгры $\bar{\Gamma}_{l-1}(x_{l-1,0}^*, t_0 + (l-1)\Delta t)$. Другими словами, покажем, что

$$\sum_{i=1}^n \int_{t_0+(l-1)\Delta t}^t \hat{\beta}_i(\tau, \hat{x}^*(\tau)) d\tau = \sum_{i=1}^n \int_{t_0+(l-1)\Delta t}^t (1 - F_{l-1}(\tau)) h_i(\hat{x}^*(\tau), \hat{u}^*(\tau)) d\tau, \quad (3.116)$$

где $\forall t \in [t_0 + (l-1)\Delta t, t_0 + l\Delta t]$. По аналогии с игрой $\bar{\Gamma}_l(x_{l,0}^*, t_0 + l\Delta t)$, при $\forall t \in [t_0 + (l-1)\Delta t, t_0 + l\Delta t]$ имеем:

$$W^{((l-1)\Delta t)}(t_0 + (l-1)\Delta t, x_{l-1,0}^*) - W^{((l-1)\Delta t)}(t, \hat{x}^*(t)) = \sum_{i=1}^n \int_{t_0+(l-1)\Delta t}^t (1 - F_{l-1}(\tau)) h_i(\hat{x}^*(\tau), \hat{u}^*(\tau)) d\tau. \quad (3.117)$$

Следовательно, равенство (3.116) верно. Далее, необходимо продолжить рассуждения в обратном порядке, вплоть до первой случайной усеченной подыгры $\bar{\Gamma}_0(x_0, t_0)$. Это позволит нам объединить результаты (3.113), (3.116) и показать, что для $\forall t \in [t_0, T]$ (3.112) выполняется. Теорема доказана. \square

3.3.2.1 Динамическая устойчивость концепции решения

Результирующее решение $\hat{W}(x_0, T - t_0)$ является динамически устойчивым по построению. Оказывается, что $\hat{W}(x_0, T - t_0)$ обладает и другим интересным

СВОЙСТВОМ.

Теорема 3.3.3 *Результирующее решение $\hat{W}(x_0, T - t_0)$ является сильно динамически устойчивым в игре $\Gamma(x_0, T - t_0)$.*

Доказательство Предположим, что в игре $\Gamma(x_0, T - t_0)$ со случайным обновлением информации игроки согласились выбрать результирующий вектор $\hat{\xi}(x_0, T - t_0) \in \hat{W}(x_0, T - t_0)$. Это означает, что в течение игры, в каждой кооперативной случайной усеченной подыгре $\bar{\Gamma}_c^j(x_{j,0}^*, t_0 + j\Delta t)$ они будут выбирать дележ $\xi_j(x_{j,0}^*, t_0 + j\Delta t) \in \bar{C}_j(x_{j,0}^*, t_0 + j\Delta t)$ с соответствующей ПРД $\beta_j(t, x_j^*) \in B_j(t, x_j^*)$, $t \in [t_0 + j\Delta t, T]$. Таким образом, в каждой усеченной подыгре игроки будут использовать ПРД $\hat{\beta}(t, \hat{x}^*) = \beta_j(t, x_j^*)$ и распределять выигрыш между собой следующим образом:

$$\int_{t_0}^T \hat{\beta}(\tau, \hat{x}^*(\tau)) d\tau = \sum_{j=0}^l \int_{t_0 + j\Delta t}^{t_0 + (j+1)\Delta t} (1 - F_j(t)) \beta_j(t, x_j^*) dt.$$

Предположим, что в момент времени $t = t_{br}$, где $t_{br} \in [t_0 + k\Delta t, T]$, $k = 0, \dots, l$ в случайной усеченной подыгре $\bar{\Gamma}_c^k(x_{k,0}^*, t_0 + k\Delta t)$ игроки решили выбрать другой дележ $\xi'_k(x_k^*(t_{br}), t_{br})$ из ПРД-ядра $\bar{C}_k(x_k^*(t_{br}), t_{br})$. Тогда существует ПРД $\beta'_k(t, x_k^*) \in B_k(t, x_k^*)$, $t \in [t_{br}, T]$, которая соответствует этому дележу:

$$\xi'_k(x_k^*(t_{br}), t_{br}) = \int_{t_{br}}^T (1 - F_j(t)) \beta'_k(t, x_k^*) dt. \quad (3.118)$$

В этом случае, суммарный выигрыш во всей игре будет распределен в соответствии с результирующим вектором $\hat{\xi}'(x_{t_0}, T - t_0)$ с помощью результирующей ПРД

$$\hat{\beta}'(t) = \begin{cases} (1 - F_k(t)) \beta_k(t, x_k^*), & t \in [t_0 + k\Delta t, t_{br}), \\ (1 - F_k(t)) \beta'_k(t, x_k^*), & t \in [t_{br}, t_0 + (k+1)\Delta t], \\ (1 - F_j(t)) \beta_j(t, x_j^*), & t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t], \end{cases}$$

где $j \neq k$, $j = 0, \dots, l$. Соответствующий результирующий дележ имеет вид

$$\begin{aligned} \hat{\xi}'(x_0, T - t_0) = & \int_{t_0}^T \hat{\beta}'(t, \hat{x}^*) dt = \sum_{\substack{j=0 \\ j \neq k}}^l \int_{t_0+j\Delta t}^{t_0+(j+1)\Delta t} (1 - F_j(t)) \beta_j(t, x_j^*) dt + \\ & \int_{t_0+k\Delta t}^{t_{br}} (1 - F_k(t)) \beta_k(t, x_k^*) dt + \int_{t_{br}}^{t_0+(k+1)\Delta t} (1 - F_k(t)) \beta'_k(t, x_k^*) dt. \end{aligned} \quad (3.119)$$

Так как $\beta'_k(t, x_k^*) \in B_k(t, x_k^*)$, $t \in [t_{br}, T]$, результирующая ПРД $\hat{\beta}'(t, \hat{x}^*)$ также принадлежит множеству $\hat{B}(t, \hat{x}^*)$. Согласно определению $\hat{W}(x_0, T - t_0)$, все векторы $\hat{\xi}'(x_0, T - t_0)$, полученные по формуле (3.111) с использованием $\hat{\beta}'(t, \hat{x}^*)$ из множества $\hat{B}(t, \hat{x}^*)$, называются результирующим решением $\hat{W}(x_0, T - t_0)$ игры $\Gamma(x_0, T - t_0)$. В (3.119) дележ $\hat{\xi}'(x_0, T - t_0)$ построен с помощью ПРД $\hat{\beta}'(t, \hat{x}^*)$ из множества $\hat{B}(t, \hat{x}^*)$, и его можно рассматривать как результирующий дележ $\hat{\xi}'(x_{t_0}, T - t_0)$, принадлежащий начальному решению $\hat{W}(x_0, T - t_0)$. Теорема доказана. \square

3.3.3 Кооперативная игра добычи ограниченного ресурса с динамическим обновлением информации и случайным горизонтом

Следующая игра добычи ресурса с двумя игроками рассмотрена Йоргенсенем и Янгом [69]. Проблема динамической устойчивости в этой игре изучена Дэвидом Янгом с соавторами [188]. Реализованы три модели конкуренции: со стохастическим прогнозом, с неслучайным прогнозом и без прогноза. Здесь рассматривается игра добычи ресурса с тремя игроками и специальной формой кооперативного решения, описанная в разделе 3.3.1.3 и [100]. Характеристическая функция для каждой коалиции получена в явном виде в соответствии с [185] и представлена ниже. Кроме того, к данной модели применен подход с динамическим обновлением информации и случайным горизонтом. В заключение демонстрируется сильная динамическая устойчивость построенного решения.

В рамках следующей модели получено частное численное решение задачи; однако, общее аналитическое решение не может быть найдено. Чтобы применить подход с динамическим обновлением информации к общему классу кооперативных дифференциальных игр, необходимо решить две основные проблемы. Первая состоит в том, чтобы решить задачу (3.96) при условии

(3.97) для каждой усеченной подыгры. Математически это классическая задача управления, и существует множество способов ее решения. При решении этой задачи получены приближенные кооперативные стратегии, кооперативные траектории $x_j^*(t)$ и соответствующий суммарный выигрыш (3.96). Вторая проблема — определить, каким образом будет распределен кооперативный выигрыш между игроками. Необходимо вычислить характеристическую функцию (3.101) для каждой усеченной подыгры вдоль кооперативной траектории. Для этого можно использовать коэволюционные алгоритмы [200], подходящие для теоретико-игровых задач. После вычисления характеристических функций может быть определено решение для каждой усеченной подыгры (например, ПРД-ядро), а затем может быть вычислено соответствующее результирующее решение (3.111).

3.3.3.1 Исходная игра

Рассмотрим экономику, наделенную одним возобновляемым ресурсом, с $n \geq 2$ фирмами, которые занимаются добычей ресурса (игроками). Обозначим через $u_i(t)$ объем ресурса, добытого игроком $i \in N$ в момент времени t , при условии, что каждый игрок выбирает скорость добычи. Пусть $x(t) \in X \subset R$ описывает запас ресурса в момент времени t . Динамика запаса возобновляемого ресурса удовлетворяет дифференциальному уравнению

$$\dot{x} = a\sqrt{x(t)} - bx(t) - \sum_{i=1}^3 u_i, \quad x(t_0) = x_0, \quad (3.120)$$

где $a\sqrt{x(t)} - bx(t)$ — скорость естественной эволюции ресурса, а $u_i \in [0, d]$, $d > 0$, $i = \overline{1, 3}$.

Стоимость добычи для игрока $i \in N$ зависит от объема добытого ресурса $u_i(t)$, запаса ресурса $x(t)$ и параметра c_i , $i = \overline{1, 3}$:

$$K_i(x_0, t_0; u) = \int_{t_0}^T \sqrt{u_i(\tau)} - \frac{c_i}{\sqrt{x(\tau)}} u_i(\tau) d\tau, \quad (3.121)$$

где c_i — константа, удовлетворяющая условию $c_i \neq c_k$, $\forall i \neq k = \overline{1, 3}$. Считаем,

что все параметры x_0 , T , a , b , d , и c_i , $i = \overline{1, 3}$ всегда неотрицательны в соответствующей задаче управления.

3.3.3.2 Случайная усеченная подыгра

Исходная игра $\Gamma(x_0, T - t_0)$ определена на временном интервале $[t_0, T]$. Предположим, что в любой момент времени $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, l$, игроки имеют усеченную информацию об игре. Она включает в себя информацию об уравнениях движения и функциях выигрыша на временном интервале $[t_0 + j\Delta t, \bar{T}_j]$, где \bar{T}_j — случайная величина на усеченном временном интервале $[\max(t_0 + (j+1)\Delta t, \bar{T}_{j-1}), T]$ с экспоненциальной функцией распределения $F_j(t)$ и функцией плотности $f_j(t)$:

$$F_j(t) = \frac{1 - \exp(-\lambda(t - \max(t_0 + (j+1)\Delta t, \bar{t}_{j-1})))}{1 - \exp(-\lambda(T - \max(t_0 + (j+1)\Delta t, \bar{t}_{j-1})))}, \quad (3.122)$$

$$f_j(t) = \frac{\lambda \exp(-\lambda(t - \max(t_0 + (j+1)\Delta t, \bar{t}_{j-1})))}{1 - \exp(-\lambda(T - \max(t_0 + (j+1)\Delta t, \bar{t}_{j-1})))}. \quad (3.123)$$

Экспоненциальное распределение широко используется для описания времени между событиями в пуассоновском процессе. Под событиями понимается изменение структуры игры. Кроме того, введем обозначение

$$\Lambda_j(t) = \begin{cases} \frac{f_j(t)}{1-F_j(t)}, & t \in [\max(t_0 + (j+1)\Delta t, \bar{t}_{j-1}), T], \\ 0, & t \in [t_0 + j\Delta t, \max(t_0 + (j+1)\Delta t, \bar{t}_{j-1})]. \end{cases}$$

Усеченная информация формализуется в случайной усеченной подыгре $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t)$. Уравнения движения и начальные условия для этой подыгры имеют следующий вид:

$$\dot{x} = a\sqrt{x(t)} - bx(t) - \sum_{i=1}^3 u_i, \quad x(t_0 + j\Delta t) = x_{j,0}. \quad (3.124)$$

В соответствии с (3.95) функция выигрыша игрока $i \in N$ может быть

представлена в виде:

$$K_i^j(x_{j,0}, t_0 + j\Delta t; u) = \int_{t_0 + j\Delta t}^T (1 - F_j(\tau)) h_i(x(\tau), u(\tau)) d\tau. \quad (3.125)$$

Рассмотрим случай, когда игроки соглашаются на кооперацию в случайной усеченной подыгре $\bar{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t)$. Следуя принципу оптимальности, они максимизируют суммарный выигрыш и разделяют ожидаемый кооперативный выигрыш используя решение пропорциональное некооперативным выигрышам агентов.

3.3.3.3 Кооперативная траектория

Далее, рассмотрим случайную усеченную подыгру $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t)$. Максимальный суммарный выигрыш в игре $\bar{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t)$ имеет следующий вид [69]:

$$W^{j\Delta t}(t, x) = A^j(t)\sqrt{x} + C^j(t), \quad (3.126)$$

где функции $A^j(t)$ и $C^j(t)$ удовлетворяют системе дифференциальных уравнений

$$\begin{aligned} \dot{A}^j(t) &= \left[\Lambda_j(t) + \frac{b}{2} \right] A^j(t) - \sum_{i=1}^3 \left[\frac{1}{4 \left[c_i + \frac{A^j(t)}{2} \right]} \right], \\ \dot{C}^j(t) &= \Lambda_j(t) C^j(t) - \frac{a}{2} A^j(t) \end{aligned} \quad (3.127)$$

с граничными условиями $\lim_{t \rightarrow T^-} A^j(t) = \lim_{t \rightarrow T^-} C^j(t) = 0$.

Оптимальная кооперативная траектория $x_j^*(t)$ случайной усеченной подыгры $\bar{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t)$ может быть в явном виде представлена явно [69] на полном временном интервале $[t_0 + j\Delta t, T]$. Траектория с начальным условием $x = x_{j,0}^*$ равна

$$x_j^*(t) = \varpi_j^2(t_0 + j\Delta t, t) \left[\sqrt{x_{j,0}^*} + \frac{1}{2}a \cdot \int_{t_0 + j\Delta t}^t \varpi_j(t_0 + j\Delta t, \tau)^{-1} d\tau \right]^2, \quad (3.128)$$

где $t \in (t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$,

$$\varpi_j(t_0 + j\Delta t, t) = \exp \int_{t_0 + j\Delta t}^t - \left[\frac{1}{2}b + \sum_{i=1}^3 \left[\frac{1}{4 \left[c_i + \frac{A^j(\tau)}{2} \right]^2} \right] \right] d\tau. \quad (3.129)$$

Начальное положение для кооперативной траектории в каждой усеченной подыгре устанавливается из предыдущей усеченной подыгры: $x_{0,0}^* = x_0$ и $x_{j,0}^* = x_{j-1}^*(t_0 + j\Delta t)$ для $j = 0, \dots, l$. Определим условно-кооперативную траекторию $\hat{x}^*(t)$ в игре с динамическим обновлением информации:

$$\hat{x}_j^*(t) = x_j^*(t), \quad t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t], \quad (3.130)$$

где $j = 0, \dots, l$.

3.3.3.4 Характеристическая функция

Чтобы распределить кооперативный выигрыш в каждой случайной усеченной подыгре, необходимо определить значения характеристической функции $V_j(S; x_{j,0}, t_0 + j\Delta t)$ ($V_j(S; x_j^*(t), t)$) для каждой коалиции $S \subset N$. Согласно формуле (3.101) максимальный кооперативный выигрыш $W_j(t_0 + j\Delta t, x_{j,0})$ (3.126) соответствует значению характеристической функции гранд коалиции $V_j(N; x_{j,0}, t_0 + j\Delta t)$ в случайной усеченной подыгре $\bar{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t)$:

$$V_j(N; x_j^*(t), t) = W^{j\Delta t}(t, x_j^*(t)), \quad (3.131)$$

где $t \in [t_0 + j\Delta t, T]$, $j = 0, \dots, l$. Далее необходимо определить значения характеристической функции для следующих коалиций:

$$\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}. \quad (3.132)$$

Согласно (3.101), равновесие по Нэшу и, как следствие, $V_j(\{i\}; x_j^*(t), t)$ определены для коалиций, состоящих из одного игрока $\{i\}$, $i = \overline{1, 3}$.

3.3.3.5 Коалиции, состоящие из одного игрока

Равновесие по Нэшу в случайной усеченной подыгре $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t)$ определяется следующими позиционными стратегиями игроков:

$$u_i^j(t, x) = \frac{x}{4[c_i + A_i^j(t)/2]^2}, \quad i = \overline{1, 3}, \quad (3.133)$$

где функции $A_i^j(t)$ находятся из системы дифференциальных уравнений

$$\begin{aligned} \dot{A}_i^j(t) &= A_i^j(t) \left[\Lambda_j(t) + \frac{b}{2} + \sum_{k \neq i} \frac{1}{8(c_k + A_k^j(t)/2)^2} \right] - \frac{1}{4(c_i + A_i^j(t)/2)}, \\ \dot{C}_i^j(t) &= \Lambda_j(t) C_i^j(t) - \frac{a}{2} A_i^j(t) \end{aligned}$$

при $i = \overline{1, 3}$, с граничными условиями $\lim_{t \rightarrow T^-} A_i^j(t) = 0$ и $\lim_{t \rightarrow T^-} C_i^j(t) = 0$.

Выигрыш игрока $i = \overline{1, 3}$ в ситуации равновесия по Нэшу определяется функцией

$$V_i^j(t, x) = A_i^j(t)\sqrt{x} + C_i^j(t), \quad i = \overline{1, 3}. \quad (3.134)$$

Таким образом, значение характеристической функции для коалиций, состоящих из одного игрока $S = \{i\}$, $i \in N$, вычисляется следующим образом:

$$V_j(\{i\}; x_j^*(t), t) = V_i^j(t, x_j^*(t)), \quad (3.135)$$

где $t \in [t_0 + j\Delta t, T]$, $j = 0, \dots, l$.

3.3.3.6 Коалиции, состоящие из двух игроков

В соответствии с формулой (3.10), значение характеристической функции $V_j(S; x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ для коалиций, состоящих из двух игроков $S = \{1, 2\}, \{1, 3\}, \{2, 3\}$, определяется как суммарный выигрыш игроков из коалиции S в ситуации равновесия по Нэшу $u_j^{NE} = (u_1^{NE,j}, u_2^{NE,j}, u_3^{NE,j})$ в игре со следующим набором игроков: коалиция S (один игрок) и игроки из множества $N \setminus S$, т.е. в игре $|N \setminus S| + 1 = 2$ лиц. В нашем случае это означает, что игроки из коалиции S действуют как один игрок, а другие игроки из множества $N \setminus S$

действуют независимо. Используя этот подход, определим равновесие между двумя игроками в следующем виде: комбинированный игрок (коалиция S) и игрок, не входящий в коалицию S (коалиция N/S).

Рассмотрим формулы для $V_j(S; x_{j,0}, t_0 + j\Delta t)$ в случае, когда $S = \{1, 2\}$, формулы для вычисления характеристической функции остальных коалиций можно получить по такому же принципу. Выигрыши игроков в этом случае имеют следующий вид:

$$\begin{aligned} V_{\{1,2\}}^j(t, x) &= A_{\{1,2\}}^j(t)\sqrt{x} + C_{\{1,2\}}^j(t), \\ V_3^j(t, x) &= A_3^j(t)\sqrt{x} + C_3^j(t), \end{aligned}$$

где функции $A_{\{1,2\}}^j(t)$, $A_3^j(t)$, $C_{\{1,2\}}^j(t)$ и $C_3^j(t)$ удовлетворяют системе дифференциальных уравнений

$$\begin{aligned} \dot{A}_{\{1,2\}}^j(t) &= A_{\{1,2\}}^j(t) \left[\Lambda_j(t) + \frac{b}{2} + \frac{1}{8(c_3 + A_3^j(t)/2)^2} \right] - \sum_{k \in S} \frac{1}{4(c_k + A_{\{1,2\}}^j(t)/2)}, \\ \dot{A}_3^j(t) &= A_3^j(t) \left[\Lambda_j(t) + \frac{b}{2} + \sum_{k \in S} \frac{1}{8(c_k + A_{\{1,2\}}^j(t)/2)^2} \right] - \frac{1}{4(c_3 + A_3^j(t)/2)}, \\ \dot{C}_{\{1,2\}}^j(t) &= \Lambda_j(t)C_{\{1,2\}}^j(t) - \frac{a}{2}A_{\{1,2\}}^j(t), \\ \dot{C}_3^j(t) &= \Lambda_j(t)C_3^j(t) - \frac{a}{2}A_3^j(t) \end{aligned}$$

с начальными условиями $\lim_{t \rightarrow T^-} A_{\{1,2\}}^j(t) = \lim_{t \rightarrow T^-} A_3^j(t) = 0$, $\lim_{t \rightarrow T^-} C_{\{1,2\}}^j(t) = \lim_{t \rightarrow T^-} C_3^j(t) = 0$.

Таким образом, значение характеристической функции коалиции $S = \{1, 2\}$ вычисляется следующим образом:

$$V_j(\{1, 2\}; x_j^*(t), t) = V_{\{1,2\}}^j(t, x_j^*(t)), \quad (3.136)$$

где $t \in [t_0 + j\Delta t, T]$, $j = 0, \dots, l$.

3.3.3.7 ПРД-ядро

Используя значения характеристической функции $V_j(S; x_{j,0}, t_0 + j\Delta t)$, $\forall S \subset N$ (3.131), (3.135), (3.136) и формулу (3.107), построим множество $B_j(t, x_j^*)$ как набор интегрируемых вектор-функций $\beta_j(t, x_j^*)$, удовлетворяющих следующим условиям:

$$\begin{aligned} \sum_{i=1}^3 (1 - F_j(t)) \beta_i^j(t, x_j^*) &= -\frac{d}{dt} V_j(\{1, 2, 3\}; x_j^*(t), t), \\ (1 - F_j(t)) (\beta_1^j(t, x_j^*) + \beta_2^j(t, x_j^*)) &\geq -\frac{d}{dt} V_j(\{1, 2\}; x_j^*(t), t), \\ (1 - F_j(t)) (\beta_1^j(t, x_j^*) + \beta_3^j(t, x_j^*)) &\geq -\frac{d}{dt} V_j(\{1, 3\}; x_j^*(t), t), \\ (1 - F_j(t)) (\beta_2^j(t, x_j^*) + \beta_3^j(t, x_j^*)) &\geq -\frac{d}{dt} V_j(\{2, 3\}; x_j^*(t), t), \\ (1 - F_j(t)) \beta_1^j(t, x_j^*) &\geq -\frac{d}{dt} V_j(\{1\}; x_j^*(t), t), \\ (1 - F_j(t)) \beta_2^j(t, x_j^*) &\geq -\frac{d}{dt} V_j(\{2\}; x_j^*(t), t), \\ (1 - F_j(t)) \beta_3^j(t, x_j^*) &\geq -\frac{d}{dt} V_j(\{3\}; x_j^*(t), t). \end{aligned} \quad (3.137)$$

Затем, объединяя множества $B_j(t, x_j^*)$, $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, l$, для всех случайных усеченных подыгр, строим множество $\hat{B}(t, \hat{x}^*)$. Далее вычисляем множество всех возможных дележей $\hat{\xi}(x_0, T - t_0) \in \hat{W}(x_0, T - t_0)$ (3.111).

Пошаговое построение ПРД-ядра для линейно-квадратичной игры управления загрязнением представлено в статье [100].

3.3.3.8 Численный пример

Рассмотрим игру, заданную на временном интервале длиной $T = 4$. На интервалах времени $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$ информация об игре известна на интервале длиной \bar{T}_j , где \bar{T}_j — случайная величина (3.122) с $\lambda = 0.5$. Информация об игре обновляется с периодом $\Delta t = 1$. Зафиксируем следующие параметры: для уравнений движения, $a = 10$ и $b = 0.5$; для функции выигрыша, $c_1 = 0.15$, $c_2 = 0.65$ и $c_3 = 0.45$; для начальных условий, $t_0 = 0$ и $x_0 = 200$.

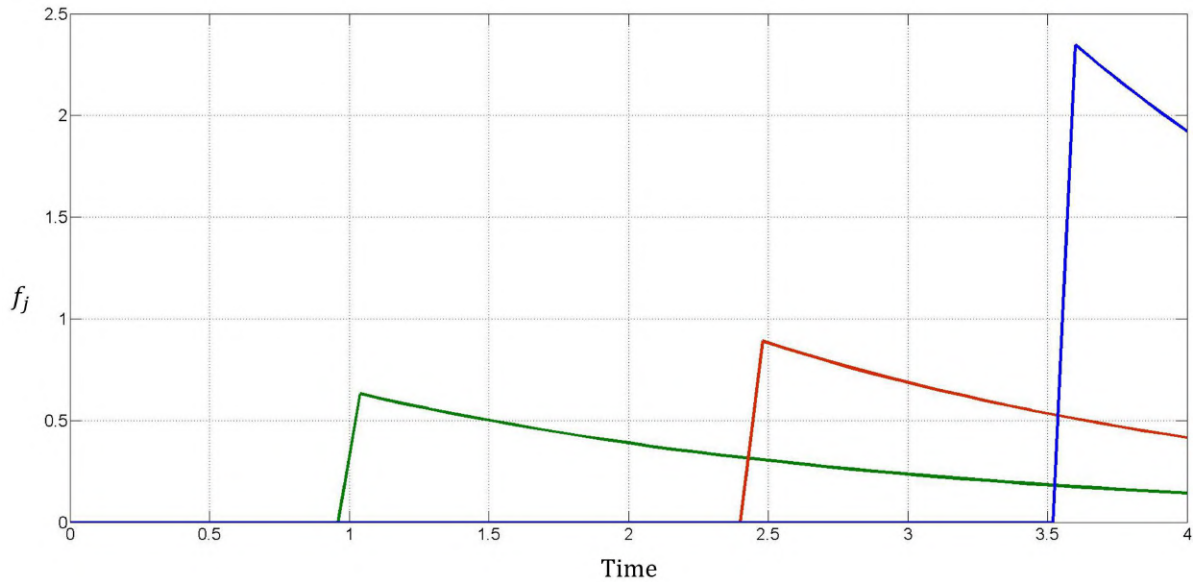


Рис. 3.10: Плотность распределения $f_j(t)$ информационного горизонта, $j = 0, 1, 2$ (3.123) для каждой случайной усеченной подыгры.

Во время обновления информации об игре происходит реализация случайного временного горизонта для текущей усеченной подыгры:

$$\bar{t}_0 = 2.423, \quad \bar{t}_1 = 3.538, \quad \bar{t}_2 = 3.871, \quad \bar{t}_3 = 4.$$

Сгенерированное значение информационного горизонта в текущей усеченной подыгре влияет на распределение продолжительности следующей усеченной подыгры. На рисунке 3.10 изображены плотности распределения. Видно, каким образом происходила генерация \bar{T}_j , и как менялась плотность распределения $f_j(t)$ (3.123) информационного горизонта.

На рисунках 3.11, 3.12 и 3.13 изображены кооперативные стратегии для каждого игрока, рассчитанные в игре с динамическим обновлением информации со случайным горизонтом (негладкая сплошная линия) и кооперативные стратегии в исходной игре [69] (гладкая пунктирная линия).

Условно-кооперативная траектория $\hat{x}^*(t)$ составлена из решений случайных усеченных подыгр $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t)$ с динамической системой (3.124). На рисунке 3.14 представлено сравнение следующих траекторий: условно-кооперативная траектория $\hat{x}^*(t)$ (толстая сплошная линия) в игре со случайным обновлением информации, условно-кооперативная траектория $\bar{x}^*(t)$ (тонкая пунктирная линия) в игре с динамическим обновлением информации [76] (детерминированный

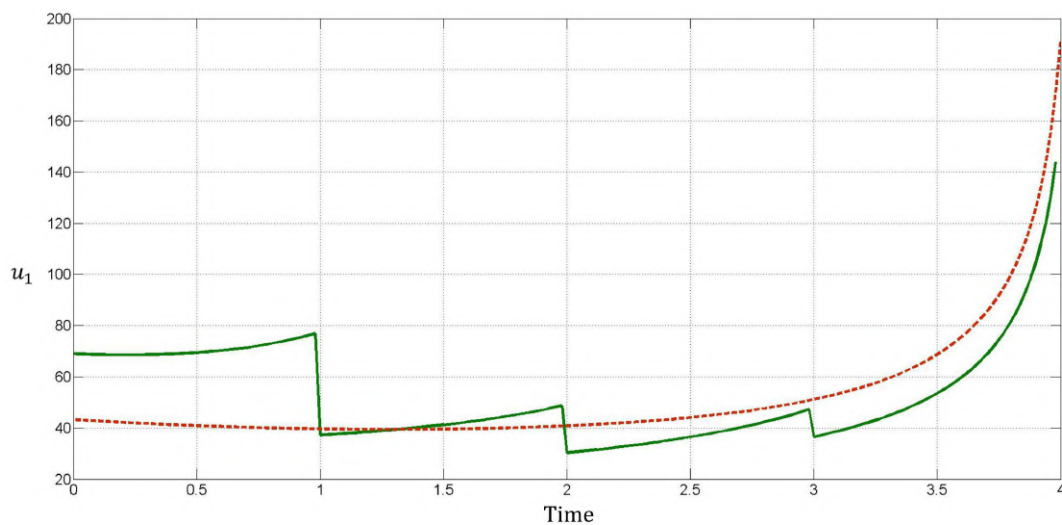


Рис. 3.11: Кооперативные стратегии для игрока 1 при динамическом обновлении информации со случайным горизонтом (негладкая сплошная линия) и кооперативные стратегии в исходной игре [69] (гладкая пунктирная линия).

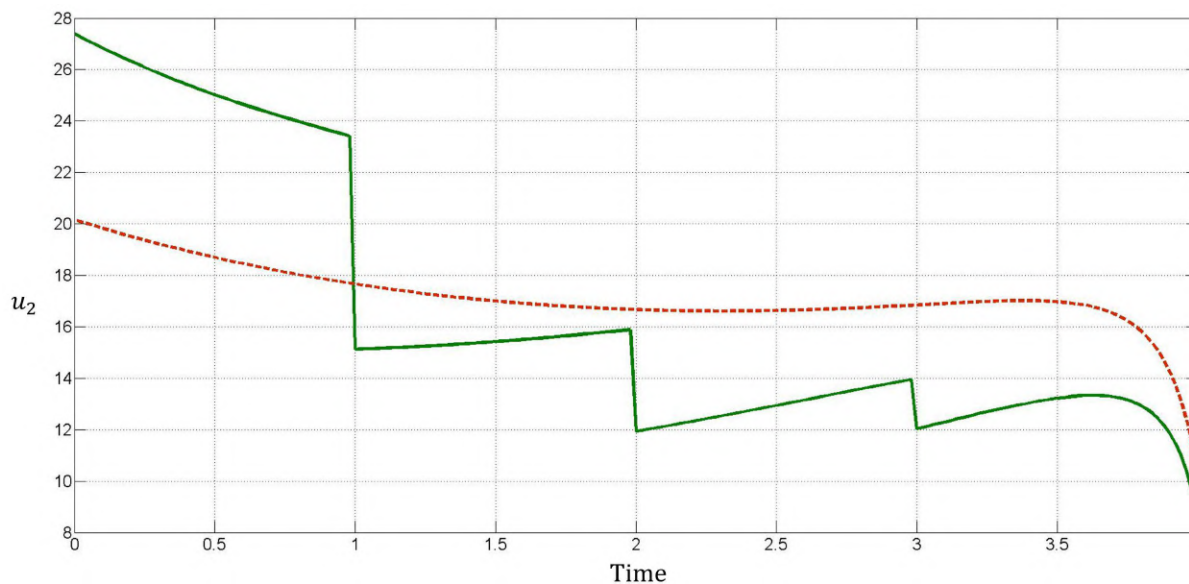


Рис. 3.12: Кооперативные стратегии для игрока 2 при динамическом обновлении информации со случайным горизонтом (негладкая сплошная линия) и кооперативные стратегии в исходной игре [69] (гладкая пунктирная линия).

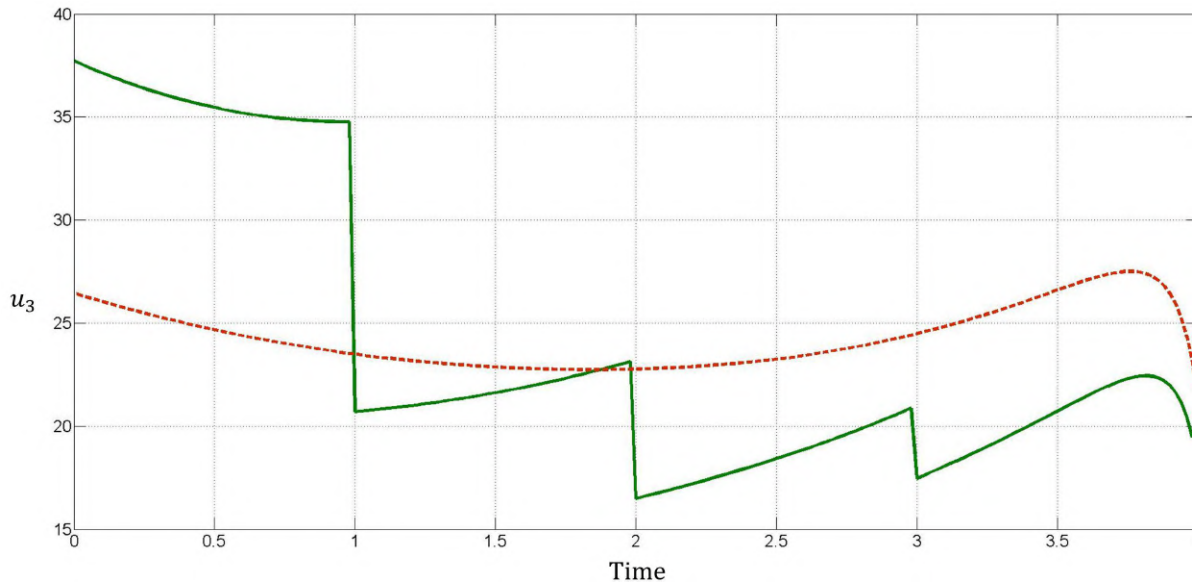


Рис. 3.13: Кооперативные стратегии для игрока 3 при динамическом обновлении информации со случайным горизонтом (негладкая сплошная линия) и кооперативные стратегии в исходной игре [69] (гладкая пунктирная линия).

горизонт $\bar{T} = 2$), и кооперативная траектория $x^*(t)$ (пунктирная линия) в исходной игре $\Gamma(x_0, T - t_0)$. Кооперативная траектория $x^*(t)$ определена в [69]. На следующих двух рисунках изображены условно-кооперативная траектория $\hat{x}^*(t)$ или $\bar{x}^*(t)$ и соответствующие кооперативные траектории в каждой из случайных усеченных подыгр.

Далее для того, чтобы распределить суммарный выигрыш между игроками, необходимо определить набор ПРД $\beta_j(t, x_j^*)$ для каждой случайной усеченной подыгры $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t)$, $j = 0, \dots, l$. Для этого при фиксированных параметрах модели значения характеристической функции $V_j(S; x_j^*(t), t)$, $S \subset N$, для каждой случайной усеченной подыгры $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t)$ найдены численно.

С помощью значений характеристической функции $V_j(S; x_j^*(t), t)$, $S \subset N$, построено множество $B_j(t, x_j^*)$, $j = 0, \dots, l$ (3.107). За счет комбинации множеств $B_j(t, x_j^*)$ получено множество ПРД для всей игры $\hat{B}(t, \hat{x}^*)$. На основе $\hat{B}(t, \hat{x}^*)$ по формуле (3.111) построена концепция решения $\hat{W}(x_0, T - t_0)$.

Продемонстрируем свойство сильной динамической устойчивости решения \hat{W} . Предположим, что в начале игры $\Gamma(x_0, T - t_0)$ игроки договорились использовать пропорциональное решение. Для каждой случайной усеченной подыгры $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t)$ с помощью ПРД определим пропорциональное

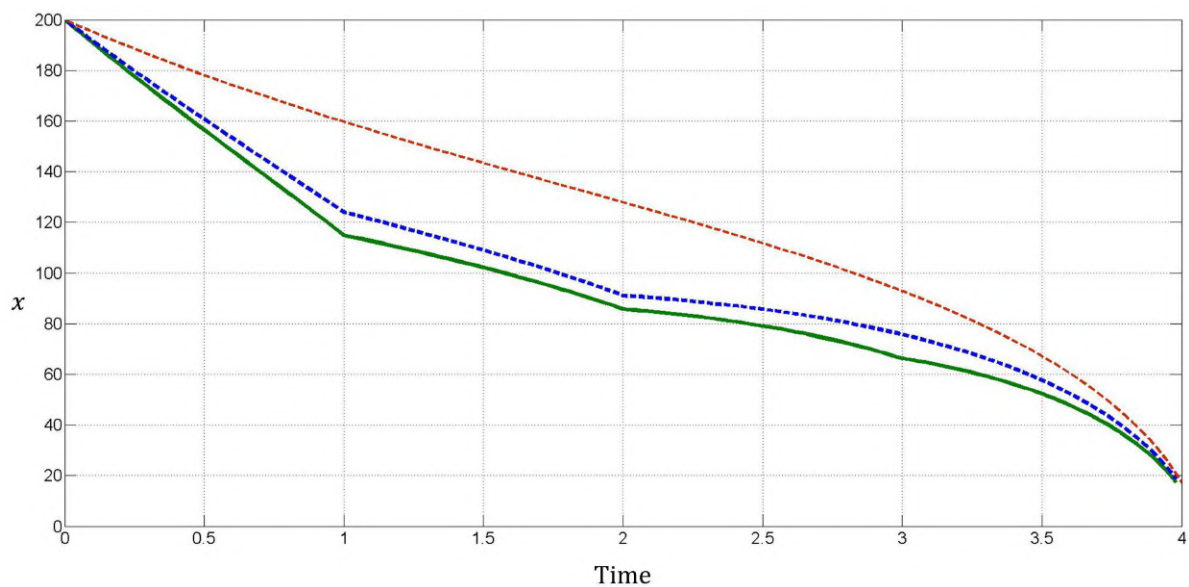


Рис. 3.14: Траектория запасов ресурса $\hat{x}^*(t)$ (толстая сплошная линия) в игре со случайным обновлением информации, траектория $\tilde{x}^*(t)$ (тонкая сплошная линия) в игре с динамическим обновлением информации, и кооперативная траектория $x^*(t)$ в исходной игре $\Gamma(x_0, T - t_0)$ (тонкая пунктирная линия).

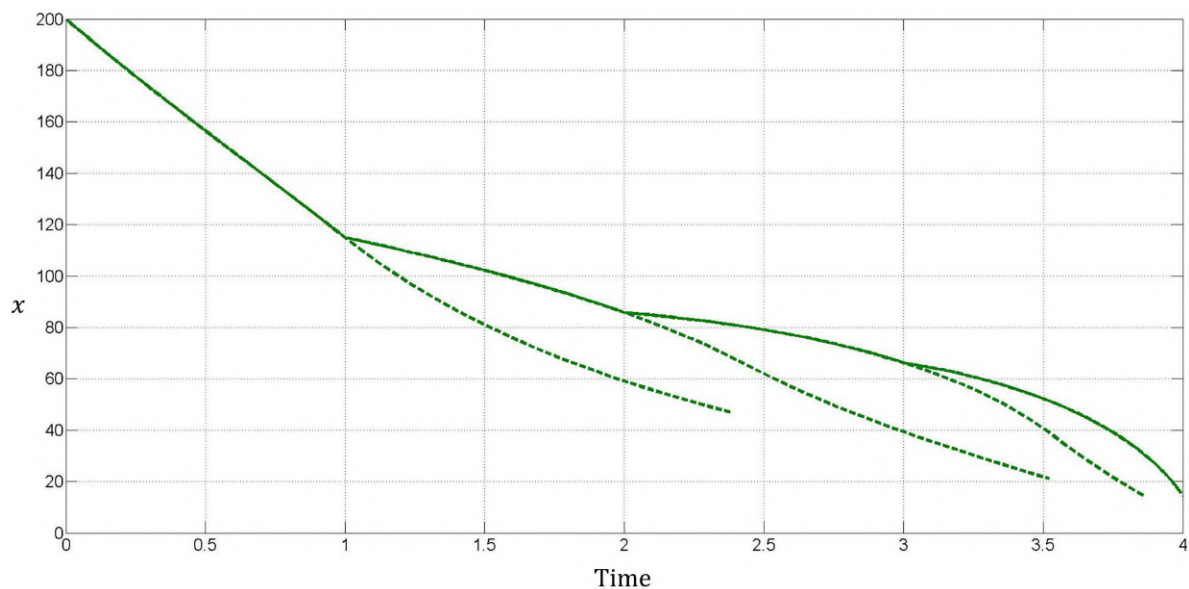


Рис. 3.15: Траектория запасов ресурса $\hat{x}^*(t)$ (толстая сплошная линия) в игре со случайным обновлением информации и соответствующие кооперативные траектории (пунктирные линии).

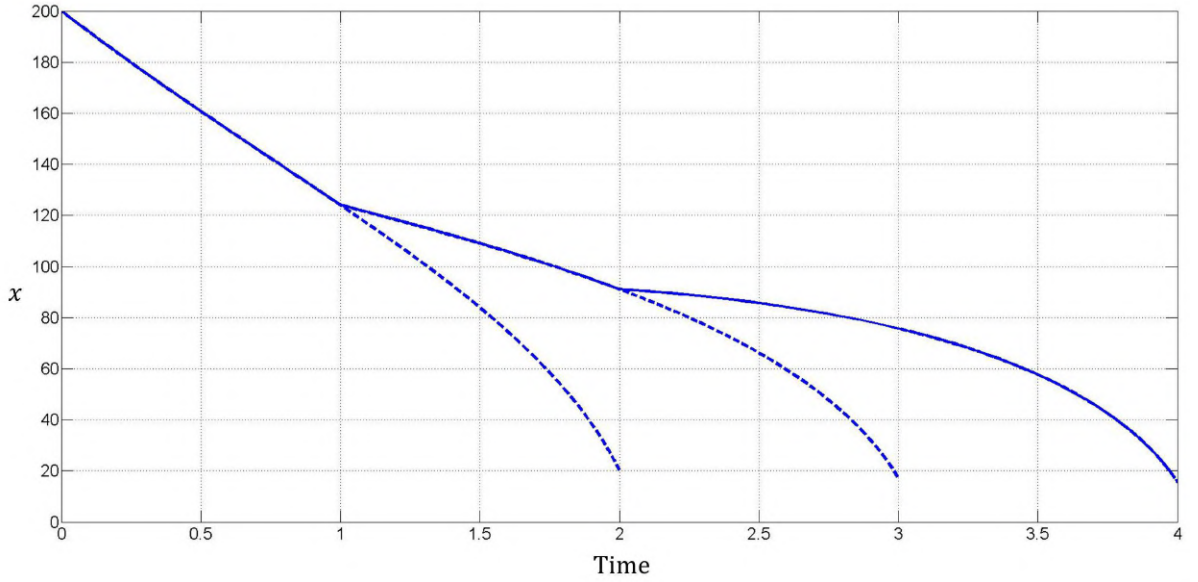


Рис. 3.16: Траектория запасов ресурса $\bar{x}^*(t)$ (сплошная линия) в игре с динамическим обновлением информации и соответствующие кооперативные траектории (пунктирные линии).

решение для игроков $i \in N$ следующим образом:

$$(1 - F_j(t))\beta_{i,j}^{Prop}(t, x_j^*) = \frac{U_j(\{i\}; x_{j,0}^*, t_0 + j\Delta t)}{\sum_{i \in N} U_j(\{i\}; x_{j,0}^*, t_0 + j\Delta t)} U_j(N; x_{j,0}^*, t_0 + j\Delta t), \quad (3.138)$$

где функция $U_j(S; x_{j,0}^*, t_0 + j\Delta t)$, $\forall S \subset N$, имеет вид (3.106). Согласно пропорциональному решению с динамическим обновлением информации следует распределять кооперативный выигрыш в течение всей игры $\Gamma(x_0, T - t_0)$, используя следующую ПРД:

$$\hat{\beta}_{Prop}(t, \hat{x}^*) = (1 - F_j(t))\beta_j^{Prop}(t, x_j^*), \quad t \in [j\Delta t, (j+1)\Delta t], \quad j = 0, \dots, l. \quad (3.139)$$

Интегрируя $\hat{\beta}_{Prop}(t, \hat{x}^*)$ по t , можно определить пропорциональный дележ $\hat{\xi}_{Prop}(\hat{x}^*(t), T - t)$ (3.111). На рисунках 3.18 и 3.19 видно, что $\hat{\beta}_{Prop}(t, \hat{x}^*)$ находится внутри множества $\hat{B}(t, \hat{x}^*)$. Это означает, что пропорциональное решение строго динамически устойчиво с заданными параметрами.

Теперь предположим, что в некоторый момент времени $t_{br} \in [t_0, T]$ игроки решили, что пропорциональное решение больше их не устраивает, и выбрали другой дележ из решения $\hat{W}(\hat{x}^*(t_{br}), T - t_{br})$, например, вектор Шепли для каждой случайной усеченной подыгры. Для каждой случайной усеченной

подыгры вектор Шепли рассчитывается по формуле

$$Sh_i^j(\hat{x}^*(t_{br}), t_{br}) = \sum_{\substack{S \subset N \\ i \in S}} \frac{(|N| - |S|)! (|S| - 1)!}{|N|!} \cdot \left(V_j(S; t_{br}, \hat{x}^*(t_{br})) - V_j(S \setminus \{i\}; t_{br}, \hat{x}^*(t_{br})) \right). \quad (3.140)$$

Используя вектор Шепли, можно определить ПРД для каждой случайной усеченной подыгры (3.104). Согласно пропорциональному решению с динамическим обновлением информации, игроки должны распределять кооперативный выигрыш в течение всей игры, используя следующую ПРД:

$$\hat{\beta}_{Sh}(t, \hat{x}^*) = (1 - F_j(t)) \beta_j^{Sh}(t, x_j^*), \quad t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t], \quad j = 0, \dots, l,$$

где $\beta_j^{Sh}(t, x_j^*)$ имеет вид (3.105). Отметим, что ПРД для вектора Шепли $\beta_j^{Sh}(t, x_j^*)$ и ПРД для пропорционального решения $\beta_j^{Prop}(t, x_j^*)$ вычисляются таким образом, чтобы гарантировать свойство динамической устойчивости, т.е. по формуле (3.105). Подробное описание пошагового решения ПРД для вектора Шепли представлено в [194].

Пусть игроки решают пересмотреть пропорциональное решение в момент времени $t_{br} = 1.2$. Тогда, согласно (3.110), формула ПРД для всей игры принимает следующий вид:

$$\hat{\beta}(t, \hat{x}^*) = \begin{cases} \hat{\beta}_{Prop}(t, \hat{x}^*), & t \in [t_0, t_{br}], \\ \hat{\beta}_{Sh}(t, \hat{x}^*), & t \in (t_{br}, T]. \end{cases} \quad (3.141)$$

На рисунке 3.17 показаны ПРД $\hat{\beta}_{Prop}(t, \hat{x}^*)$ для пропорционального решения (3.139) (жирная сплошная линия) и ПРД $\hat{\beta}(t, \hat{x}^*)$ для комбинированного решения (3.141) (пунктирная линия).

Посредством прямого интегрирования $\hat{\beta}(t, \hat{x}^*)$ (3.141) по t можно вывести формулу (3.111) для результирующего дележа $\hat{\xi}(\hat{x}^*(t), T - t)$. Согласно $\hat{\xi}(\hat{x}^*(t), T - t)$, игроки распределяют кооперативный выигрыш в игре $\Gamma(x_0, T - t_0)$ следующим образом:

$$\hat{\xi}(\hat{x}^*(t), T - t) = (12.3, 30.2, 16.8). \quad (3.142)$$

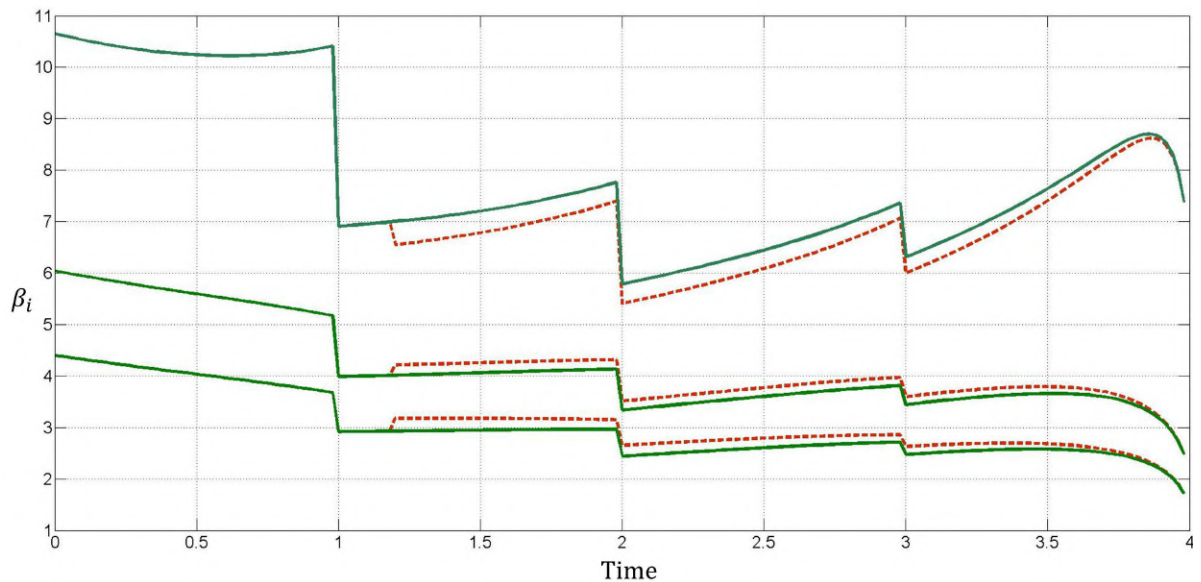


Рис. 3.17: ПРД $\hat{\beta}_{Prop}(t, \hat{x}^*)$ для пропорционального решения (3.138) (жирная сплошная линия) и ПРД $\hat{\beta}(t, \hat{x}^*)$ для комбинированного решения (3.141) (пунктирная линия).

На рисунках 3.18 и 3.19 видно, что $\hat{\beta}(t, \hat{x}^*)$ (3.141) находится внутри множества $\hat{B}(t, \hat{x}^*)$ (3.107). Это означает, что соответствующий дележ $\hat{\xi}(\hat{x}^*(t), T - t) \in \hat{W}(\hat{x}^*(t), T - t)$ при заданных параметрах. Этот факт демонстрирует свойство сильной динамической устойчивости концепции решения $\hat{W}(x_0, T - t_0)$. Кроме того, на рисунках 3.18 и 3.19 видно, что пропорциональное решение $\hat{\beta}_{Prop}(t, \hat{x}^*)$ (3.139) находится внутри множества $\hat{B}(t, \hat{x}^*)$.

На рисунке 3.20 представлена разность между $\hat{\xi}(\hat{x}^*(t), T - t)$ и $\hat{\xi}_{Prop}(\hat{x}^*(t), T - t)$.

3.4 Дифференциальная игра нефтяного рынка с динамическим обновлением информации

3.4.1 Введение

Этот раздел посвящен построению теоретико-игровой модели мирового рынка нефти на основе подхода с динамическим обновлением информации. Игры с динамическим обновлением информации позволяют учитывать изменчивость рыночных требований, адаптацию действий участников рынка к изменяющейся

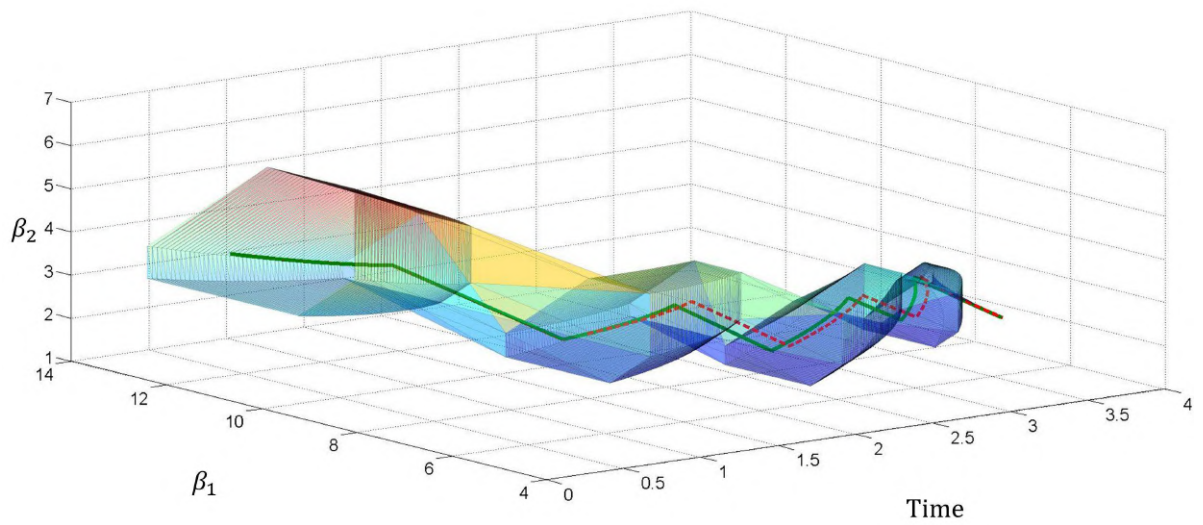


Рис. 3.18: Оси: β_1 , β_2 , t . β_3 можно вычислить по (3.107).

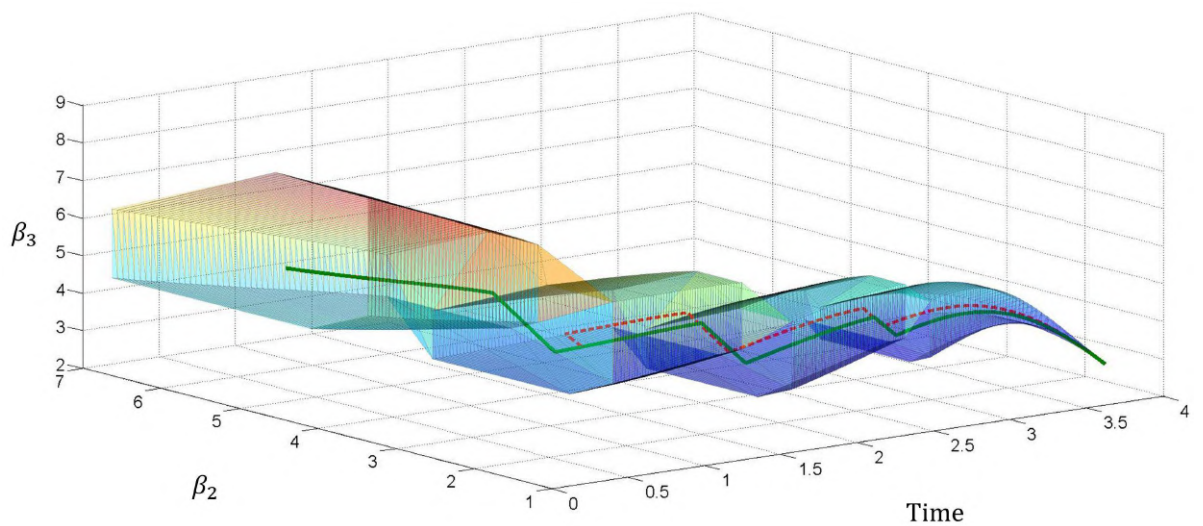


Рис. 3.19: Оси: β_2 , β_3 , t . β_1 можно вычислить по (3.107).

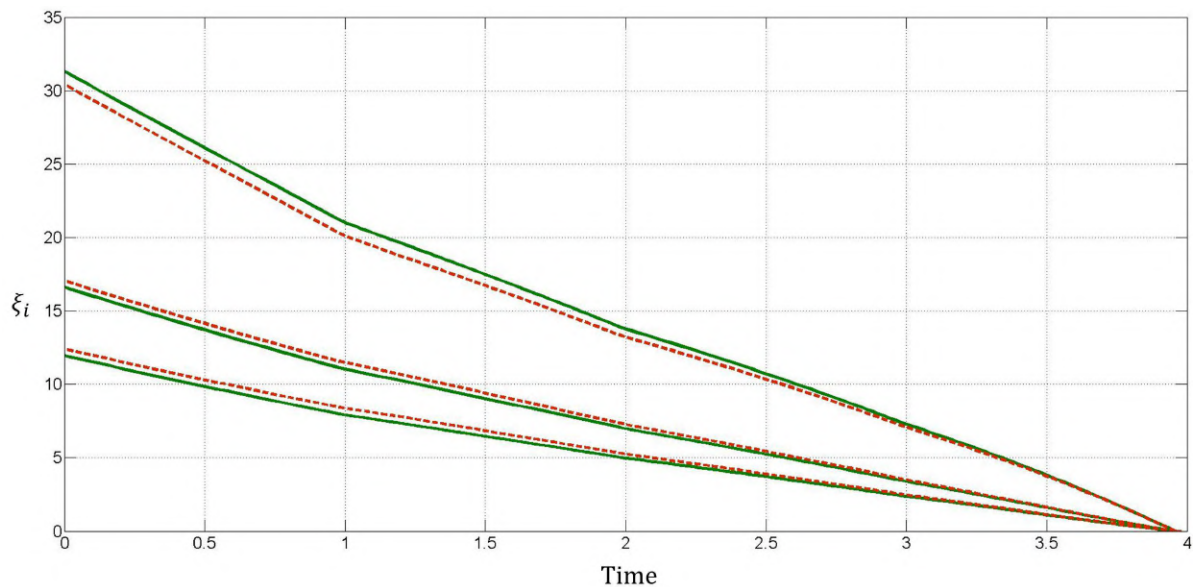


Рис. 3.20: Дележ $\hat{\xi}_{Prop}(\hat{x}^*(t), T - t)$ для пропорционального решения (толстая сплошная линия) и дележ $\hat{\xi}(\hat{x}^*(t), T - t)$ для комбинированного решения (пунктирная линия).

среде, а также реальные горизонты планирования спроса. Рынок нефти отличается высокой волатильностью цен, поэтому к рынку нефти применяется метод динамического обновления информации и исследуется полученная модель. Предполагается, что страны не имеют или не могут использовать долгосрочные прогнозы для параметров нефтяного рынка. Следовательно, их поведение можно смоделировать с помощью предложенного подхода.

Целью этого раздела является моделирование динамики рынка нефти в течение двух конкретных периодов, с декабря 2015 г. по ноябрь 2016 г. и с декабря 2016 г. по май 2017 г. Страны ОПЕК и одиннадцать стран, не входящих в ОПЕК, достигли шестимесячного соглашения о сокращении добычи нефти на второй указанный период на саммите в Вене 30 ноября 2016 г. Соглашение было направлено на рост цен на нефть. В совокупности на эти страны приходится более 60% мировой добычи нефти. Чтобы показать эффект от соглашения, рассмотрим две теоретико-игровые модели, соответствующие двум упомянутым периодам. Предполагается, что страны не сотрудничали официально в отношении объемов нефти, которые должны быть добыты до 30 ноября 2016 г. Следовательно, их поведение можно смоделировать с помощью некооперативной игры. Во-первых, строится некооперативная игра для временного интервала с декабря 2015 г. по ноябрь 2016 г. и адаптируется к реальным данным цен на

нефть. В качестве принципа оптимальности используется равновесие по Нэшу в позиционных стратегиях. Достижение соглашения об объемах добычи нефти — пример чисто кооперативного поведения. Таким образом, коалиционная игра строится для временного интервала с декабря 2016 г. по май 2017 г. Кооперация моделируется с помощью игры с нетрансферабельной полезностью. Далее для этого типа игр строится классическое кооперативное решение, т.е. состоятельное во времени решение, предложенное в [165]. Кроме того, моделируется соглашение о сотрудничестве (историческое решение), которое фактически использовалось странами. Согласно историческому соглашению, решение основано на объемах добычи нефти в ноябре 2016 г. Оно предписывает игрокам 1 и 2 линейно снизить объемы добычи нефти до ранее согласованного уровня. Очевидно, что второе решение не является временно состоятельным, хотя позволяет нам моделировать последнее реальное рыночное соглашение.

Мы представляем результаты статистического моделирования и сравнения исторических данных и данных модели за оба периода. В частности, теоретическая траектория цены на нефть за период с декабря 2015 г. по ноябрь 2016 г. сравнивается со статистическими данными. За период с декабря 2016 г. по май 2017 г. сравниваются траектория цен на нефть, соответствующая статистическим данным, состоятельное во времени кооперативное решение, определенное в [165], и кооперативное решение на основе соглашения, подписанного крупнейшими экспортерами нефти в конце 2016 г. Для определения параметров модели используются данные из открытых источников о мировом рынке нефти: ежемесячные данные о поставках сырой нефти с января 2015 г. по май 2017 г. (Международное энергетическое агентство), ежемесячные данные о ценах на нефть марки Brent и светлую нефть с января 2015 г. по май 2017 г. (агентство «Финам»), а также оценки стоимости добычи барреля нефти в 2016 г. (Rystad Energy Ucube и новости нефтяного рынка).

Крупнейшие экспортеры нефти достигли договоренности о сокращении добычи нефти с целью повышения цен после 30 ноября 2016 г., и во всех наших моделях экспортеры рассматриваются в группах от пяти игроков до одного игрока. Это не первый случай, когда ОПЕК ограничивает добычу нефти в надежде контролировать цены на нефть. Такие действия членов организации представляются в определенном смысле рациональными, поскольку в

2014 г. цены на нефть начали неуклонно падать. Таким образом, в наших игровых моделях все члены ОПЕК вместе называются игроком 1. Остальные одиннадцать стран, подписавших соглашение и у которых явно есть общие интересы, хотя и с большей свободой в принятии решений по сравнению со странами ОПЕК, называются игроком 2. В то же время устойчивое падение цен на нефть было в значительной степени связано с революцией в сфере сланцевой нефти и газа в США, которая началась в 2012 г. Поэтому в наши модели в качестве основных конкурентов на рынке включены компании, добывающие сланцевую и несланцевую нефть (соответственно игроки 3 и 4). Все остальные страны-экспортеры нефти составляют игрока 5. Страны объединяются в игроков, от одного до пяти, по описанному правилу, так как в наших моделях учитывается только процесс изменения цены. Кроме того, такая комбинация значительно упрощает вычислительный процесс, не оказывая серьезного влияния на достоверность результатов.

Мировой рынок нефти является одним из наиболее значимых рынков и важнейшим компонентом темпов роста и бюджетных сценариев для некоторых ресурсозависимых стран. Справедливости ради надо отметить, что это подробно освещено в литературе. Большинство существующих моделей рассматривают рынок нефти как рынок с несовершенной конкуренцией. Модели рыночной конкуренции для нефтяного рынка объясняют присутствие ОПЕК картельным поведением, доминирующей фирмой или, в большинстве случаев, целевым поведением. Например, Даль и Юсель [201] описали поведение ОПЕК как слабую координацию или дуополию. В статье [202] показано, что модель добычи на рынке нефти несовместима с моделями теории конкуренции. Даниэльсен и Ким [203] представили важные свидетельства сотрудничества между странами ОПЕК. Смит [204] утверждал, что ОПЕК — это гораздо больше, чем просто некооперативная олигополия, но меньше, чем картель «без трений». Согласно Юханна [205], модель картеля с частичным разделением рынка доминирует над всеми другими моделями. В работе [206] Гюлен привел доказательства координации объемов производства и предположил, что ОПЕК действовала как картель в период 1982–1993 гг. Автор [207] считал, что поведение ОПЕК согласуется с теорией картеля. Как отмечено Бокемом [208], рынок сырой нефти лучше всего можно описать с помощью модели ценового лидера, в которой ОПЕК является лидером, а все остальные страны (не входящие в ОПЕК)

— пассивными участниками рынка. В статье [209] утверждалось, что ОПЕК ведет себя скорее как олигополия, где Саудовская Аравия является лидером цен и крупнейшим производителем. Согласно статье [210], некооперативная модель Нэша–Курно потенциально может объяснить рынок нефти лучше, чем конкурентная модель. Моран [211] и Краснер [212] проанализировали основные черты нефтяной олигополии. Авторы [213] представили выборочный обзор моделей олигополии для производства энергии. В статье [214] исследовано формирование картелей на мировом рынке нефти в постановке Курно. Авторы [215] также использовали количественную конкурентную среду для моделирования столкновений и пропорционального регулирования уровней производства. Следуя этой устоявшейся традиции, для моделирования рынка нефти будем использовать количественную оценку олигополии.

В первой части этого раздела подход с динамическим обновлением информации применяется к дифференциальной модели некооперативной олигополии [216] нефтяного рынка с крупнейшими экспортерами нефти и другими нефтедобывающими странами. Во второй части раздела рассматривается частично кооперативная дифференциальная игра с нетрансферабельной полезностью, в которой в качестве исходной модели используется описанная выше некооперативная игра. Предполагается, что первый и второй игроки сотрудничают, а остальные используют против них равновесные по Нэшу стратегии.

3.4.2 Модель некооперативной игры

3.4.2.1 Исходная игра

Рассмотрим дифференциальную игровую модель олигополии Курно [216] на рынке нефти. Олигополистический рынок n асимметричных стран (игроков) из множества $N = \{1, \dots, n\}$, добывающих нефть и конкурирующих за объем производства q_i при условии жестких цен, описывается дифференциальной игрой $\Gamma(p_0, T - t_0)$ с заданной продолжительностью $T - t_0$ и начальным состоянием $p(t_0) = p_0 \in P \subset R$.

В рамках модели, динамика рыночной цены p_i удовлетворяет дифференци-

альному уравнению

$$\dot{p}(t) = s(\hat{p}(t) - p(t)), \quad p(t_0) = p_0, \quad (3.143)$$

где $\hat{p}(t) \in P \subset R$ — условная цена в момент времени t , $p(t)$ — текущая цена, а параметр $s : 0 < s < 1$ — скорость корректирования. Таким образом, цена корректируется в зависимости от разности между ее условным и текущим значениями.

Далее предполагается, что условные цены в любой момент времени t определяются линейной обратной функцией спроса

$$\hat{p}(t) = a - d \sum_{i \in N} q_i(t). \quad (3.144)$$

Каждый игрок $i \in N$ выбирает объем производства $q_i(t) \in U_i \subset R$, чтобы максимизировать свою прибыль:

$$K_i(p_0, T - t_0; q_1, \dots, q_n) = \int_{t_0}^T e^{-\rho(t-t_0)} [q_i(t)(p(t) - c_i - g_i q_i(t))] dt, \quad (3.145)$$

где $0 \leq \rho \leq 1$ — положительная ставка дисконтирования. Будем считать, что она одинакова для всех временных периодов и для всех игроков, чтобы упростить модель и уравнять игроков как симметричных участников глобального рынка капитала. Функция общих затрат для каждого игрока i имеет вид $C_i(t) = c_i q_i(t) + g_i q_i^2(t)$.

3.4.2.2 Усеченная подыгра

Предположим, что информация для игроков обновляется в фиксированные моменты времени $t = t_0 + j\Delta t$, $j = 0, \dots, l$, где $l = \frac{T}{\Delta t} - 1$. В течение временного интервала $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$ игроки имеют полную информацию о динамике (уравнениях движения) игры (3.143) и функции выигрыша (3.147) на временном интервале времени $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, где $\Delta t \leq \bar{T} \leq T$. В момент времени $t = t_0 + (j+1)\Delta t$ информация об игре обновляется, и такая же процедура повторяется для временного интервала с номером $j+1$.

Для моделирования такого поведения введем следующее определение, где векторы $p_{j,0} = p(t_0 + j\Delta t)$ и $p_{j,1} = p(t_0 + (j+1)\Delta t)$.

Определение 3.4.0.1 Пусть $j = 0, \dots, l$. Усеченная подыгра $\bar{\Gamma}_j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ определена на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$. Уравнения движения и начальное условие в усеченной подыгре $\bar{\Gamma}_j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ имеют вид

$$\dot{p}(t) = s \left(a_j - d_j \sum_{i \in N} q_i^j(t) - p(t) \right), \quad p(t_0 + j\Delta t) = p_{j,0}. \quad (3.146)$$

Функция выигрыша игрока i в усеченной подыгре j равна

$$\begin{aligned} K_i^j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^j, \dots, q_n^j) &= \\ &= \int_{t_0 + j\Delta t}^{t_0 + j\Delta t + \bar{T}} e^{-\rho(t-t_0)} \left[q_i^j(t)(p(t) - c_i - g_i q_i^j(t)) \right] dt. \end{aligned} \quad (3.147)$$

Уравнения движения и функция выигрыша усеченной подыгры на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ совпадают с уравнениями движения и функцией выигрыша исходной игры $\Gamma(x_0, T - t_0)$ на том же временном интервале.

3.4.2.3 Некооперативный исход усеченной подыгры

Согласно [1; 217], некооперативное равновесное по Нэшу решение игры $\bar{\Gamma}_j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ можно определить с помощью уравнений Гамильтона–Якоби–Беллмана в частных производных. Рассмотрим семейство подыгр $\bar{\Gamma}_j(p(t), t, t_0 + j\Delta t + \bar{T})$ со структурой выигрышей (3.145) и динамикой (3.146), начинающихся в момент времени $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ из начального состояния $p(t)$. При $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ обозначим через $q_j^{NE}(t, p) = (q_1^{jNE}(t, p), \dots, q_n^{jNE}(t, p))$ набор позиционных стратегий, образующих равновесное по Нэшу решение для игры $\bar{\Gamma}_j(p(t), t, t_0 + j\Delta t + \bar{T})$, а через $V_i^j(\tau, p) : [t, T] \times R^n \rightarrow R$ — функцию значения игрока $i \in N$, которая удовлетворяет соответствующим уравнениям Гамильтона–Якоби–Беллмана [1; 217].

Теорема 3.4.1 Предположим, что существует непрерывно дифференцируе-

мая функция $V_i^j(t, p) : [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}] \times R \rightarrow R$, удовлетворяющая дифференциальному уравнению в частных производных

$$-V_t^{j,i}(t, p) = \max_{q_i^j} \left\{ e^{-\rho(t-t_0)} \left[q_i^j(p - c_i - g_i q_i^j) \right] + \right. \\ \left. + V_p^{j,i}(t, p) s \left(a - d \left[q_i^j(t) + \sum_{k \neq i} q_k^{jNE}(t) \right] - p(t) \right) \right\}, \quad i = 1, \dots, n. \quad (3.148)$$

где $V_i^j(t_0 + j\Delta t + \bar{T}, p) = 0$. Пусть управления $q_j^{NE}(t)(t, p)$ доставляют максимум правой части уравнения (3.148). Тогда $q_j^{NE}(t)(t, p)$ является равновесием по Нэшу в позиционных стратегиях в усеченной подыгре $\bar{\Gamma}_j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$.

Поскольку рассматриваемая дифференциальная игра является линейно-квадратичной, равновесие по Нэшу в позиционных стратегиях единственное [217].

В этой игре функцию Беллмана $V_i^j(t, p)$ можно получить в следующем виде:

$$V_i^j(t, p) = e^{-\rho(t-t_0)} \left[A_i^j(t)p^2 + B_i^j(t)p + C_i^j(t) \right], \quad i = \overline{1, n}. \quad (3.149)$$

При подстановке выражения (3.149) в уравнение (3.148) получаем равновесные по Нэшу позиционные стратегии

$$q_i^{jNE}(t, p) = -\frac{(c_i - p) + sd \left[B_i^j(t) + 2A_i^j(t)p \right]}{2g_i}, \quad i = \overline{1, n}, \quad (3.150)$$

где функции $A_i^j(t)$, $B_i^j(t)$ и $C_i^j(t)$, $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, находятся из

системы дифференциальных уравнений

$$\begin{aligned}
\dot{A}_i^j(t) &= A_i^j(t) [\rho + 2s] + \frac{(2A_i^j(t)d_j s - 1)^2}{4g_i} - \sum_{k \neq i} \frac{A_i^j(t)d_j s - 2A_i^j(t)A_k^j(t)d^2 s^2}{g_k} \\
\dot{B}_i^j(t) &= B_i^j(t) [\rho + s] - \frac{c_i}{2g_i} - 2A_i^j(t)a_j s - \sum_{k \neq i} \frac{A_i^j(t)B_k^j(t)d_j^2 s^2}{g_k} - \\
&\quad - \sum_{k \in N} \frac{B_i^j(t)d_j s - A_i^j(t)c_k d_j s - A_k^j(t)B_i^j(t)d_j^2 s^2}{g_k} \\
\dot{C}_i^j(t) &= C_i^j(t)\rho - B_i^j(t)a_j s + \frac{c_i^2 + (B_i^j(t)d_j s)^2}{4g_i} + \sum_{k \neq i} \frac{B_i^j(t)B_k^j(t)d_j^2 s^2}{2g_k} + \\
&\quad + \sum_{k \in N} \frac{B_i^j(t)c_k d_j s}{2g_k}
\end{aligned}$$

соответственно с граничными условиями $A_i^j(t_0 + j\Delta t + \bar{T}) = 0$, $B_i^j(t_0 + j\Delta t + \bar{T}) = 0$ и $C_i^j(t_0 + j\Delta t + \bar{T}) = 0$.

Подставляя функцию $q_j^{NE}(t, p)$ (3.150) в систему (3.146), получаем динамику равновесной по Нэшу траектории в виде

$$\dot{p}(t) = s \left(a - d \sum_{i \in N} q_i^{jNE}(t, p) - p(t) \right), \quad p(t_0 + j\Delta t) = p_{j,0}. \quad (3.151)$$

Обозначим через $p_j^{NE}(t)$ решение дифференциального уравнения (3.151).

3.4.2.4 Условно-некооперативная траектория

Предположим, что каждая усеченная подыгра $\bar{\Gamma}_j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ развивается вдоль траектории $p_j^{NE}(t)$. Тогда некооперативная игра с динамическим обновлением информации развивается вдоль следующей траектории.

Определение 3.4.0.2 *Условно-некооперативная траектория $\{\hat{p}_{NE}(t)\}_{t=t_0}^T$ — это комбинация траекторий $p_j^{NE}(t)$ для каждой усеченной подыгры $\bar{\Gamma}_j(p_{j,0}^{NE}, t_0 +$*

$j\Delta t, t_0 + j\Delta t + \bar{T})$:

$$\{\hat{p}_{NE}^*(t)\}_{t=t_0}^T = \begin{cases} p_0^{NE}(t), & t \in [t_0, t_0 + \Delta t), \\ \dots, \\ p_j^{NE}(t), & t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t), \\ \dots, \\ p_l^{NE}(t), & t \in [t_0 + l\Delta t, t_0 + (l+1)\Delta t]. \end{cases} \quad (3.152)$$

Вдоль условно-некооперативной траектории игроки получают выигрыш по следующей формуле:

Определение 3.4.0.3 *Результирующий некооперативный исход для игрока $i = 1, \dots, n$ в подыгре игры $\Gamma(p_0, T - t_0)$ с динамическим обновлением информации, начинающейся в момент времени $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, имеет следующий вид:*

$$\hat{V}_i(t, \hat{p}_{NE}(t)) = \sum_{m=j+1}^l \left[V_i^m(t_0 + m\Delta t, p_{m,0}^{NE}) - V_i^m(t_0 + (m+1)\Delta t, p_{m,1}^{NE}) \right] + \left[V_i^j(t, p_j^{NE}(t)) - V_i^j(t_0 + (j+1)\Delta t, p_{j,1}^{NE}) \right], \quad i \in N. \quad (3.153)$$

3.4.2.5 Численное моделирование

Первая игра начинается в декабре 2015 г. и заканчивается до саммита в Вене в ноябре 2016 г. Нефть рассматривается как однородный продукт. Оценивается функция спроса с параметрами средней мировой цены на нефть и совокупного мирового предложения нефти. Средние цены на нефть рассчитываются для каждого периода на основе только двух основных торговых классификаций, а именно, нефти марки Brent и легкой нефти (источник данных — агентство «Финам»). В качестве начальной цены принята средняя цена на декабрь 2015 г., равная $p_0 = 34,51$.

Период	Общее мировое предложение, млн баррелей в сутки	Brent, \$ за баррель	Light, \$ за баррель	Средняя цена, \$ за баррель
12.2015	96,411	35,910	33,110	34,510
01.2016	95,875	36,640	33,990	35,315
02.2016	95,420	40,140	37,820	38,980
03.2016	95,294	47,320	45,990	46,655
04.2016	95,400	49,520	48,750	49,135
05.2016	95,187	49,740	48,640	49,190
06.2016	95,954	43,270	41,760	42,515
07.2016	96,891	46,970	45,000	45,985
08.2016	95,894	49,990	48,050	49,020
09.2016	96,001	48,510	46,970	47,740
10.2016	97,362	44,520	43,120	43,820
11.2016	97,241	47,034	45,908	46,471

Чтобы гарантировать, что наклон спроса будет отрицательным, принято решение не регрессировать исторические цены на нефть для произведенного количества и принять во внимание среднюю цену и общее мировое предложение (см. таблицу выше) только за один предыдущий период. Таким образом, предполагается, что параметр спроса a_j равен 300 в каждый период. Следовательно, параметр d_j можно найти из (3.144):

$$d_j = (a_j - \hat{p}(t-1)) \sum_{i \in N} q_i(t-1). \quad (3.154)$$

Длина каждого временного интервала Δt составляет один месяц. Игроки используют оцененный спрос с параметрами a_j и d_j в качестве прогноза на следующие $\bar{T} = 3$ периода. Значения устанавливаются на основе параметров функции затрат с использованием общей стоимости добычи барреля и средних объемов добычи нефти для игроков в 2016 г., путем фиксации параметра g_i на уровне 0,7 для каждого игрока и каждого периода. Оба параметра c_i

и g_i остаются неизменными во время игры. Предполагается, что скорость корректировки и коэффициент дисконтирования равны соответственно $s = 0,2$ и $r = 10\%$.

Период	a	d
12.2015	300	2,717
01.2016	300	2,754
02.2016	300	2,761
03.2016	300	2,735
04.2016	300	2,659
05.2016	300	2,630
06.2016	300	2,635
07.2016	300	2,683
08.2016	300	2,622
09.2016	300	2,617
10.2016	300	2,628
11.2016	300	2,631

i	Производитель	c	g
1	Члены ОПЕК	3,169	0,7
2	Страны, не входящие в ОПЕК	17,333	0,7
3	США, сланцевая нефть	20,238	0,7
4	США, несланцевая нефть	18,182	0,7
5	Другие	20,867	0,7

На рисунке 2 представлено сравнение условно-некооперативной траектории $\{\hat{p}_{NE}(t)\}_{t=t_0}^T$ и исторической динамики средней цены на нефть. На рисунке 3 позиционные равновесные по Нэшу стратегии (3.150), соответствующие $\{\hat{p}_{NE}(t)\}_{t=t_0}^T$ и исторические объемы добычи нефти показаны по каждой группе стран.

Согласно рисункам 2 и 3, параметры c_i и g_i из приведенной ниже таблицы можно использовать для моделирования кооперативного соглашения на временном интервале с декабря 2016 г. по май 2017 г.

На рисунке 4 показан выигрыш игроков $\hat{V}_i(t, \hat{p}_{NE}(t))$, $i \in N$, соответствующий равновесным по Нэшу стратегиям вдоль траектории $\hat{p}_{NE}(t)$.

3.4.3 Частично кооперативная игра

Как упоминалось во введении, в конце ноября 2016 г. две группы стран, члены ОПЕК и страны, не входящие в ОПЕК (игроки 1 и 2), подписали соглашение о

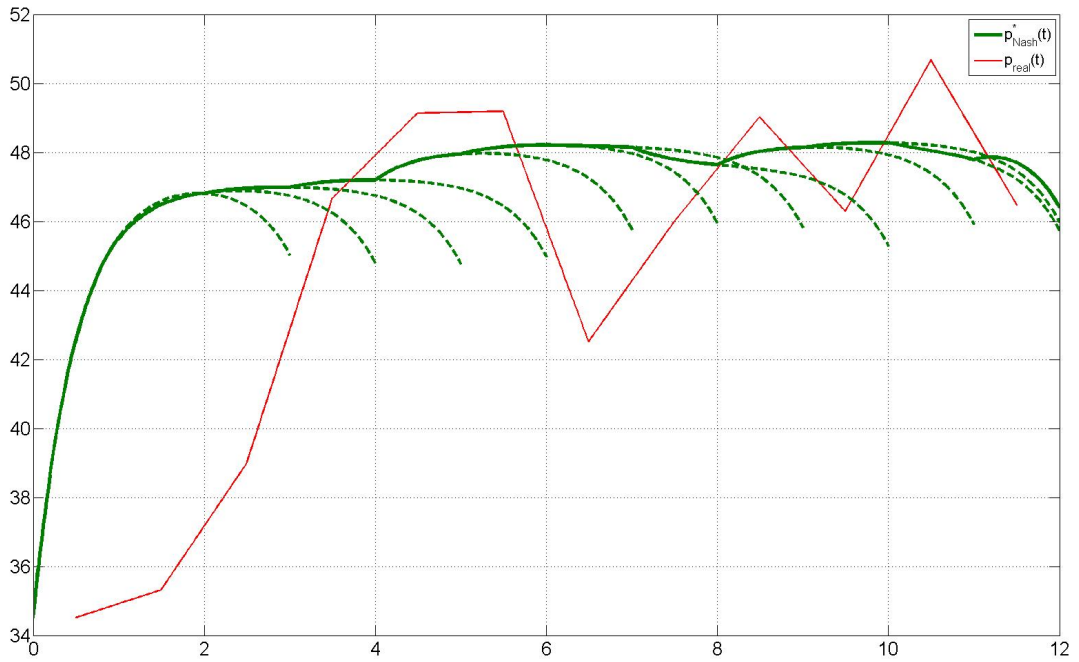


Рис. 3.21: Условно-некооперативная траектория цены на нефть $\hat{p}_{NE}(t)$ (толстая сплошная линия) с динамическим обновлением информации и историческая траектория цен на нефть (толстая штриховая линия).

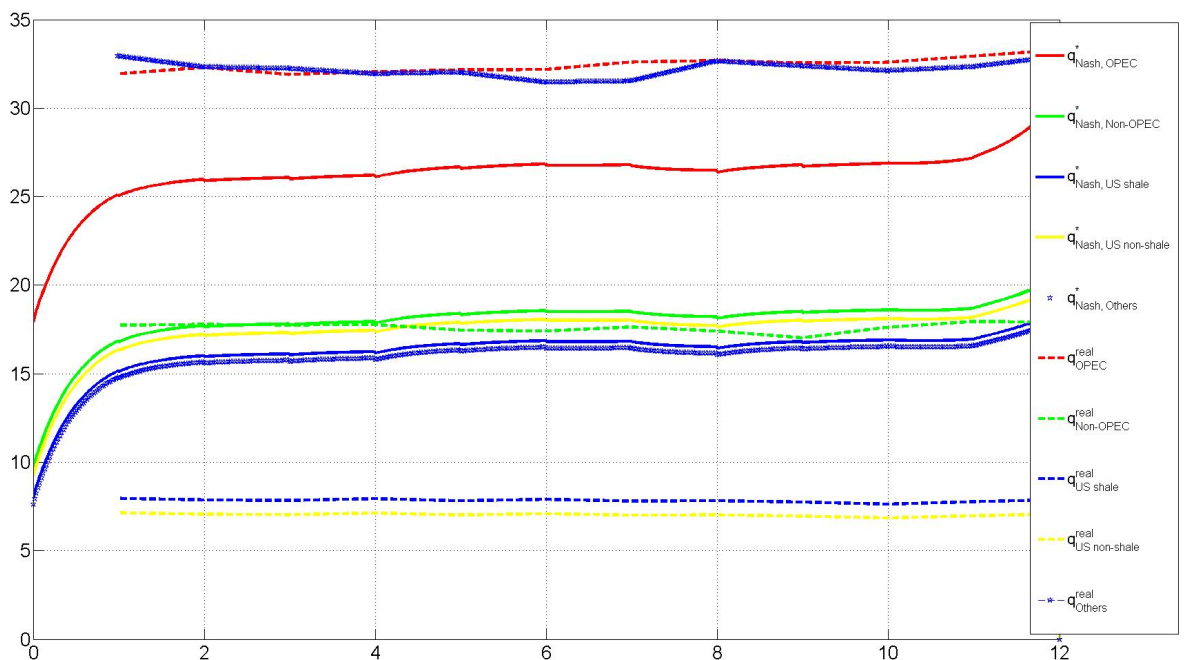


Рис. 3.22: Позиционные равновесные по Нэшу стратегии с динамическим обновлением информации (сплошные линии) и соответствующие исторические объемы добычи нефти (штриховые линии).

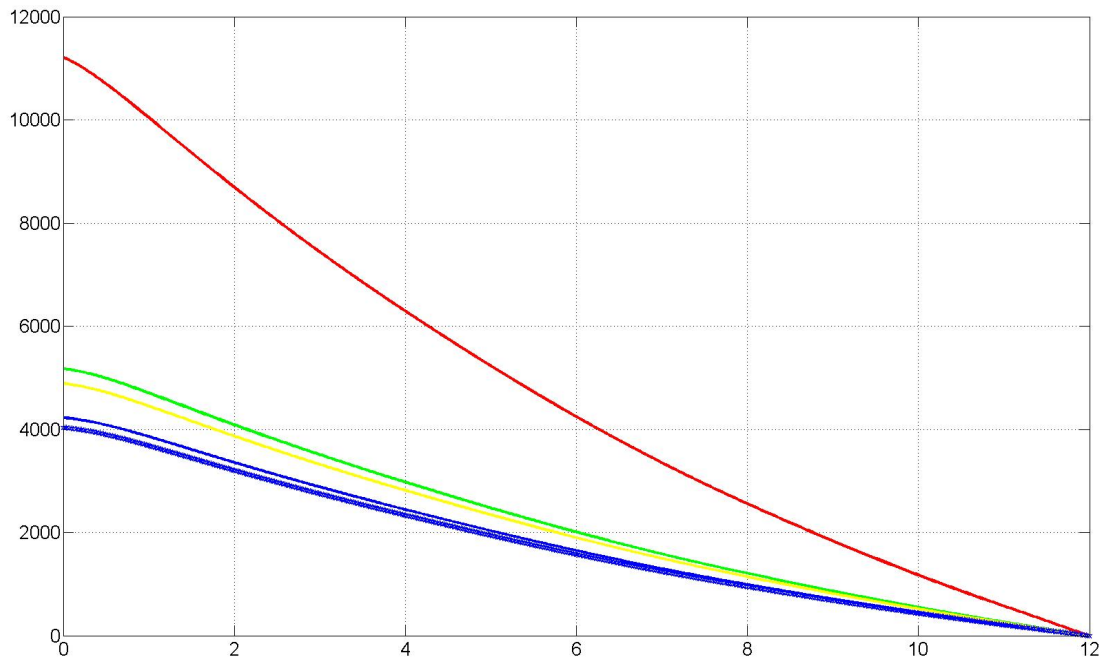


Рис. 3.23: Выигрыши игроков $\hat{V}_i(t, \hat{p}_{NE}(t))$ в равновесии по Нэшу.

сотрудничестве. Используя модель из раздела 1, такую частичную кооперацию будем описывать игрой 4-х лиц следующего вида: игроки 1 и 2 действуют как один игрок, а остальные игроки (игроки 3, 4 и 5) действуют самостоятельно. В качестве исторического кооперативного решения используется линейное сокращение производимых объемов до согласованного уровня. Кроме того, строится и кооперативное решение, обладающее свойством временной состоятельности, впервые представленного в статье [165]. При численном моделировании ниже обсуждается разница между историческим решением и решением [165].

3.4.3.1 Частично-кооперативная усеченная подыгра

Чтобы смоделировать поведение всех игроков, при условии, что игроки 1 и 2 взаимодействуют друг с другом, а остальные действуют независимо, сначала построим дифференциальную игру 4-х лиц. Множество игроков в этой игре обозначим через $\tilde{N} = \{(1, 2), 3, 4, 5\}$, $|\tilde{N}| = 4$. Игроки 1 и 2 сотрудничают, комбинируя наборы своих стратегий, но фактические выигрыши остаются индивидуальными. Таким образом, для игроков 1 и 2 будем использовать аналог кооперативной дифференциальной игры с нетрансферабельными выигрышами. Игроки 1 и 2 ориентируются на выигрыш в виде линейной комбинации их

индивидуальных выигрышей с фиксированными коэффициентами α_j :

$$\begin{aligned} \tilde{K}_{1,2}^j(\alpha_j, p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^j, \dots, q_5^j) = \\ = \sum_{i \in \{1,2\}} \alpha_i^j K_i^j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^j, \dots, q_5^j), \end{aligned} \quad (3.155)$$

где $K_i^j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^j, \dots, q_5^j)$ определены в (3.145) и

$$\alpha_j \in \Lambda_j = \{\alpha : \alpha_1^j + \alpha_2^j = 1, \alpha_i^j \geq 0, i = 1, 2\}. \quad (3.156)$$

Индивидуальные выигрыши $\tilde{K}_i^j(\alpha_j, p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^j, \dots, q_5^j)$, $i = 1, 2$, которые игроки получают во время игры, вычисляются по формуле (3.145). Выигрыши игроков $i = 3, 4, 5$ также находятся по формуле (3.145).

Оказывается, для любого $\alpha_j \in \Lambda_j$ получается единственная дифференциальная игра.

Определение 3.4.0.4 Пусть $j = 0, \dots, l$. Частично-кооперативная усеченная подыгра $\tilde{\Gamma}_j^{\alpha_j}(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ определена на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$. Уравнения движения и начальное условие усеченной подыгры $\tilde{\Gamma}_j^{\alpha_j}(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ имеют следующий вид:

$$\dot{p}(t) = s \left(a_j - d_j \sum_{i \in N} q_i^j(t) - p(t) \right), \quad p(t_0 + j\Delta t) = p_{j,0}. \quad (3.157)$$

Функция выигрыша $\tilde{K}_{1,2}^j(\alpha_j, p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^j, \dots, q_5^j)$ игроков $\{1, 2\}$ в усеченной подыгре j определяется выражением (3.155), а функция выигрыша $\tilde{K}_i^j(\alpha_j, p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^j, \dots, q_5^j)$ игроков $i = 3, 4, 5$ — выражением (3.145).

3.4.3.2 Парето-оптимальные исходы

Предположим, что игроки 1 и 2 могут договориться о форме совместной функции выигрыша, т.е. о коэффициентах $\alpha_j \in \Lambda_j$. Множество исходов при всевозможных $\alpha_j \in \Lambda_j$ для игроков 1 и 2 определяет множество Парето-

оптимальных исходов для игроков 1 и 2. Рассмотрим равновесие по Нэшу в позиционных стратегиях $q_{\alpha_j}^*(t, p)$ в игре 4-х лиц $\tilde{\Gamma}_j^{\alpha_j}(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$. Тогда выигрыши игроков 1 и 2 в равновесии по Нэшу при всевозможных α_j образуют множество Парето-оптимальных исходов $\tilde{K}_1^j(\alpha_j, p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^{*\alpha_j}, \dots, q_5^{*\alpha_j})$ и $\tilde{K}_2^j(\alpha_j, p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^{*\alpha_j}, \dots, q_5^{*\alpha_j})$. Позиционные равновесные по Нэшу стратегии игроков определены как в [217].

Пусть выигрыши игроков $\{1, 2\}$ в равновесии по Нэшу в позиционных стратегиях с выбранным значением α_j описывается функцией $W_{1,2}^{\alpha_j}(t, p)$, где t и p — соответственно время и начальное состояние частично-кооперативной усеченной подыгры $\tilde{\Gamma}_j^{\alpha_j}(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$. По аналогии, обозначим через $W_i^{\alpha_j}(t, p)$, $i = 3, 4, 5$, функцию выигрыша игроков $i = 3, \dots, 5$ при использовании равновесных по Нэшу стратегий $q_{\alpha_j}^*(t, p)$. Достаточные условия существования решения и равновесия по Нэшу в позиционных стратегиях даются следующим утверждением.

Теорема 3.4.2 *Предположим, что существуют непрерывно дифференцируемые функции $W_{1,2}^{\alpha_j}(t, p)$, $W_3^j(t, p)$, $W_4^j(t, p)$, $W_5^j(t, p)$: $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}] \times R \rightarrow R$, удовлетворяющие следующей системе уравнений в частных производных:*

$$\begin{aligned} -W_t^{1,2, \alpha_j}(t, p) &= \max_{q_1^j, q_2^j} \left\{ \sum_{i=1}^2 e^{-\rho(t-t_0)} \alpha_i^j \left[q_i^j (p - c_i - g_i q_i^j) \right] + \right. \\ &+ W_p^{1,2, \alpha_j}(t, p) s \left(a - d \left[\sum_{i=1}^2 q_i^j(t) + \sum_{k=3}^5 q_k^{*\alpha_j}(t) \right] - p(t) \right) \left. \right\}, \\ -W_t^{i, \alpha_j}(t, p) &= \max_{q_i^j} \left\{ e^{-\rho(t-t_0)} \left[q_i^j (p - c_i - g_i q_i^j) \right] + \right. \\ &+ W_p^{i, \alpha_j}(t, p) s \left(a - d \left[q_i^j(t) + \sum_{k \neq i}^5 q_k^{*\alpha_j}(t) \right] - p(t) \right) \left. \right\}, \quad i = 3, \dots, 5, \quad (3.158) \end{aligned}$$

где $W_{1,2}^{\alpha_j}(t_0 + j\Delta t + \bar{T}, p) = 0$ и $W_i^{\alpha_j}(t_0 + j\Delta t + \bar{T}, p) = 0$, $i = 3, \dots, 5$. Тогда стратегии $q_{\alpha_j}^*(t, p)$ образуют равновесие по Нэшу в позиционных стратегиях в частично-кооперативной усеченной подыгре $\tilde{\Gamma}_j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$.

В этой игре функции Беллмана $W_{1,2}^{\alpha_j}(t, p)$ и $W_i^{\alpha_j}(t, p)$, $i = 3, \dots, 5$, можно

получить в виде

$$\begin{aligned} W_{1,2}^{\alpha_j}(t, p) &= e^{-\rho(t-t_0)} \left[A_j^{1,2}(t)p^2 + B_j^{1,2}(t)p + C_j^{1,2}(t) \right], \\ W_i^{\alpha_j}(t, p) &= e^{-\rho(t-t_0)} \left[A_j^i(t)p^2 + B_j^i(t)p + C_j^i(t) \right], \quad i = 3, 4, 5. \end{aligned} \quad (3.159)$$

Позиционные равновесные по Нэшу стратегии для частично-кооперативной усеченной подыгры равны

$$\begin{aligned} q_i^{*\alpha_j}(t, p) &= -\frac{(c_i - p) + ds \left[B_j^{1,2}(t) + 2A_j^{1,2}(t)p \right]}{2\alpha_j g_i}, \quad i = 1, 2, \\ q_i^{*\alpha_j}(t, p) &= -\frac{(c_i - p) + sd \left[B_j^i(t) + 2A_j^i(t)p \right]}{2g_i}, \quad i = 3, 4, 5. \end{aligned} \quad (3.160)$$

Функции $A_j^{1,2}(t)$, $B_j^{1,2}(t)$, $C_j^{1,2}(t)$, $A_j^i(t)$, $B_j^i(t)$ и $C_j^i(t)$, $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, находятся из систем дифференциальных уравнений, представленных ниже.

Для функций $A_j^{1,2}(t_0 + j\Delta t + \bar{T})$, $B_j^{1,2}(t_0 + j\Delta t + \bar{T})$ и $C_j^{1,2}(t_0 + j\Delta t + \bar{T})$:

$$\begin{aligned} \dot{A}_j^{1,2}(t) &= A_j^{1,2}(t)(\rho + 2s) + \sum_{k \in \{1,2\}} \left(\frac{\alpha_k^j}{4g_k} + \frac{(A_j^{1,2}(t)d_j s)^2}{\alpha_k^j g_k} \right) + \\ &+ \sum_{k \in \{3,4,5\}} \frac{2A_j^k(t)A_j^{1,2}(t)d_j^2 s^2}{g_k} - \sum_{k \in N} \frac{A_j^{1,2}(t)d_j s}{g_k}, \\ \dot{B}_j^{1,2}(t) &= B_j^{1,2}(t)(\rho + s) - 2A_j^{1,2}(t)a_j s - \sum_{k \in \{1,2\}} \left(\frac{\alpha_k^k c_k}{2g_k} - \frac{A_j^{1,2}(t)B_j^{1,2}(t)d_j^2 s^2}{\alpha_k^j g_k} \right) \\ &- \sum_{k \in \{3,4,5\}} \left(\frac{A_j^{1,2}(t)B_j^k(t)d_j^2 s^2}{g_k} + \frac{A_j^k(t)B_j^{1,2}(t)d_j^2 s^2}{g_k} \right) + \\ &- \sum_{k \in N} \left(\frac{B_j^{1,2}(t)d_j s}{2g_k} - \frac{A_j^{1,2}(t)c_k d_j s}{g_k} \right) \\ \dot{C}_j^{1,2}(t) &= C_j^{1,2}(t)\rho - B_j^{1,2}(t)a_j s + \sum_{k \in \{1,2\}} \left(\frac{\alpha_k^k c_k^2}{4g_k} + \frac{(B_j^{1,2}(t)d_j s)^2}{4\alpha_k^j g_k} \right) + \\ &+ \sum_{k \in \{3,4,5\}} \frac{B_j^{1,2}(t)B_j^k(t)d_j^2 s^2}{2g_k} + \sum_{k \in N} \frac{B_j^{1,2}(t)c_k d_j s}{2g_k}, \end{aligned}$$

с граничными условиями $A_j^{1,2}(t_0 + j\Delta t + \bar{T}) = 0$, $B_j^{1,2}(t_0 + j\Delta t + \bar{T}) = 0$ и $C_j^{1,2}(t_0 + j\Delta t + \bar{T}) = 0$.

Для функций $A_j^i(t_0 + j\Delta t + \bar{T})$, $B_j^i(t_0 + j\Delta t + \bar{T})$ и $C_j^i(t_0 + j\Delta t + \bar{T})$, $i = 3, \dots, 5$:

$$\begin{aligned} \dot{A}_j^i(t) &= A_j^i(t)(\rho + 2s) + \sum_{k \in \{1,2\}} \left(\frac{1}{4g_i} + \frac{2A_j^i(t)A_j^{1,2}(t)d_j^2s^2}{\alpha_k^j g_k} \right) + \\ &+ \sum_{k \in \{3,4,5\}} \frac{2A_j^k(t)A_j^i(t)d_j^2s^2}{g_i} - \sum_{k \in N} \frac{A_j^i(t)d_j s}{g_k}, \\ \dot{B}_j^i(t) &= B_j^i(t)(\rho + s) - 2A_j^i(t)a_j s - \frac{c_i}{2g_i} - \sum_{k \in N} \left(\frac{B_j^i(t)d_j s}{2g_k} - \frac{A_j^i(t)c_k d_j s}{g_k} \right) \\ &+ \sum_{k \in \{1,2\}} \left(\frac{A_j^i(t)B_j^{1,2}(t)d_j^2s^2}{\alpha_k^j g_k} + \frac{A_j^{1,2}(t)B_j^i(t)d_j^2s^2}{\alpha_k^j g_k} \right) + \\ &+ \sum_{k \in \{3,4,5\}} \left(\frac{A_j^i(t)B_j^k(t)d_j^2s^2}{g_k} + \frac{A_j^k(t)B_j^i(t)d_j^2s^2}{g_k} \right) \\ \dot{C}_j^i(t) &= C_j^i(t)\rho + \frac{c_i^2}{4g_i} - B_j^i(t)a_j s + \sum_{k \in \{1,2\}} \frac{B_j^i(t)B_j^{1,2}(t)d_j^2s^2}{2\alpha_k^j g_k} + \\ &+ \sum_{k \in \{3,4,5\}} \frac{B_j^i(t)B_j^k(t)d_j^2s^2}{2g_k} + \sum_{k \in N} \frac{B_j^k(t)c_k d_j s}{2g_k}, \end{aligned}$$

с граничными условиями $A_j^i(t_0 + j\Delta t + \bar{T}) = 0$, $B_j^i(t_0 + j\Delta t + \bar{T}) = 0$ и $C_j^i(t_0 + j\Delta t + \bar{T}) = 0$.

Подставляя функцию $q_{\alpha_j}^*(t, p)$ (3.160) в систему (3.157), получаем динамику равновесной по Нэшу траектории, соответствующей выбранным весовым коэффициентам игроков 1 и 2:

$$\dot{p}(t) = s \left(a - d \sum_{i \in N} q_i^{*\alpha_j}(t, p) - p(t) \right), \quad p(t_0 + j\Delta t) = p_{j,0}. \quad (3.161)$$

Обозначим через $p_{\alpha_j}^*(t)$ решение дифференциального уравнения (3.161), соответствующее фиксированным весовым коэффициентам $\alpha_j = (\alpha_1^j, \alpha_2^j)$. Далее Парето-оптимальную траекторию будем обозначать через $p_{\alpha_j}^*(t)$.

3.4.3.3 Индивидуальные выигрыши игроков при кооперации

В предыдущем разделе определены выигрыши игроков 3, 4 и 5, но не индивидуальные выигрыши игроков 1 и 2. Следующим шагом является определение индивидуальных выигрышей для сотрудничающих игроков. Это можно сделать, подставив Парето-оптимальные траекторию $p_{\alpha_j}^*(t)$ (3.161) и стратегии $q_{\alpha_j}^*(t, p)$ (3.160) в формулу выигрыша каждого игрока (3.145). Выигрыш игроков 1 и 2 в усеченной подыгре $\tilde{\Gamma}_j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ вдоль Парето-оптимальной траектории $p_{\alpha_j}^*(t)$ при стратегиях $q_{\alpha_j}^*(t, p)$ обозначим соответственно через $W_1^{\alpha_j}(t, p)$ и $W_2^{\alpha_j}(t, p)$. Другой (и более предпочтительный) способ — применить результаты статьи [165], где специальная форма уравнения Гамильтона–Якоби–Беллмана решена для каждого игрока, но с фиксированными стратегиями $q_{\alpha_j}^*(t, p)$ (3.160). (Таким образом, максимизация в правой части уравнения не задействуется.)

3.4.3.4 Состоятельные во времени решения

В статье [165] свойство временной состоятельности кооперативного решения игры с нетрансферабельной полезностью определено следующим образом:

Определение 3.4.0.5 Вектор выигрыша $W(t, p^*(t)) = (W_1(t, p^*(t)), \dots, W_n(t, p^*(t)))$ вдоль траектории $p^*(t)$ является состоятельным во времени в игре $\Gamma_c(p_0, T - t_0)$ с нетрансферабельной полезностью, если он удовлетворяет следующим условиям:

1 вектор выигрыша

$$W(t, p^*(t)) = (W_1(t, p^*(t)), \dots, W_n(t, p^*(t))), \quad t \in [t_0, T] \quad (3.162)$$

является Парето-оптимальным;

2 вектор выигрыша обладает свойством индивидуальной рациональности

$$W_i(t, p^*(t)) \geq V_i(t, p^*(t)), \quad t \in [t_0, T], \quad i = \overline{1, n}, \quad (3.163)$$

где $V_i(t, p^*(t))$ — индивидуальный выигрыш игрока $i = \overline{1, n}$.

Определение 3.4.0.6 Определим множество $S_j^\tau = \bigcap_{t_0+j\Delta t \leq t < \tau} S_t^j$ для $\tau \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t)$, $j = 0, \dots, l$, где S_t^j — множество значений $\alpha_j = (\alpha_1^j, \alpha_2^j)$, удовлетворяющих индивидуальной рациональности в момент времени t для $i = \overline{1, n}$:

$$W_i^{\alpha_j}(t, p_{\alpha_j}^*(t)) \geq V_i^j(t, p_{\alpha_j}^*(t)), \quad (3.164)$$

где $V_i(t, p^*(t))$ — выигрыш игрока $i \in N$ в некооперативной игре с равновесными по Нэшу стратегиями, но рассчитанный вдоль Парето-оптимальной траектории $p_{\alpha_j}^*(t)$.

Множество $S_j^{t_0+(j+1)\Delta t}$ состоит из значений $\alpha_j = (\alpha_1^j, \alpha_2^j)$, удовлетворяющих индивидуальной рациональности (3.164) на протяжении временного интервала до обновления информации об игре в момент времени $t = t_0 + (j+1)\Delta t$. Множество $S_j^{t_0+(j+1)\Delta t}$ предоставляет соответствующий набор состоятельных во времени решений:

$$\left\{ W^{\alpha_j}(t, p_{\alpha_j}^*(t)) = (W_1^{\alpha_j}(t, p_{\alpha_j}^*(t)), \dots, W_n^{\alpha_j}(t, p_{\alpha_j}^*(t))), \right. \\ \left. t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t], \alpha_j \in S_j^{t_0+(j+1)\Delta t}, i = \overline{1, n}, j = 0, \dots, l \right\}. \quad (3.165)$$

Для более общего случая, когда свойство индивидуальной рациональности (3.164) выполняется на всем временном интервале, в статье [165] был установлен следующий результат.

Теорема 3.4.3 Принцип оптимальности решения, согласно которому игроки соглашаются выбирать одинаковые весовые коэффициенты $\alpha = (\alpha_1, \dots, \alpha_n) \in \Lambda$ во всех подыграх $\Gamma_c(p_\alpha^*(t), T - t)$, $t \in [t_0, T]$, таким образом, что

$$W_i^\alpha(t, p_\alpha^*(t)) \geq V_i(t, p_\alpha^*(t)), \quad t \in [t_0, T], \quad i = \overline{1, n}, \quad (3.166)$$

дает состоятельное во времени решение кооперативной игры $\Gamma_c(p_0, T - t_0)$.

3.4.3.5 Концепция решения

Предположим, что весовые коэффициенты $\alpha_j \in S_j^{t_0+(j+1)\Delta t}$ выбраны в начале каждой усеченной подыгры $\tilde{\Gamma}_j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$. Тогда каждая усеченная подыгра развивается вдоль Парето-оптимальной траектории $p_{\alpha_j}^*(t)$ и, таким образом, вся игра $\Gamma_c(p_0, T - t_0)$ развивается вдоль траектории следующего вида, где $p_{j,0}^* = p_{\alpha_{j-1}}^*(t_0 + j\Delta t)$.

Определение 3.4.0.7 *Условно-Парето-оптимальная траектория $\{\hat{p}^*(t)\}_{t=t_0}^T$ представляет собой комбинацию Парето-оптимальных траекторий $p_{\alpha_j}^*(t)$ для каждой усеченной подыгры $\tilde{\Gamma}_j(p_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ и заданных весовых коэффициентов $\alpha_j \in S_j^{t_0+(j+1)\Delta t}$:*

$$\{\hat{p}^*(t)\}_{t=t_0}^T = \begin{cases} p_{\alpha_0}^*(t), & t \in [t_0, t_0 + \Delta t), \\ \dots, \\ p_{\alpha_j}^*(t), & t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t), \\ \dots, \\ p_{\alpha_l}^*(t), & t \in [t_0 + l\Delta t, t_0 + (l+1)\Delta t]. \end{cases} \quad (3.167)$$

Теперь определим решение для всей игры $\Gamma_c(p_0, T - t_0)$. Введем обозначение $p_{j,1}^* = p_{\alpha_j}^*(t_0 + (j+1)\Delta t)$.

Определение 3.4.0.8 *Пусть $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$. Результирующее решение для игры $\Gamma_c(p_0, T - t_0)$ с динамическим обновлением информации, соответствующее выбранным весовым коэффициентам $\alpha_j \in S_j^{t_0+(j+1)\Delta t}$ для каждой усеченной подыгры $\tilde{\Gamma}_j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, $j = 0, \dots, l$, определяется следующим образом:*

$$\begin{aligned} \hat{W}_i(t, \hat{p}^*(t)) = & \sum_{m=j+1}^l \left[W_i^{\alpha_m}(t_0 + m\Delta t, p_{m,0}^*) - W_i^{\alpha_m}(t_0 + (m+1)\Delta t, p_{m,1}^*) \right] + \\ & + \left[W_i^{\alpha_j}(t, p_{\alpha_j}^*(t)) - W_i^{\alpha_j}(t_0 + (j+1)\Delta t, p_{j,1}^*) \right]. \end{aligned} \quad (3.168)$$

Следовательно, множество всех возможных решений в игре $\Gamma_c(p_0, T - t_0)$ имеет

следующий вид:

$$\left\{ \hat{W}(t, \hat{p}^*(t)) = (\hat{W}_1(t, \hat{p}^*(t)), \dots, \hat{W}_n(t, \hat{p}^*(t))), t \in [t_0, T], \right. \\ \left. \alpha_j \in S_j^{t_0+(j+1)\Delta t}, \forall j = 0, \dots, l \right\}. \quad (3.169)$$

В игре с подвижным информационным горизонтом решение называется условно-Парето-оптимальным, если оно строится с использованием Парето-оптимальных решений в каждой усеченной подыгре. В этом смысле результирующее решение $\hat{W}(t, \hat{p}^*(t))$ называется условно-Парето-оптимальным.

3.4.3.6 Временная состоятельность концепции решения

В статье [111] свойство временной состоятельности результирующего решения (3.168) доказано для случая полной кооперации. Аналогичный результат можно получить при частичной кооперации.

Теорема 3.4.4 *Результирующее решение $\hat{W}(t, \hat{p}^*(t))$ является состоятельным во времени в игре $\Gamma_c(\hat{p}^*(t), T - t)$ с динамическим обновлением информации, если выполняется следующее условие при $\forall t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, $j = 0, \dots, l$, $i = \overline{1, n}$:*

$$W_i^{\alpha_j}(t, p_{\alpha_j}^*(t)) - V_i^j(t, p_{\alpha_j}^*(t)) \geq \\ \geq W_i^{\alpha_j}(t_0 + (j + 1)\Delta t, p_{j,1}^*) - V_i^j(t_0 + (j + 1)\Delta t, p_{j,1}^*), \quad (3.170)$$

3.4.4 Соглашение о сотрудничестве

Чтобы провести всестороннее сравнение теоретических результатов и реальной ситуации, необходимо отметить сложность такого соглашения в масштабах стран. Оно влияет не только на ситуацию на нефтяном рынке внутри стран, но и на его баланс и все макроэкономические показатели. Соглашение носит как экономический, так и политический характер. Согласно соглашению от 30 ноября 2016 г. члены ОПЕК и 11 стран, не входящих в ОПЕК, решили сократить добычу на 1,18–7,69% для разных стран. При этом Иран, Ливия

и Нигерия могут увеличить свою добычу. После того, как договоренность была достигнута, цены на нефть действительно начали расти. Однако к концу мая 2017 г. они упали почти до прежнего уровня. Страны-аутсайдеры, в частности сланцевые производители США, увеличили добычу. Большинству стран-участниц соглашения удалось его выполнить. По прошествии шести месяцев все эти страны решили продлить договор на следующий год.

Для моделирования этого кооперативного соглашения построены линейные стратегии для игроков 1 и 2. Равновесные по Нэшу стратегии используются в некооперативной дифференциальной игре для других игроков. Оптимальные стратегии для игроков 3, 4 и 5 находятся так же, как в разделе 1, с использованием уравнения Гамильтона–Якоби–Беллмана при стратегиях (3.171).

Согласно статистическим данным, используются стратегии игроков $i = 1, 2$ по кооперативному соглашению, которые имеют следующий вид:

$$\bar{q}_1^*(t) = 3.5\%q_1^{avr}(t_{nov})\left(1 - 3.5\%t/T\right) \quad (3.171)$$

$$\bar{q}_2^*(t) = 3.1\%q_2^{avr}(t_{nov})\left(1 - 3.1\%t/T\right), \quad (3.172)$$

где $q_{avr}(t_{nov}) = (q_1^{avr}(t_{nov}), q_2^{avr}(t_{nov})) = (39.838, 18.771)$ — средний объем добытой нефти в момент времени t_{nov} (ноябрь 2016 г.), а T соответствует маю 2017 г.

3.4.4.1 Численное моделирование

Вторая игра начинается сразу после саммита и длится с декабря 2016 г. по май 2017 г. Как и на предыдущем временном интервале, средние цены на нефть за каждый период рассчитываются только на основе двух основных торговых классификаций нефти марки Brent и легкой нефти (источник данных — агентство «Финам»). В качестве начальной цены для второй игры принята средняя цена в декабре 2016 г., $p_0 = 53,73$.

Период	Общее мировое предложение, млн баррелей в сутки	Brent, \$ за баррель	Light, \$ за баррель	Средняя цена, \$ за баррель
12.2016	97,324	55,008	52,449	53,728
01.2017	96,627	55,607	52,951	54,279
02.2017	96,412	56,122	53,661	54,892
03.2017	96,411	52,592	49,873	51,232
04.2017	97,075	53,841	51,216	52,529
05.2017	97,529	51,533	48,744	50,139

Для тестирования построенной модели с динамическим обновлением информации выбраны те же параметры игроков c_i и g_i , что и для предыдущего временного интервала, в предположении, что параметр спроса $a_j = 300$ в каждый Δt -период. Игроки используют оцененный спрос с параметрами a_j и b_j в качестве прогноза на следующие $\bar{T} = 3$ периода. Параметр d_j получен по формуле (3.154).

Период	a	d	i	Производитель	c	g
01.2017	300	2,530	1	Члены ОПЕК	3,169	0,7
02.2017	300	2,543	2	Страны, не входящие в ОПЕК	17,333	0,7
03.2017	300	2,542	3	США, сланцевая нефть	20,238	0,7
04.2017	300	2,580	4	США, несланцевая нефть	18,182	0,7
05.2017	300	2,549	5	Другие	20,867	0,7

На рисунке 5 изображены условно-Парето-оптимальная траектория $\hat{p}^*(t)$ (сплошная линия) и множество Парето-оптимальных траекторий $p_{\alpha_j}^*(t)$ (штриховые линии), соответствующих весовым коэффициентам $\alpha_j \in S_j^{t_0+(j+1)\Delta t}$ в каждой усеченной подыгре $\tilde{\Gamma}_j(p_{j,0}^*, t_0 + \Delta t, t_0 + \Delta t + \bar{T})$. На рисунке 6 для каждой группы стран представлены оптимальные стратегии, соответствующие $\{\hat{p}_{\alpha_j}^*(t)\}_{t=t_0}^T$, а также исторические объемы добычи нефти.

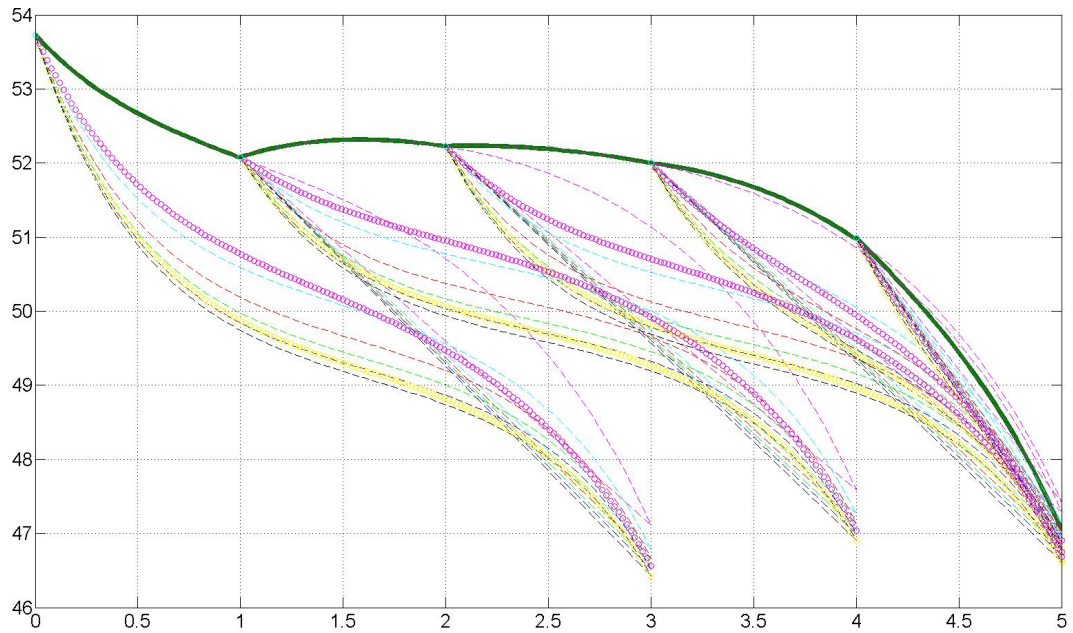


Рис. 3.24: Условно-Парето-оптимальная траектория $\hat{p}^*(t)$ (сплошная линия) и множество Парето-оптимальных траекторий $p_{\alpha_j}^*(t)$ (штриховые линии), соответствующих весовым коэффициентам $\alpha_j \in S_j^{t_0+(j+1)\Delta t}$ в каждой усеченной подыгре $\tilde{\Gamma}_j(p_{j,0}^*, t_0 + \Delta t, t_0 + \Delta t + \bar{T})$.

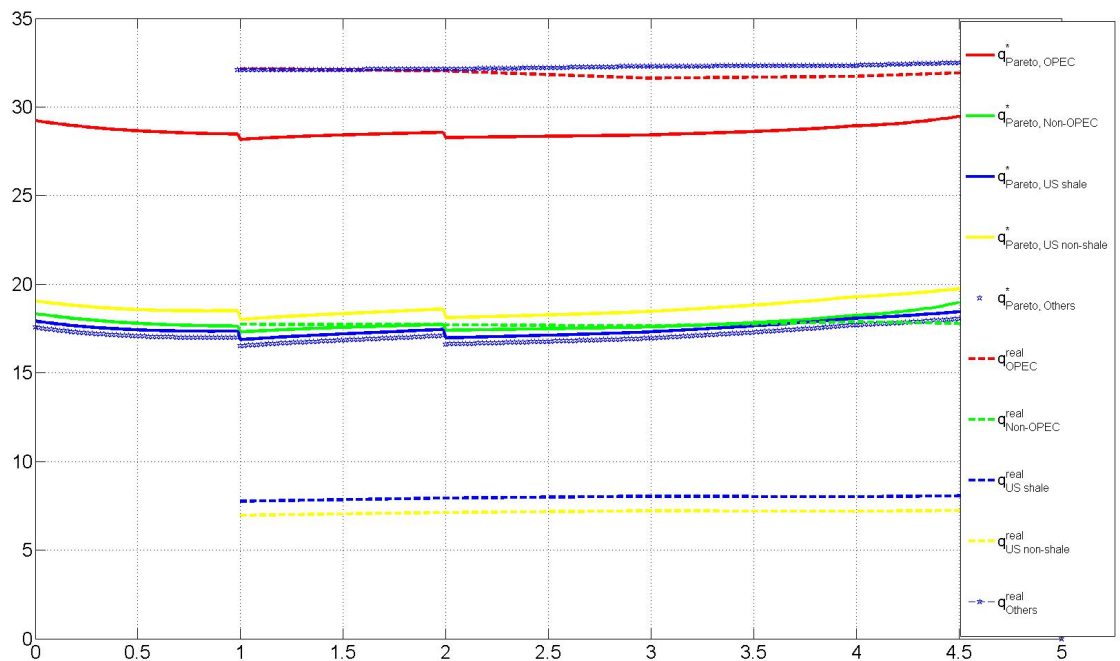


Рис. 3.25: Оптимальные стратегии с динамическим обновлением информации (сплошные линии) и соответствующие исторические объемы добычи нефти (штриховые линии).

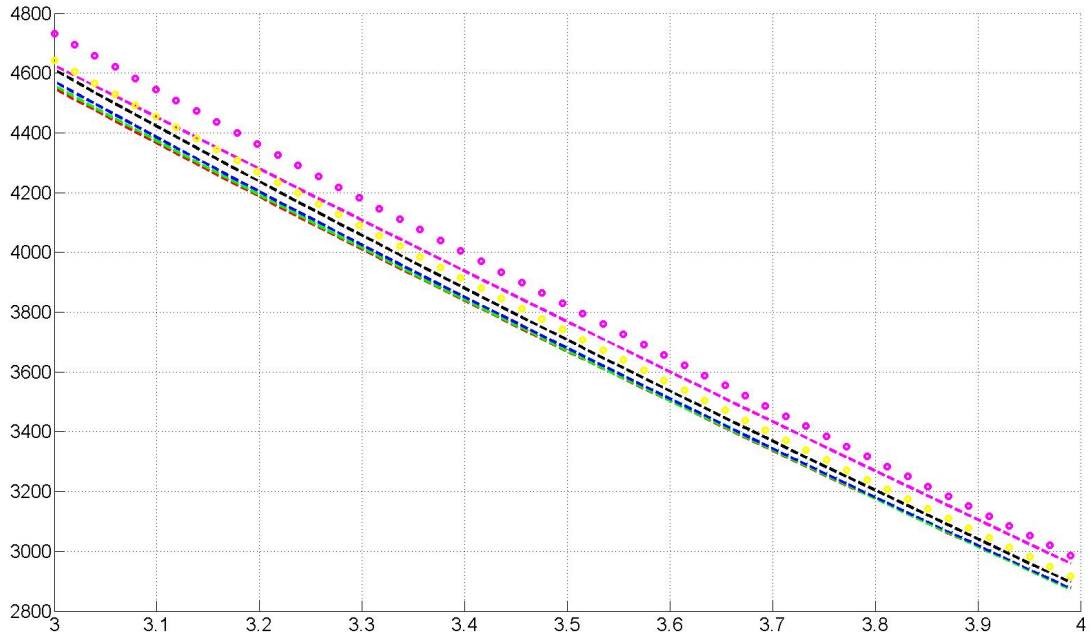


Рис. 3.26: Множество индивидуальных выигрышей $W_1^j(t, p_{\alpha_j}^*(t))$ игрока 1 при всевозможных весовых коэффициентах $\alpha_j \in S_j^{t_0+(j+1)\Delta t}$ в усеченной подыгре $\tilde{\Gamma}_j(p_{j,0}^*, t_0 + \Delta t, t_0 + \Delta t + \bar{T})$, $j = 3$.

На рисунке 7 показан индивидуальный выигрыш $W_1^j(t, p_{\alpha_j}^*(t))$ игрока $i = 1$ при кооперации (сплошные линии) при всевозможных весовых коэффициентах $\alpha_j \in S_j^{t_0+(j+1)\Delta t}$ и $\alpha_j^* \in S_j^{t_0+(j+1)\Delta t}$, $j = 3$. На рисунке 8 представлены выигрыш игроков $\hat{W}_i(t, \hat{p}^*(t))$, $i \in N$ (сплошные линии), а также выигрыш игроков $\hat{V}_i(t, \hat{p}^*(t))$, $i \in N$ (штриховые линии), соответствующие равновесным по Нэшу стратегиям вдоль условно-Парето-оптимальной траектории $\hat{p}^*(t)$.

3.4.4.2 Сравнение исторического кооперативного решения и предлагаемого решения

В этом разделе сравним историческое соглашения о сотрудничестве (см. (3.171) при стратегиях $\bar{q}^*(t)$) со состоятельным во времени кооперативным решением (3.168). На рисунке 9 представлена условно-Парето-оптимальная траектория $\hat{p}_{\alpha_j}^*(t)$, соответствующая выбранным весовым коэффициентам $\alpha_j^* \in S_j^{t_0+(j+1)\Delta t}$ во всех усеченных подыграх $\tilde{\Gamma}_j(p_{j,0}^*, t_0 + \Delta t, t_0 + \Delta t + \bar{T})$, $j = 0, \dots, l$, а также траектория $\bar{p}^*(t)$, соответствующая историческому соглашению о сотрудничестве и историческая динамика цен на нефть.

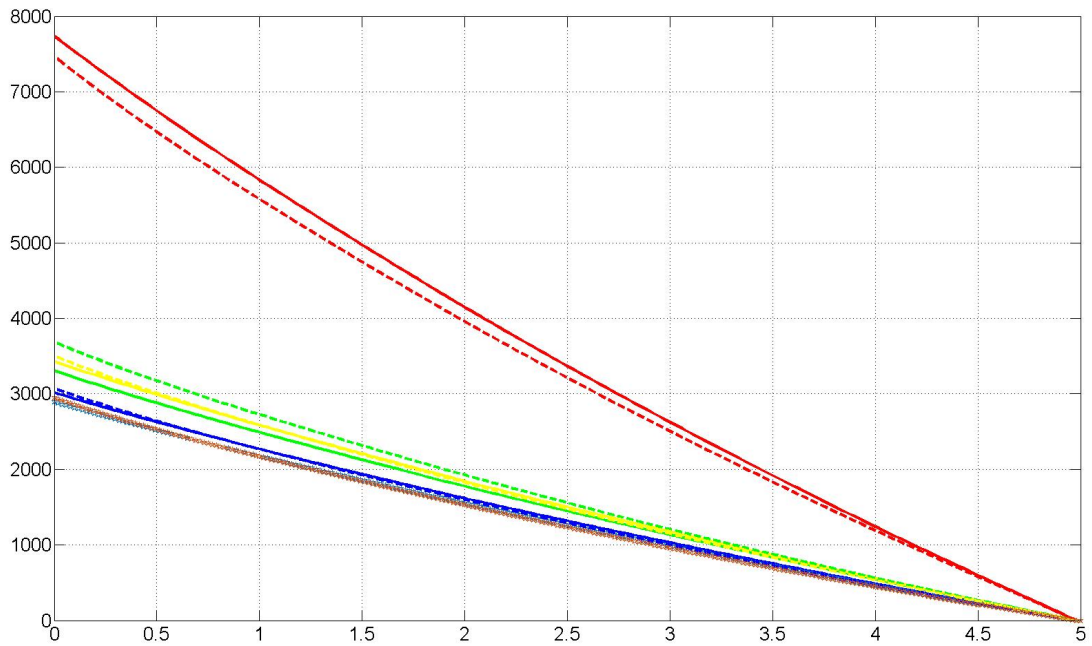


Рис. 3.27: $\hat{W}_i(t, \hat{p}^*(t))$ (сплошные линии) и $\hat{V}_i(t, \hat{p}^*(t))$ (штриховые линии), $i \in N$.

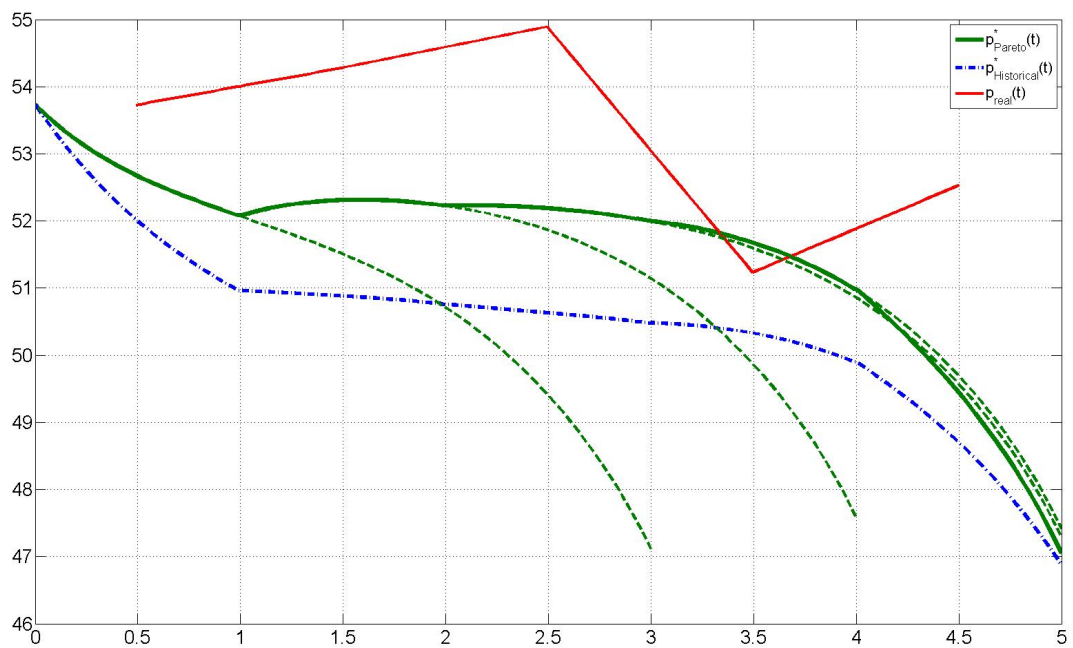


Рис. 3.28: Условно-Парето-оптимальная траектория $\hat{p}_{\alpha_j}^*(t)$, соответствующая выбранным весовым коэффициентам $\alpha_j^* \in S_j^{t_0+(j+1)\Delta t}$ $j = 0, \dots, l$ (толстая сплошная линия), траектория $\bar{p}^*(t)$, соответствующая историческому соглашению о сотрудничестве (3.171) (штриховая линия) и историческая динамика цен на нефть (тонкая сплошная линия).

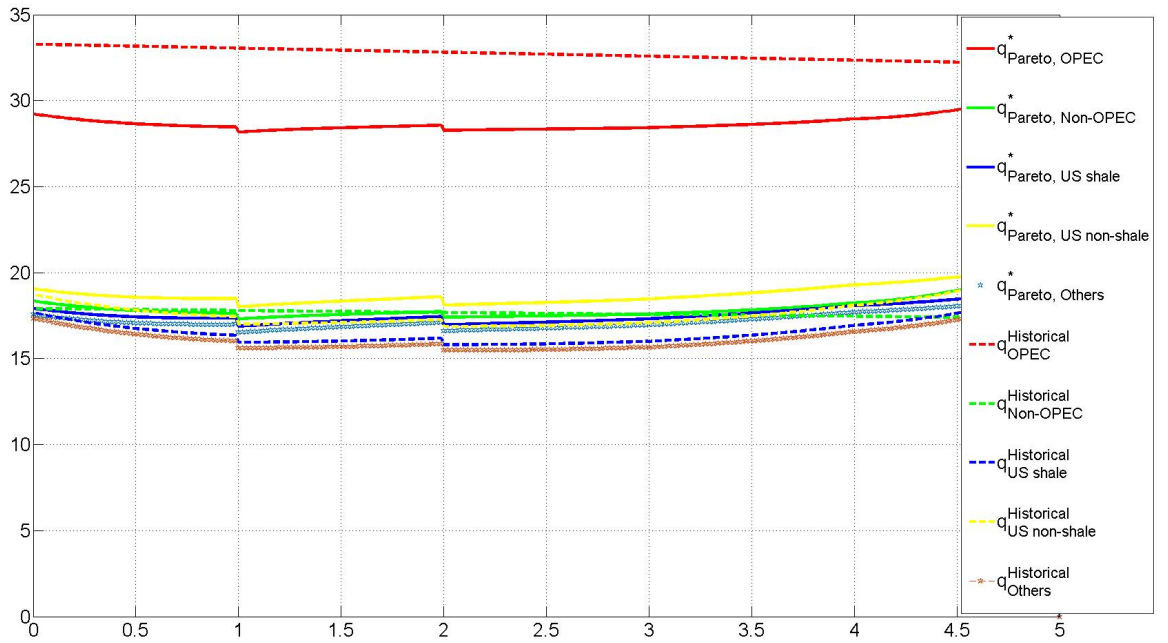


Рис. 3.29: Оптимальные стратегии (3.150) с динамическим обновлением информации (сплошные линии) и стратегии (3.171), соответствующие историческому соглашению о сотрудничестве (штриховые линии).

На рисунке 10 для каждой группы стран показаны оптимальные стратегии (3.150), соответствующие $\hat{p}_{\alpha_j}^*(t)$, стратегии $\bar{q}^*(t)$ (3.171), соответствующие историческому соглашению о сотрудничестве, а также исторические объемы добычи нефти. На рисунке 11 показан выигрыш игроков $\hat{W}_i(t, \hat{p}^*(t))$, $i \in N$ (сплошные линии) вдоль траектории $\hat{p}^*(t)$ и выигрыши, соответствующие историческому соглашению о сотрудничестве (штриховые линии).

3.4.4.3 Анализ оптимальности соглашения о сотрудничестве

В целом, соглашение о сотрудничестве было оправданным для рассматриваемых стран. Это можно увидеть, сравнив кооперативное и некооперативное решения на втором временном интервале. На рисунке 12 изображена условно-Парето-оптимальная траектория $\hat{p}_{\alpha_j}^*(t)$ (сплошная линия), соответствующая выбранным весовым коэффициентам $\alpha_j^* \in S_j^{t_0+(j+1)\Delta t}$ во всех усеченных подыграх $\tilde{\Gamma}_j(p_{j,0}^*, t_0 + \Delta t, t_0 + \Delta t + \bar{T})$, $j = 0, \dots, l$, а также условно-некооперативная траектория динамики цен на нефть $\hat{p}_{NE}(t)$, соответствующая равновесию по Нэшу (штриховая линия).

Тем не менее, анализ выигрышей отдельных стран подтверждает выгоду

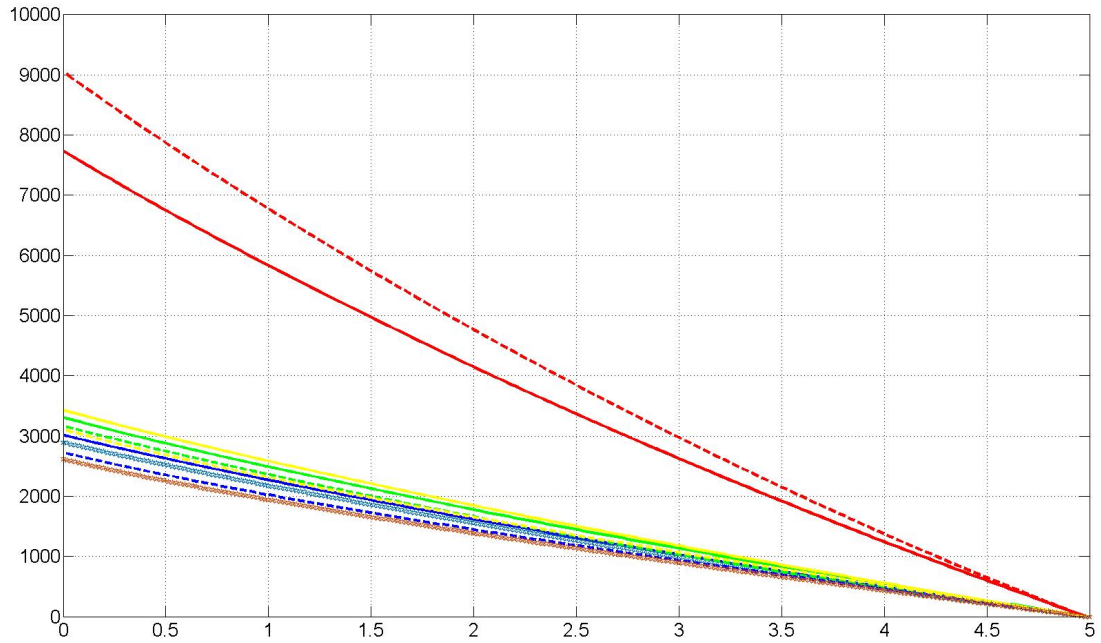


Рис. 3.30: $\hat{W}_i(t, \hat{p}^*(t))$, $i \in N$ (сплошные линии) и выигрыши, соответствующие историческому соглашению о сотрудничестве (штриховые линии).

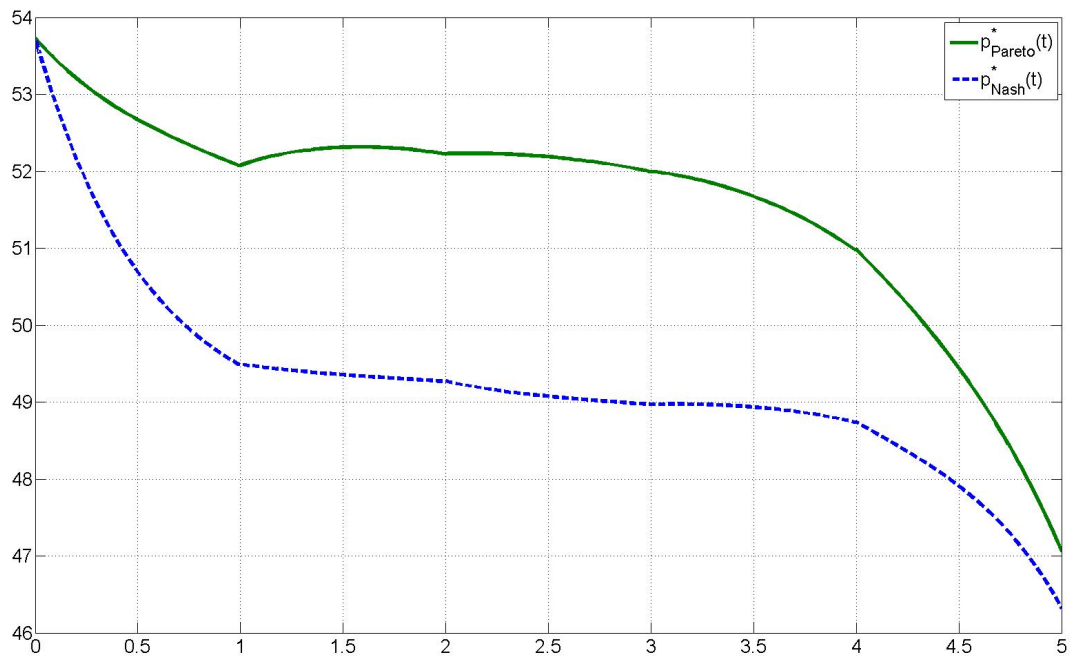


Рис. 3.31: Условно-Парето-оптимальная траектория $\hat{p}_{\alpha_j}^*(t)$ (сплошная линия) и условно-некооперативная траектория динамики цен на нефть $\hat{p}_{NE}(t)$ (штриховая линия).

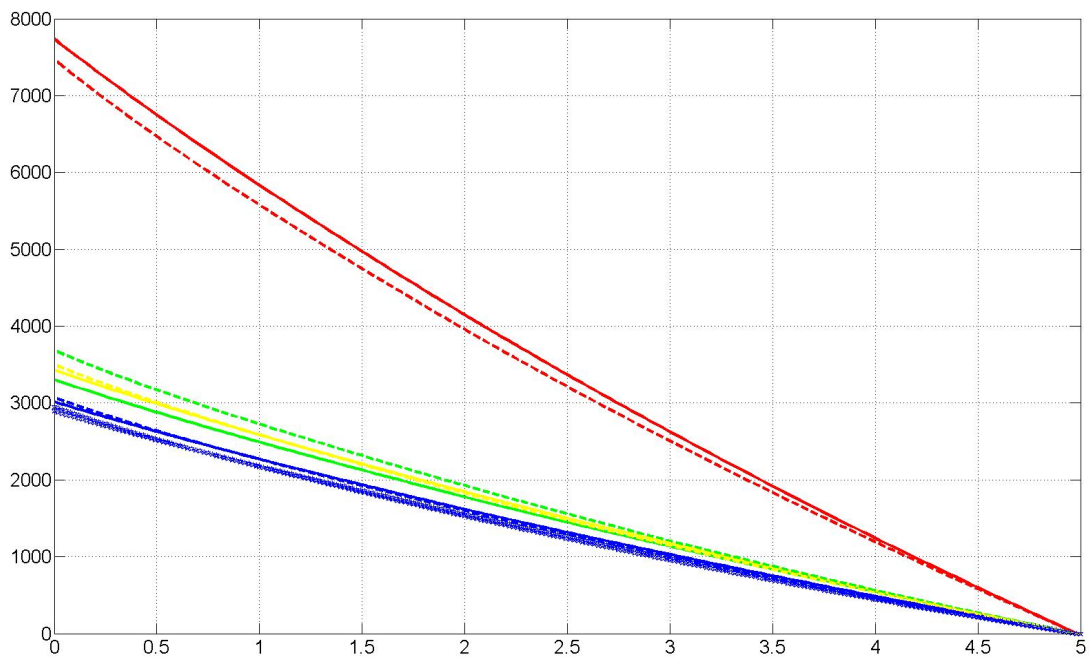


Рис. 3.32: $\hat{W}_i(t, \hat{p}^*(t))$, $i \in N$ (сплошные линии) и выигрыши игроков, соответствующие равновесию по Нэшу $\hat{V}_i(t, \hat{p}_{NE}(t))$ (штриховые линии).

соглашения только для игрока 1. На рисунке 13 показаны выигрыши игроков $\hat{W}_i(t, \hat{p}^*(t))$, $i \in N$ (сплошные линии) вдоль траектории $\hat{p}^*(t)$, а также выигрыши игроков, соответствующие равновесию по Нэшу $\hat{V}_i(t, \hat{p}_{NE}(t))$ (штриховые линии).

Этот анализ соответствует историческому решению. Фактически выгоды получают только страны ОПЕК. Предлагаемое моделью решение предполагает, что страны должны еще больше сократить добычу нефти. Это привело бы к еще большему росту цен (рисунок 9) и, следовательно, к прибыли для всех участников рынка. Таким образом, у стран есть рациональные стимулы к такому кооперативному поведению. Наличие этих стимулов для сговора и соблюдения достигнутых договоренностей побудило его участников продлить договор на встрече в мае 2017 г.

Глава 4

Дифференциальные игры с непрерывным обновлением информации

Эта глава посвящена новому классу дифференциальных игр с непрерывным обновлением информации. В отличие от предыдущих глав, здесь предполагается, что в каждый момент времени игроки имеют или используют информацию об игре, определенную на замкнутом временном интервале. Однако с течением времени информация об игре обновляется, а именно происходит непрерывный сдвиг временного интервала, на котором определена доступная игрокам информация. Информация об игре состоит из информации об уравнениях движения и функциях выигрыша игроков. Для этого класса игр прямое применение классических подходов к определению принципов оптимальности, таких как кооперативные стратегии или равновесие по Нэшу, невозможно. Предметом исследования являются уравнения Гамильтона–Якоби–Беллмана, принцип максимума Понтрягина и условия для линейно-квадратичных игровых моделей с непрерывным обновлением информации, как в кооперативном, так и некооперативном случае.

4.1 Модель дифференциальной игры с непрерывным обновлением информации

4.1.1 Исходная дифференциальная игра

Рассмотрим дифференциальную игру n лиц $\Gamma(x_0, T - t_0)$ с заданной продолжительностью, определенную на временном интервале $[t_0, T]$. В этой игре $|N| = n$.

Уравнения движения имеют вид

$$\begin{aligned} \dot{x}(t) &= f(t, x, u), \\ x(t_0) &= x_0, \\ x, x_0 &\in \mathbb{R}^l, \quad u = (u_1, \dots, u_n), \quad u_i = u_i(t, x) \in U_i \subset \text{comp} \mathbb{R}^k, \quad t \in [t_0, T]. \end{aligned} \quad (4.1)$$

Определим функцию выигрыша игрока i следующим образом:

$$K_i(x_0, T - t_0; u) = \int_{t_0}^T g^i[t, x(t), u(t, x)] dt, \quad i \in N, \quad (4.2)$$

где $g^i[t, x(t), u(t, x(t))]$ и $f(t, x, u)$ — интегрируемые функции, $x(t)$ — решение задачи Коши (4.1) при фиксированном $u(t, x) = (u_1(t, x), \dots, u_n(t, x))$. Набор стратегий $u(t, x) = (u_1(t, x), \dots, u_n(t, x))$ называется допустимым, если задача (4.1) имеет единственное продолжимое решение на отрезке $[t_0, T]$ (см. п.1.1.1). Здесь и ниже, при определении усеченной подыгры используются условия существования, единственности и продолжимости А.Ф. Филиппова [135].

На основе исходной дифференциальной игры с заданной продолжительностью T , построим соответствующую дифференциальную игру с непрерывным обновлением информации.

4.1.2 Игра с непрерывным обновлением информации

Рассмотрим дифференциальную игру n лиц $\Gamma(x_0, t_0, \bar{T})$, определенную на временном интервале $[t_0, t_0 + \bar{T}]$, где $0 < \bar{T} < +\infty$.

Уравнения движения имеют вид

$$\begin{aligned} \dot{x}_{t_0}(s) &= f(s, x_{t_0}, u^{t_0}), \\ x_{t_0}(t_0) &= x_0, \\ x_{t_0}, x_0 &\in \mathbb{R}^l, \quad u^{t_0} = (u_1^{t_0}, \dots, u_n^{t_0}), \quad u_i^{t_0} = u_i^{t_0}(s, x_{t_0}) \in U_i \subset \text{comp} \mathbb{R}^k. \end{aligned} \quad (4.3)$$

Определим функцию выигрыша игрока i следующим образом:

$$K_i^{t_0}(x_0, t_0, \bar{T}; u^{t_0}) = \int_{t_0}^{t_0 + \bar{T}} g^i[s, x_{t_0}(s), u^{t_0}(s, x_{t_0})] ds, \quad i \in N, \quad (4.4)$$

где $x_{t_0}(s)$ и $u^{t_0}(s, x_{t_0})$ обозначают соответственно траекторию и стратегии в игре $\Gamma(x_0, t_0, \bar{T})$, а $\dot{x}_{t_0}(s)$ — производная функции x_{t_0} по s .

Подыгра дифференциальной игры с непрерывным обновлением информации. Рассмотрим дифференциальную игру n лиц $\Gamma(x, t, \bar{T})$, $t \in [t_0, +\infty)$, определенную на временном интервале $[t, t + \bar{T}]$, где $0 < \bar{T} < +\infty$.

Уравнения движения для подыгры $\Gamma(x, t, \bar{T})$ имеют вид

$$\begin{aligned} \dot{x}_t(s) &= f(s, x_t, u^t), \\ x_t(t) &= x, \\ x_t, x &\in \mathbb{R}^l, \quad u^t = (u_1^t, \dots, u_n^t), \quad u_i^t = u_i^t(s, x_t) \in U_i \subset \text{comp} \mathbb{R}^k, \quad s \in [t, t + \bar{T}]. \end{aligned} \quad (4.5)$$

Функция выигрыша игрока i в подыгре $\Gamma(x, t, \bar{T})$ определяется как

$$K_i^t(x, t; u^t) = \int_t^{t + \bar{T}} g^i[s, x_t(s), u^t(s, x_t)] ds, \quad i \in N, \quad (4.6)$$

где $x_t(s)$, $u^t(s, x_t)$ обозначают соответственно траектории и стратегии в подыгре $\Gamma(x, t, \bar{T})$, а $\dot{x}_t(s)$ — производная функции x_t по s .

Дифференциальная игра с непрерывным обновлением информации протекает по следующему правилу: *текущее время $t \in [t_0, +\infty)$ меняется, и в результате игроки непрерывно получают новую информацию об уравнениях движения и функциях выигрыша в игре $\Gamma(x, t, \bar{T})$.*

Ситуация $u(t, x)$ в дифференциальной игре с непрерывным обновлением информации имеет вид

$$u(t, x) = u^t(s, x_t)|_{s=t}, \quad t \in [t_0, +\infty), \quad (4.7)$$

где $u^t(s, x_t)$, $s \in [t, t + \bar{T}]$, обозначает набор стратегий в подыгре $\Gamma(x, t, \bar{T})$.

Траектория $x(t)$ в дифференциальной игре с непрерывным обновлением информации определяется в соответствии с уравнением

$$\begin{aligned} \dot{x}(t) &= f(t, x, u), \\ x(t_0) &= x_0, \\ x, x_0 &\in \mathbb{R}^l, \end{aligned} \quad (4.8)$$

где $u = u(t, x)$ — набор стратегии в игре с непрерывным обновлением информации (4.7), а $\dot{x}(t)$ обозначает производную x по t . Предположим, что стратегия с непрерывным обновлением информации, полученная с помощью (4.7), допустима или что задача (4.8) имеет единственное и продолжимое решение на отрезке $[t_0, +\infty]$ (см. п.1.1.1). Соответствующие условия существования, единственности и продолжимости для системы (4.1)–(4.2) были представлены А.Ф. Филипповым [135].

Существенное отличие игровой модели с непрерывным обновлением информации от классической дифференциальной игры $\Gamma(x_0, T - t_0)$ с заданной продолжительностью заключается в следующем: в исходной игре игроки ориентируются на выигрыши, которые они в конечном итоге получают на временном интервале $[t_0, T]$, а в игре с непрерывным обновлением информации в момент времени t — на их ожидаемые выигрыши (4.6), которые вычисляются на основе имеющейся информации на временном интервале $[t, t + \bar{T}]$ или на основе информации, доступной в момент времени t .

4.2 Модель некооперативной дифференциальной игры с непрерывным обновлением информации

4.2.1 Равновесие по Нэшу в игре с непрерывным обновлением информации

В рамках непрерывно обновляемой информации важно моделировать поведение игроков. Для этого используем концепцию равновесия по Нэшу в позиционных стратегиях. Для класса дифференциальных игр с непрерывным обновлением информации хотелось бы ввести это равновесие в следующем виде:

- для любого фиксированного $t \in [t_0, +\infty)$, $u^{NE}(t, x) = (u_1^{NE}(t, x), \dots, u_n^{NE}(t, x))$ совпадает с равновесием по Нэшу в игре (4.5),(4.6), определенной на временном интервале $[t, t + \bar{T}]$ в момент времени t .

Однако прямое применение классических подходов для определения равновесия по Нэшу в позиционных стратегиях невозможно. Рассмотрим два интервала, $[t, t + \bar{T}]$ и $[t + \epsilon, t + \bar{T} + \epsilon]$, где $\epsilon \ll \bar{T}$. Тогда согласно постановке задачи:

- $u^{NE}(t, x)$ в момент времени t совпадает с равновесием по Нэшу в позиционных стратегиях в игре, определенной на временном интервале $[t, t + \bar{T}]$,
- $u^{NE}(t + \epsilon, x)$ в момент времени $t + \epsilon$ должно совпадать с равновесием по Нэшу в позиционных стратегиях в игре, определенной на временном интервале $[t + \epsilon, t + \epsilon + \bar{T}]$.

Для построения таких стратегий рассмотрим концепцию обобщенного равновесия по Нэшу в позиционных (и программных) стратегиях с непрерывным обновлением информации как принцип оптимальности:

$$\tilde{u}_{fb}^{NE}(t, x; s, x_t) = (\tilde{u}_{fb,1}^{NE}(t, x; s, x_t), \dots, \tilde{u}_{fb,n}^{NE}(t, x; s, x_t)), \quad t \in [t_0, T], \quad s \in [t, t + \bar{T}], \quad (4.9)$$

$$(\tilde{u}_{ol}^{NE}(t, x; s) = (\tilde{u}_{ol,1}^{NE}(t, x; s), \dots, \tilde{u}_{ol,n}^{NE}(t, x; s)), \quad t \in [t_0, T], \quad s \in [t, t + \bar{T}]), \quad (4.10)$$

которое будет использовано в дальнейшем для построения стратегий $u_{fb}^{NE}(t, x)$ ($u_{ol}^{NE}(t, x)$).

Определение 4.2.1 Ситуация $\tilde{u}_{fb}^{NE}(t, x; s, x_t) = (\tilde{u}_{fb,1}^{NE}(t, x; s, x_t), \dots, \tilde{u}_{fb,n}^{NE}(t, x; s, x_t))$ ($\tilde{u}_{ol}^{NE}(t, x; s) = (\tilde{u}_{ol,1}^{NE}(t, x; s), \dots, \tilde{u}_{ol,n}^{NE}(t, x; s))$) называется обобщенным равновесием по Нэшу в позиционных (программных) стратегиях с непрерывным обновлением информации, если для любого фиксированного $t \in [t_0, +\infty)$ ситуация $\tilde{u}_{fb}^{NE}(t, s, x)$ ($\tilde{u}_{ol}^{NE}(t, s, x)$) является равновесием по Нэшу в позиционных (программных) стратегиях в игре $\Gamma(x, t, \bar{T})$.

Важно отметить, что обобщенное равновесие по Нэшу в позиционных (программных) стратегиях $\tilde{u}_{fb}^{NE}(t, x; s, x_t)$ ($\tilde{u}_{ol}^{NE}(t, x; s)$) с непрерывным обновлением информации при фиксированном t является функцией s и x , где s определено на временном интервале $[t, t + \bar{T}]$. Однако, при каждом фиксированном значении параметра t ситуация \tilde{u}^{NE} является равновесием по Нэшу в игре $\Gamma(x, t, \bar{T})$ (без обновления информации). При этом подчеркнем, что в игре $\Gamma(x, t, \bar{T})$ равновесие по Нэшу понимается в том смысле, что выигрыш i -го игрока не увеличивается при отклонении на другую допустимую ситуацию, которая отличается от предыдущей лишь стратегией i -го игрока (т.е. не один игрок изменяет свою стратегию независимо от других игроков, а изменение возможно лишь в классе допустимых ситуаций). Используя обобщенное равновесие по Нэшу с непрерывным обновлением информации, можно определить концепцию решения для игры с непрерывным обновлением информации.

Определение 4.2.2 Ситуация $u_{fb}^{NE}(t, x)$ ($u_{ol}^{NE}(t, x)$) называется равновесием по Нэшу в позиционных (программных) стратегиях с непрерывным обновлением информации, если

$$\begin{aligned} u_{fb}^{NE}(t, x) &= \tilde{u}_{fb}^{NE}(t, x; s, x_t)|_{s=t} = (\tilde{u}_{fb,1}^{NE}(t, x; s, x_t)|_{s=t}, \dots, \tilde{u}_{fb,n}^{NE}(t, x; s, x_t)|_{s=t}), \\ (u_{ol}^{NE}(t, x) &= \tilde{u}_{ol}^{NE}(t, x; s)|_{s=t} = (\tilde{u}_{ol,1}^{NE}(t, x; s)|_{s=t}, \dots, \tilde{u}_{ol,n}^{NE}(t, x; s)|_{s=t})), \end{aligned} \quad (4.11)$$

где $\tilde{u}_{fb}^{NE}(t, x; s, x_t)$ ($\tilde{u}_{ol}^{NE}(t, x; s)$) — обобщенное равновесие по Нэшу в позиционных (программных) стратегиях с непрерывным обновлением информации согласно Определению 4.2.1.

Траектория $x_{fb}^{NE}(t)$ ($x_{ol}^{NE}(t)$), соответствующая равновесию по Нэшу с непрерывным обновлением информации $\tilde{u}_{fb}^{NE}(t, x; s, x_t)$ ($\tilde{u}_{ol}^{NE}(t, x; s)$) может быть получена из системы

$$\begin{aligned} \dot{x}(t) &= f(t, x, u_{fb}^{NE}) \quad (\dot{x}(t) = f(t, x, u_{ol}^{NE})) \\ x(t_0) &= x_0, \\ x &\in \mathbb{R}^l. \end{aligned} \tag{4.12}$$

В отличие от обобщенного равновесия по Нэшу в позиционных (программных) стратегиях с непрерывным обновлением информации, $u_{fb}^{NE}(t, x)$ ($u_{ol}^{NE}(t, x)$) не содержит равновесных по Нэшу позиционных (программных) стратегий для любых $s \in [t, t + \bar{T}]$. Ситуация $u_{fb}^{NE}(t, x)$ ($u_{ol}^{NE}(t, x)$) содержит только стратегии, реализуемые игроками в соответствии с процедурой непрерывного обновления информации с $s = t$ (см. описание в разделе 4.1.2). Вектор $u_{fb}^{NE}(t, x)$ ($u_{ol}^{NE}(t, x)$) будет использоваться как концепция решения в некооперативной игре с непрерывным обновлением информации. Понятие равновесия по Нэшу с непрерывным обновлением информации $u_{fb}^{NE}(t, x)$ ($u_{ol}^{NE}(t, x)$) близко к равновесию по Нэшу, совершенному по подыграм. В смысле ожидаемого выигрыша (4.6), для любого фиксированного текущего момента времени t индивидуальное отклонение от равновесия по Нэшу с непрерывным обновлением информации невыгодно для игроков, или для любой подыгры игры с непрерывным обновлением информации индивидуальное отклонение не выгодно в силу информационной структуры.

4.2.2 Уравнения Гамильтона–Якоби–Беллмана с непрерывным обновлением информации

В этом разделе рассмотрим только равновесие по Нэшу в позиционных стратегиях с непрерывным обновлением информации $u_{fb}^{NE}(t, x)$. Поэтому будем использовать упрощенное обозначение $u^{NE}(t, x)$.

Чтобы определить ситуацию $u^{NE}(t, x)$, необходимо определить обобщенное равновесие по Нэшу в позиционных стратегиях $\tilde{u}^{NE}(t, x; s, x_t)$ с непрерывным обновлением информации в игре с непрерывным обновлением информации

$\Gamma(x_0, t_0, \bar{T})$. Для этого воспользуемся модернизированной версией динамического программирования. В рамках этого подхода функция Беллмана $V^i(t; s, x)$ определяется как выигрыш игрока i в равновесии по Нэшу в позиционных стратегиях $\tilde{u}^{NE}(t, x; s, x_t)$ в подыгре, начинающейся в момент времени s из состояния x в игре, определенной на временном интервале $[t, t + \bar{T}]$:

$$V^i(t; s, x) = \int_s^{t+\bar{T}} g^i[\tau, x_t(\tau), \tilde{u}^{NE}(t, \tau, x_t(\tau))]d\tau, \quad i \in N \quad (4.13)$$

при условии

$$\begin{aligned} \dot{x}_t(\tau) &= f(\tau, x, \tilde{u}^{NE}), \\ x_t(s) &= x. \end{aligned} \quad (4.14)$$

Имеет место следующая теорема.

Теорема 4.2.1 *Ситуация $\tilde{u}^{NE}(t, x; s, x_t)$ является обобщенным равновесием по Нэшу в позиционных стратегиях с непрерывным обновлением информации в дифференциальной игре с непрерывным обновлением информации $\Gamma(x_0, t_0, \bar{T})$, если существуют непрерывно дифференцируемые по s и x функции $V^i(t; s, x) : [t_0, +\infty) \times [t, t + \bar{T}] \times R \rightarrow R$, $i \in N$, которые удовлетворяют следующей системе уравнений в частных производных:*

$$\begin{aligned} -V_s^i(t; s, x) &= \max_{\phi_i} \{g^i(s, x, \tilde{u}_{-i}^{NE}) + V_x^i(t; s, x)f(s, x, \tilde{u}_{-i}^{NE})\} = \\ &= g^i(s, x, \tilde{u}^{NE}) + V_x^i(t; s, x)f(s, x, \tilde{u}^{NE}), \\ V^i(t; t + \bar{T}, x) &= 0, \quad i \in N, \end{aligned} \quad (4.15)$$

где $\tilde{u}_{-i}^{NE}(\phi_i) = (\tilde{u}_1^{NE}, \dots, \phi_i, \dots, \tilde{u}_n^{NE})$.

Доказательство Согласно определению обобщенного равновесия по Нэшу с непрерывным обновлением информации, ситуация $\tilde{u}^{NE}(t, x; s, x_t)$ должна быть равновесием по Нэшу в позиционных стратегиях для любого фиксированного t .

Зафиксировав t в формулировке Теоремы 4.2.1 и, в частности, в (4.15), по-

лучим классические достаточные условия равновесия по Нэшу в позиционных стратегиях в дифференциальной игре с заданной продолжительностью $[t, t + \bar{T}]$ (см. [80]). Следовательно, для любого фиксированного t выполнены условия определения обобщенного равновесия по Нэшу с непрерывным обновлением информации. Теорема доказана. \square

Если можно найти обобщенное равновесие по Нэшу с непрерывным обновлением информации $\tilde{u}^{NE}(t, x; s, x_t)$ с помощью уравнений (4.15), то в результате применения процедуры (4.11) получаем желаемую ситуацию $u^{NE}(t, x)$.

4.2.2.1 Дифференциальная игра инвестиций в общественные блага

В качестве иллюстративного примера, рассмотрим дифференциальную игру инвестиций в общественные блага следующего вида. Два игрока инвестируют в общественный фонд знаний. Предполагается, что знание — чистое общественное благо, и каждый человек имеет к нему доступ. Модель такого типа впервые была сформулирована в [218] (см. также [50; 219]). В рамках этих моделей, каждый игрок имеет квадратичную функцию полезности и стоимость инвестиций увеличивается квадратично с увеличением объема инвестиций. Ниже мы используем другие функции для определения выигрышей игроков.

4.2.2.1.1 Модель исходной игры

Обозначим через $x(t)$ запас знаний в момент времени t , а через $u_i(t, x)$ — инвестицию игрока i в общественное знание в момент времени t . Предположим, что запас знания накапливается согласно дифференциальному уравнению

$$\dot{x}(t) = -\beta x(t) + u_1(t, x) + u_2(t, x), \quad x(0) = x_0, \quad (4.16)$$

где β — скорость устаревания знаний, а $x_0 > 0$ — заданный начальный запас знаний. Пусть возможные состояния в каждый момент времени t удовлетворяют условию $x(t) > 0$. Считаем, что каждый индивидуум i несет инвестиционные затраты, которые зависят от состояния и описываются функцией $\frac{u_i^2}{x}$. Предположим, что полезность каждого игрока является линейной функцией потребления

запаса знаний, $q_i x(t)$, где $q_i > 0$. Другими словами, функция выигрыша каждого игрока равна

$$K_i(x_0, T; u) = \int_0^T \left(q_i x(t) - \frac{u_i^2(t, x)}{x(t)} \right) dt. \quad (4.17)$$

Для определения равновесия по Нэшу в позиционных стратегиях $u^*(t, x)$ используем достаточные условия в виде уравнений Гамильтона–Якоби–Беллмана:

$$\begin{aligned} -V_t^i(t, x) &= \max_{u_i} \left(q_i x - \frac{u_i^2}{x} + V_x^i(t, x)(-\beta x + u_i + u_j^*) \right), \quad i \neq j, \\ V_i(T, x) &= 0, \quad i = 1, 2. \end{aligned} \quad (4.18)$$

Пусть функция Беллмана имеет вид

$$V^i(t, x) = a_i(t)x, \quad i = 1, 2.$$

В результате максимизации в (4.18) находим стратегию игрока i :

$$u_i^*(t, x) = \frac{x V_x^i(t, x)}{2} = \frac{a_i(t)x}{2}. \quad (4.19)$$

Подставляя (4.19) в (4.18), получаем следующую систему дифференциальных уравнений для $a_i(t)$:

$$\begin{aligned} \dot{a}_i(t) &= \beta a_i(t) + \frac{a_i^2(t)}{4} - a_i(t)a_j(t) - q_i, \quad i \neq j, \quad i = 1, 2, \\ a_i(T) &= 0. \end{aligned} \quad (4.20)$$

В качестве примера, рассмотрим симметричный случай $q_1 = q_2 = q$. Тогда решением системы (4.20) является

$$a_i(t) = \frac{2q(e^{v(t-T)} - 1)}{(\beta - v)e^{v(t-T)} - \beta - v}, \quad i = 1, 2,$$

где $v = \sqrt{\beta^2 - 3q}$, в предположении $\beta^2 > 3q$. Решение системы (4.20) определено на интервале $(-\infty; T + \frac{1}{v} \ln \frac{\beta+v}{\beta-v})$. Заметим, что $[0; T] \subset (-\infty; T + \frac{1}{v} \ln \frac{\beta+v}{\beta-v})$. В этом случае, равновесные по Нэшу позиционные стратегии имеют вид

$$u_i^*(t, x) = \frac{q(e^{v(t-T)} - 1)}{(\beta - v)e^{v(t-T)} - \beta - v} x(t), \quad i = 1, 2.$$

Результирующая равновесная траектория представляет собой функцию

$$x^*(t) = x_0 e^{-\frac{1}{3}(\beta+2v)t} \left(\frac{(\beta-v)e^{vt} - (\beta+v)e^{vT}}{\beta-v - (\beta+v)e^{vT}} \right)^{\frac{4}{3}}.$$

Ограничение на вектор состояния $x^*(t) > 0$ выполнено для всех $t \in [0; T]$ при условии $x_0 > 0$.

4.2.2.1.2 Модель игры с непрерывным обновлением информации

Теперь рассмотрим эту модель как игру с непрерывным обновлением информации. Предположим, что информация об уравнениях движения и функциях выигрыша непрерывно обновляется во времени. В каждый момент времени $t \in [0, +\infty)$ игрокам доступна информация только на временном интервале $[t, t + \bar{T}]$. Это означает, что из-за возможности изменения количества инвесторов или полезностей игроков в каждый момент времени они могут рассчитывать на устойчивость процесса только в течение периода времени \bar{T} .

Согласно Теореме 4.2.1, уравнения Гамильтона–Якоби–Беллмана для обобщенного равновесия по Нэшу в позиционных стратегиях имеют вид

$$\begin{aligned} -V_s^i(t; s, x) &= \max_{u_i} \left(q_i x - \frac{u_i^2}{x} + V_x^i(t; s, x)(-\beta x + u_i + \tilde{u}_j^{NE}) \right), \quad i \neq j, \\ V_i(t; t + \bar{T}, x) &= 0, \quad i = 1, 2. \end{aligned} \quad (4.21)$$

Определим функцию Беллмана

$$V_i(t; s, x) = a_i(t, s)x.$$

Максимизация в (4.21) дает стратегию игрока i :

$$\tilde{u}_i^{NE}(t, x; s, x_t) = \frac{xV_x^i(t; s, x)}{2} = \frac{xa_i(t, s)}{2}. \quad (4.22)$$

Подставляя выражение (4.22) в формулу (4.21), получаем следующую систему

дифференциальных уравнений для $a_i(t, s)$:

$$\begin{aligned} \dot{a}_i(t, s) &= \beta a_i(t, s) + \frac{a_i^2(t, s)}{4} - a_i(t, s)a_j(t, s) - q_i, \quad i \neq j, \quad i = 1, 2, \\ a_i(t, t + \bar{T}) &= 0. \end{aligned} \quad (4.23)$$

С учетом симметрии игроков, решение можно записать в виде

$$a_i(t, s) = \frac{2q(e^{v(s-t-\bar{T})} - 1)}{(\beta - v)e^{v(s-t-\bar{T})} - \beta - v}, \quad i = 1, 2,$$

где $v = \sqrt{\beta^2 - 3q}$.

Итак, обобщенным равновесием по Нэшу в позиционных стратегиях с непрерывным обновлением информации является ситуация

$$\tilde{u}_i^{NE}(t, x; s, x_t) = \frac{q(e^{v(s-t-\bar{T})} - 1)}{(\beta - v)e^{v(s-t-\bar{T})} - \beta - v} x(t).$$

Следуя процедуре (4.11), построим равновесие по Нэшу в позиционных стратегиях с непрерывным обновлением информации:

$$u_i^{NE}(t, x) = \tilde{u}_i^{NE}(t, x; s, x_t) \Big|_{s=t} = \frac{q(e^{-v\bar{T}} - 1)}{(\beta - v)e^{-v\bar{T}} - \beta - v} x(t).$$

Равновесная траектория $x^{NE}(t)$ с непрерывным обновлением информации имеет вид

$$x^{NE}(t) = x_0 e^{\left(-\beta + \frac{2q(e^{-v\bar{T}} - 1)}{(\beta - v)e^{-v\bar{T}} - \beta - v}\right)t}.$$

Ограничение на вектор состояния $x^{NE}(t) > 0$ выполнено для всех $t \in [0; T]$ при $x_0 > 0$.

На рисунках 4.1, 4.2 показаны графики решений, полученных при следующих значениях параметров:

$$\beta = 1/2, \quad q = 1/15, \quad T = 10, \quad \bar{T} = 1/2, \quad x_0 = 5.$$

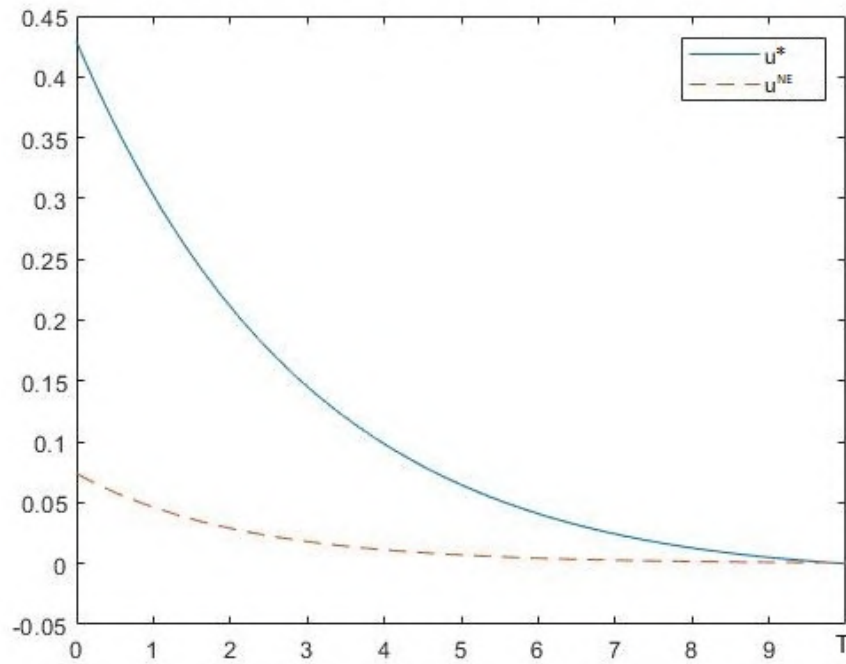


Рис. 4.1: Равновесие по Нэшу с непрерывным обновлением информации $u^{NE}(t, x)$ (красная пунктирная линия) и равновесные по Нэшу стратегии $u^*(t)$ в исходной игре (голубая сплошная линия).

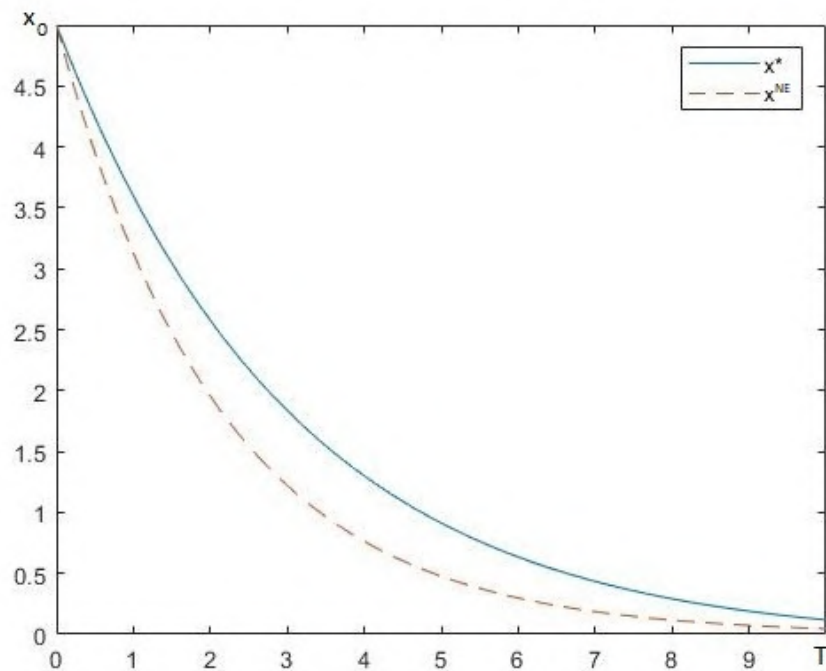


Рис. 4.2: Результирующая равновесная траектория $x^{NE}(t)$ с непрерывным обновлением информации (красная пунктирная линия) и результирующая равновесная траектория $x^*(t)$ в исходной игре (голубая сплошная линия).

Результаты численного моделирования показывают, что в случае непрерывного обновления информации игроки более осторожны, а их инвестиции меньше, чем в случае исходной игры (рис. 4.1). Как следствие, запас знаний уменьшается по сравнению с исходной игрой (рис. 4.2).

4.2.3 Принцип максимума Понтрягина с непрерывным обновлением информации

В этом разделе рассмотрим только равновесие по Нэшу в программных стратегиях с непрерывным обновлением информации $u_{ol}^{NE}(t, x)$. Поэтому будем использовать упрощенное обозначение $u^{NE}(t, \cdot)$.

Чтобы определить ситуацию $u^{NE}(t, x)$, необходимо определить обобщенное равновесие по Нэшу в программных стратегиях с непрерывным обновлением информации $\tilde{u}^{NE}(t, s, x)$ в игре с непрерывным обновлением информации. Для этого воспользуемся модернизированным вариантом принципа максимума Понтрягина. Прежде всего, определим действительную функцию H_i^t вида, $t \in [t_0, +\infty)$, $\tau \in [t, t + \bar{T}]$

$$H_i^t(\tau, x^t, u^t, \lambda^t) = \bar{T}g^i(\bar{T}\tau + t, x^t, u^t) + \lambda_i^t \bar{T}f(\bar{T}\tau + t, x^t, u^t). \quad (4.24)$$

Функция $H_i^t, i \in N$, называется текущим гамильтонианом и играет важную роль в принципе максимума Понтрягина. Переменная λ_i^t называется текущей переменной сопряженного состояния, связанной с переменной состояния x^t , или просто текущей сопряженной переменной. Важно отметить, что функция H_i^t , как и в уравнениях Гамильтона–Якоби–Беллмана, явно зависит от текущего времени t .

Как и в случае уравнений Гамильтона–Якоби–Беллмана, можно вывести необходимые условия для обобщенного равновесия по Нэшу с непрерывным обновлением информации следующего вида.

Теорема 4.2.2 Пусть функция $f(s, \cdot, u^t)$ непрерывно дифференцируема в R^l , $\forall s \in [t, t + \bar{T}]$, а функции $g^i(s, \cdot, u^t)$, $i \in N$, непрерывно дифференцируемы в R^l , $\forall s \in [t, t + \bar{T}]$. Если $\tilde{u}^{NE}(t, s, x)$ образует обобщенное равновесие по

Нэш в программных стратегиях с непрерывным обновлением информации в дифференциальной игре с непрерывным обновлением информации и для всех $t \in [t_0, +\infty)$ функция $\tilde{x}^t(s)$, $s \in [t, t + \bar{T}]$, является соответствующей траекторией состояния в игре $\Gamma(x, t, t + \bar{T})$, то для всех $t \in [t_0, +\infty)$ существуют n функций сопряженного состояния $\lambda_i^t(\tau, x)$, $\tau \in [0, 1]$, $i \in N$, удовлетворяющих следующим соотношениям:

1 для любого $\tau \in [0, 1]$

$$H_i^t(\tau, \tilde{x}^t, \tilde{u}^{NE}(t, \tau, x), \lambda^t) = \max_{\phi_i} \{H_i^t(\tau, \tilde{x}^t, \tilde{u}_{-i}^{NE}(t, \tau, x), \lambda^t)\}, i \in N, \quad (4.25)$$

$$\text{где } \tilde{u}_{-i}^{NE} = (\tilde{u}_1^{NE}, \dots, \phi_i, \dots, \tilde{u}_n^{NE}),$$

2 $\lambda_i^t(\tau, x)$ — набор решений системы сопряженных уравнений

$$\begin{aligned} \frac{d\lambda_i^t(\tau, x)}{d\tau} &= - \frac{\partial H_i^t(\tau, \tilde{x}^t(\tau), \tilde{u}^{NE}(t, \tau, x), \lambda^t)}{\partial x^t} = \\ &= -\bar{T} \frac{\partial g^i(\bar{T}\tau + t, \tilde{x}^t, \tilde{u}^{NE})}{\partial x^t} - \lambda_i^t(\tau, x) \bar{T} \frac{\partial f(\bar{T}\tau + t, \tilde{x}^t, \tilde{u}^{NE})}{\partial x_t}, i \in N, \end{aligned} \quad (4.26)$$

с условиями трансверсальности

$$\lambda_i^t(1, x) = 0, i \in N, \quad (4.27)$$

3 для любых $t \in [t_0, +\infty)$

$$\dot{\tilde{x}}^t(\tau) = \bar{T} f(\bar{T}\tau + t, \tilde{x}^t, \tilde{u}^{NE}), \quad \tilde{x}^t(0) = x, \quad \tau \in [0, 1]. \quad (4.28)$$

Доказательство Зафиксировав $t \geq t_0$, рассмотрим игру $\Gamma(x, t, t + \bar{T})$. С помощью замены переменных

$$\tau = \frac{s - t}{\bar{T}} \quad (4.29)$$

приведем уравнения движения (4.14) к виду

$$\dot{x}^t(\tau) = \bar{T} f(\bar{T}\tau + t, x^t, u^t), \quad x^t(0) = x, \quad \tau \in [0, 1]. \quad (4.30)$$

Функция выигрыша игрока $i \in N$ задается выражением

$$K_i^t(x, t, \bar{T}; u^t) = \int_0^1 \bar{T} g^i[\bar{T}\tau + t, x^t(\tau), u^t(\tau, x)] d\tau, \quad i \in N. \quad (4.31)$$

Для задачи оптимизации (4.30)–(4.31) гамильтониан принимает вид

$$\begin{aligned} H_i^t(\tau, x^t, u^t, \lambda^t) = \\ = \bar{T} g^i(\bar{T}\tau + t, x^t(\tau), u^t(\tau, x)) + \lambda_i^t(\tau, x) \bar{T} f(\bar{T}\tau + t, x^t(\tau), u^t(\tau, x)). \end{aligned} \quad (4.32)$$

Если $\tilde{u}^{NE}(t, \tau, x)$ является обобщенным равновесием по Нэшу в программных стратегиях с непрерывным обновлением информации в дифференциальной игре с непрерывным обновлением информации, то (см. Определение 4.2.1) при каждом фиксированном $t \geq t_0$, $\tilde{u}^{NE}(t, \tau, x)$ образует равновесие по Нэшу в программных стратегиях в игре $\Gamma(x, t, t + \bar{T})$. Следовательно, при любом фиксированном $t \geq t_0$ условия 1–3 теоремы выполнены как необходимые условия равновесия по Нэшу в программных стратегиях (см. [220]). Теорема доказана. \square

Отметим, что если для любых $t \geq t_0$ функции H_i^t вогнуты по (x_t, u_t) при всех $i \in N$, то условия теоремы достаточны для обобщенного равновесия по Нэшу в программных стратегиях с непрерывным обновлением информации [221].

4.2.3.1 Дифференциальная модель логарифмической рекламной игры с непрерывным обновлением информации

В качестве наглядного примера рассмотрим логарифмическую модель дуополии с избыточной рекламой, предложенную Йоргенсенем в [222]. На рынке работают две фирмы. Предполагается, что рыночный потенциал остается постоянным во времени. Реклама — единственный маркетинговый инструмент, который используют фирмы. Реклама имеет убывающую отдачу из-за увеличения предельных затрат. Оптимальные по Нэшу программные рекламные стратегии были определены в [222]. Найдем равновесие по Нэшу в программных стратегиях с непрерывным обновлением информации, используя Теорему 4.2.2.

4.2.3.1.1 Модель исходной игры

Рассмотрим модель, исследованную в [222]. Пусть $x_i(t)$ обозначает скорость продаж фирмы i в момент времени t , $i = 1, 2$. Предположим, что $x_1 + x_2 = M$, т.е. в каждый момент времени потенциал рынка полностью истощается игроками. Игра протекает на временном интервале $[0, T]$, где T — произвольное, но фиксированное положительное число. В силу предположения $x_1 + x_2 = M$, имеем $\dot{x}_2 = -\dot{x}_1$. Уравнение состояния имеет вид

$$\begin{aligned} \dot{x}_1 &= k \log \frac{u_1}{u_2} = k(\log u_1 - \log u_2), \\ \dot{x}_2 &= -\dot{x}_1 = k(\log u_2 - \log u_1), \\ x_1(0) &= x_1^0, x_2(0) = x_2^0, \end{aligned} \quad (4.33)$$

где k — положительная константа, $x_i(0)$ — заданная начальная скорость продаж фирмы i . Уравнение состояния (4.33) описывает рынок, на котором покупатели совершенно мобильны и мгновенно переключаются на фирму с наибольшими рекламными расходами (т.е. на фирму, рекламирующую активнее оппонента). В этой модели доля рынка увеличивается линейно в зависимости от размера избыточной рекламы. Показатели эффективности задаются в виде

$$K_i = \int_0^T (\varphi_i x_i - u_i) \exp\{-r_i t\} dt, \quad i = 1, 2, \quad (4.34)$$

где $x_2 = M - x_1$. Предположим, что $r_i > 0$, $i = 1, 2$. Равновесие по Нэшу в программных стратегиях было построено в [222]:

$$u_i^{initial, NE} = \frac{k\varphi_i}{r_i} [1 - \exp\{-r_i(T - t)\}]. \quad (4.35)$$

В случае $r_1 = r_2$ оптимальные траектории находятся по формулам

$$\begin{aligned} x_1(t) &= (k \log \frac{\varphi_1}{\varphi_2})t + x_1(0), \\ x_2(t) &= M - (k \log \frac{\varphi_1}{\varphi_2})t - x_1(0). \end{aligned} \quad (4.36)$$

При $r_1 \neq r_2$ траектория x_1 удовлетворяет дифференциальному уравнению

$$\dot{x}_1 = k \log \frac{\varphi_1 r_2 [1 - \exp\{-r_1(T - t)\}]}{\varphi_2 r_1 [1 - \exp\{-r_2(T - t)\}]} \quad (4.37)$$

Решение уравнения (4.37) имеет вид

$$x_1(t) = x_1^0 + k \log \frac{\varphi_1 r_2}{\varphi_2 r_1} t + k \int_0^t \log \frac{1 - \exp\{-r_1(T - s)\}}{1 - \exp\{-r_2(T - s)\}} ds.$$

4.2.3.1.2 Модель игры с непрерывным обновлением информации

Теперь рассмотрим эту модель как игру с непрерывным обновлением информации. Предположим, что информация об уравнениях движения и функциях выигрыша непрерывно обновляется во времени. В каждый момент времени $t \in [0, +\infty)$ игроки имеют информацию только на временном интервале $[t, t + \bar{T}]$.

Следовательно, для каждого момента времени t , можно получить функцию выигрыша игрока i на временном интервале $[t, t + \bar{T}]$:

$$K_i^t = \int_t^{t+\bar{T}} (\varphi_i x_i^t - u_i^t) \exp(-r_i s) ds, \quad i = 1, 2.$$

Для упрощения решаемой задачи можно использовать замену переменных $\tau = \frac{s-t}{\bar{T}}$:

$$\begin{aligned} \dot{x}_1^t(\tau) &= \bar{T} k \log \frac{u_1^t(\tau, x)}{u_2^t(\tau, x)} = \bar{T} k (\log u_1^t(\tau) - \log u_2^t(\tau)), \quad \tau \in [0, 1], \\ \dot{x}_2^t(\tau) &= -\dot{x}_1^t(\tau), \\ x_1^t(0) &= x_1, \quad x_2^t(0) = x_2, \\ K_i^t &= \int_0^1 \bar{T} (\varphi_i x_i^t(\tau) - u_i^t(\tau, x)) \exp\{-r_i(\bar{T}\tau + t)\} d\tau, \quad i = 1, 2. \end{aligned} \quad (4.38)$$

Тогда гамильтонианы принимают вид

$$\begin{aligned} H_1^t(t, \tau, x, u^t, \lambda^t) &= (\varphi_1 x_1^t - u_1^t) \bar{T} + \lambda_1^t(\tau, x) \bar{T} k (\log u_1^t - \log u_2^t), \\ H_2^t(t, \tau, x, u^t, \lambda^t) &= (\varphi_2 x_2^t - u_2^t) \bar{T} - \lambda_2^t(\tau, x) \bar{T} k (\log u_1^t - \log u_2^t). \end{aligned}$$

Обращаем внимание, что текущий гамильтониан равен произведению $\exp(r_i(\bar{T}\tau + t))$ на стандартный гамильтониан. Необходимые условия оптимальности H_i^t при $u_i^t \in (0, +\infty)$ выражаются системой уравнений

$$\begin{aligned}\frac{\partial H_1^t}{\partial u_1^t} &= -\bar{T} + \lambda_1^t(\tau, x)\bar{T}k\frac{1}{u_1^t} = 0, \\ \frac{\partial H_2^t}{\partial u_2^t} &= -\bar{T} + \lambda_2^t(\tau, x)\bar{T}k\frac{1}{u_2^t} = 0.\end{aligned}$$

Следовательно, оптимальное управление u_i^t находится по формуле

$$u_1^t(\tau, x) = \lambda_1^t(\tau, x)k, \quad u_2^t(\tau, x) = \lambda_2^t(\tau, x)k. \quad (4.39)$$

Сопряженные переменные $\lambda_i^t(\tau)$ должны удовлетворять уравнениям

$$\begin{aligned}\dot{\lambda}_1^t(\tau, x) &= -\frac{\partial H_1^t}{\partial x_1^t} + \bar{T}r_1\lambda_1^t(\tau, x) = -\varphi_1\bar{T} + \bar{T}r_1\lambda_1^t(\tau, x), \\ \dot{\lambda}_2^t(\tau, x) &= -\frac{\partial H_2^t}{\partial x_2^t} + \bar{T}r_2\lambda_2^t(\tau, x) = -\varphi_2\bar{T} + \bar{T}r_2\lambda_2^t(\tau, x).\end{aligned} \quad (4.40)$$

Отметим, что эти уравнения несвязанные, и условия трансверсальности принимают вид

$$\lambda_i^t(1, x) = 0, \quad i = 1, 2.$$

Решая дифференциальные уравнения (4.40), находим:

$$\begin{aligned}\lambda_1^t(\tau, x) &= \frac{\varphi_1}{r_1}[1 - \exp\{\bar{T}r_1(\tau - 1)\}], \\ \lambda_2^t(\tau, x) &= \frac{\varphi_2}{r_2}[1 - \exp\{\bar{T}r_2(\tau - 1)\}].\end{aligned} \quad (4.41)$$

Подстановка выражения (4.41) в формулу (4.39) дает

$$u_i^{tNE}(\tau, x) = \frac{k\varphi_i}{r_i}[1 - \exp\{\bar{T}r_i(\tau - 1)\}]. \quad (4.42)$$

Отметим, что x является начальным состоянием в подыгре $\Gamma(x, t, t + \bar{T})$. В нашем примере программные стратегии $u_i^{tNE}(\tau, x)$ фактически не зависят от начального состояния x .

Покажем, что построенное решение удовлетворяет достаточным условиям. Поскольку $\frac{\partial^2 H_i^t}{\partial x^t \partial x^t} = 0$, $\frac{\partial^2 H_i^t}{\partial x^t \partial u_i^t} = 0$ и $\frac{\partial^2 H_i^t}{\partial u_i^t \partial u_i^t} = -\lambda_i^t(\tau) \bar{T} k \frac{1}{(u_i^t)^2} \leq 0$, согласно [223] ситуация $u^{tNE}(\tau, x)$ действительно образует равновесие по Нэшу в подыгре $\Gamma(x, t, t + \bar{T})$.

Наконец, при переходе от τ к исходным переменным t, s , получаем обобщенное равновесие по Нэшу в программных стратегиях с непрерывным обновлением информации вида

$$\begin{aligned}\tilde{u}_1^{NE}(t, s, x) &= \frac{k\varphi_1}{r_1} [1 - \exp\{r_1(s - t - \bar{T})\}], \\ \tilde{u}_2^{NE}(t, s, x) &= \frac{k\varphi_2}{r_2} [1 - \exp\{r_2(s - t - \bar{T})\}].\end{aligned}\tag{4.43}$$

Следуя Определению 4.2.2, построим равновесие по Нэшу в программных стратегиях с непрерывным обновлением информации:

$$u_i^{NE}(t, x) = \tilde{u}_i^{NE}(t, s, x)|_{s=t} = \frac{k\varphi_i}{r_i} [1 - \exp\{-r_i \bar{T}\}] \quad i = 1, 2.\tag{4.44}$$

Отметим, что в данном примере стратегии u_i^{tNE} не зависят от начальных значений переменных состояния подыгры $\Gamma(x, t, t + \bar{T})$. Следовательно, стратегии $u_i^{NE}(t, x)$ не зависят от x .

Рассмотрим различие между оптимальными стратегиями в исходной игре и игре с непрерывным обновлением информации:

$$u_i^{initial, NE} - u_i^{NE} = \frac{k\varphi_i}{r_i} \exp\{-r_i \bar{T}\} [1 - \exp\{-r_i(T - t - \bar{T})\}].$$

Очевидно, что рекламные расходы игроков меньше в игре с непрерывным обновлением информации при $t < T - \bar{T}$. Это происходит из-за того, что при непрерывном обновлении информации игроки не ведут себя рационально с учетом информации, доступной на всем временном интервале.

Соответствующие равновесные по Нэшу траектории $x_1^{NE}(t)$ и $x_2^{NE}(t)$ в игре с непрерывным обновлением информации представляют собой решения

дифференциальных уравнений

$$\begin{aligned}
 \dot{x}_1(t) &= k \log \left(\frac{\varphi_1 r_2 [1 - \exp\{-r_1 \bar{T}\}]}{r_1 \varphi_2 [1 - \exp\{-r_2 \bar{T}\}]} \right), \\
 \dot{x}_2(t) &= -\dot{x}_1(t), \\
 x_1(0) &= x_1^0, \\
 x_2(0) &= x_2^0,
 \end{aligned} \tag{4.45}$$

где $r_i > 0, i = 1, 2$. Таким образом, динамика состояния системы имеет вид

$$\begin{aligned}
 x_1^{NE}(t) &= x_1^0 + k \log \left(\frac{\varphi_1 r_2 [1 - \exp\{-r_1 \bar{T}\}]}{r_1 \varphi_2 [1 - \exp\{-r_2 \bar{T}\}]} \right) t, \\
 x_2^{NE}(t) &= M - x_1^0 - k \log \left(\frac{\varphi_2 r_1 [1 - \exp\{-r_2 \bar{T}\}]}{r_2 \varphi_1 [1 - \exp\{-r_1 \bar{T}\}]} \right) t.
 \end{aligned} \tag{4.46}$$

Отметим, что в случае $r_1 = r_2$ оптимальные траектории в исходной игре и игре с непрерывным обновлением информации совпадают. На рисунках 4.3 и 4.4 показаны результаты, полученные в рамках исходной игры и игры с непрерывным обновлением информации при следующих значениях параметров: $\frac{\varphi_1}{r_1} = 0.1, \frac{\varphi_2}{r_2} = 0.5, k = 1, T = 10, \bar{T} = 0.2, r_1 = 5, r_2 = 3, x_1^0 = 8, x_2^0 = 10$.

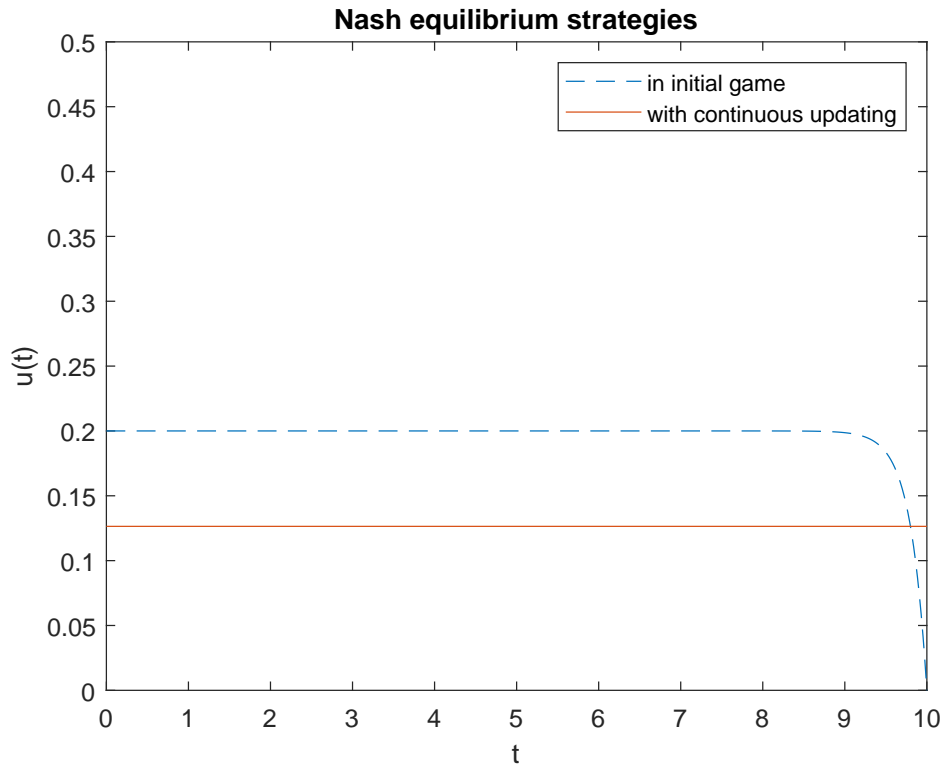


Рис. 4.3: Сравнение равновесных по Нэшу стратегий в исходной игре и игре с непрерывным обновлением информации.

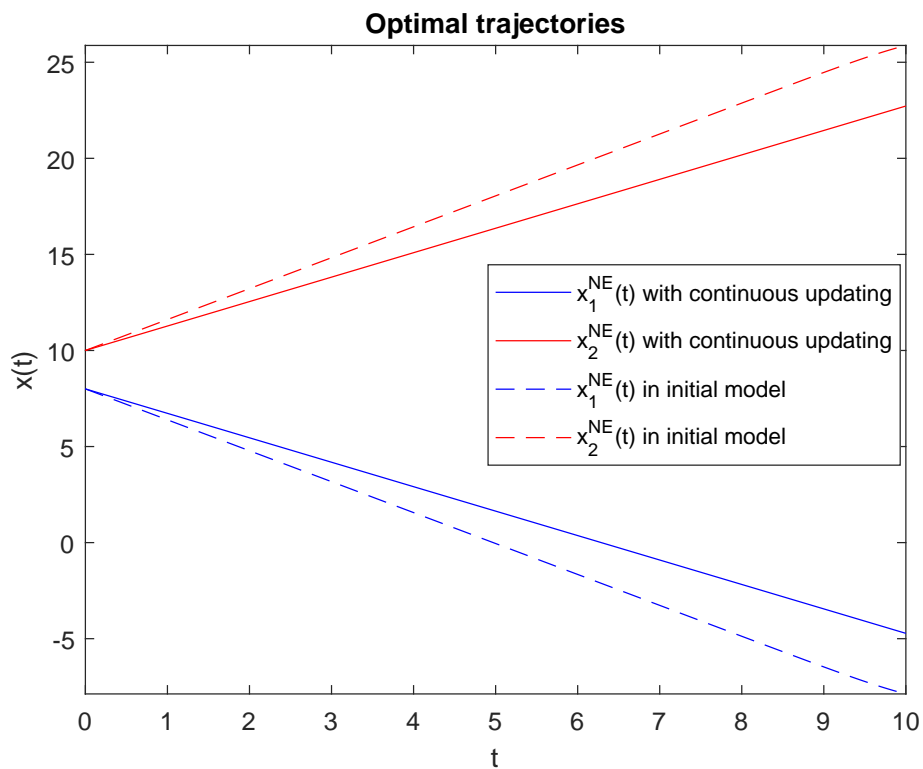


Рис. 4.4: Сравнение оптимальных траекторий в исходной игре и игре с непрерывным обновлением информации.

Очевидно, что скорость продаж для игрока 1 в игре с непрерывным обновлением информации меньше, чем в исходной игре. Это происходит из-за того, что при непрерывном обновлении информации игроки не являются рациональными с учетом информации, доступной на всем временном интервале.

4.2.4 Автономный случай линейно-квадратичной игры с непрерывным обновлением информации

В этом разделе рассмотрим только равновесие по Нэшу в позиционных стратегиях с непрерывным обновлением информации $u_{fb}^{NE}(t, x)$. Поэтому будем использовать упрощенное обозначение $u^{NE}(t, x)$. Здесь рассматривается особый класс дифференциальных игр с непрерывным обновлением информации, в которых уравнения движения представляют собой линейную систему, а функция выигрыша каждого игрока — квадратичный функционал.

4.2.4.1 Модель исходной автономной линейно-квадратичной игры

Рассмотрим линейно-квадратичную дифференциальную игру n лиц $\Gamma(x_0, T - t_0)$, определенную на временном интервале $[t_0, T]$. В этой игре $|N| = n$.

Уравнения движения имеют вид

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B_1 u_1(t, x) + \dots + B_n u_n(t, x), \\ x(t_0) &= x_0, \\ x \in \mathbb{R}^l, u &= (u_1, \dots, u_n), u_i = u_i(t, x) \in U_i \subset \text{comp} \mathbb{R}^k, t \in [t_0, T]. \end{aligned} \quad (4.47)$$

Определим функцию выигрыша игрока $i \in N$ следующим образом:

$$K_i(x_0, t_0, T; u) = \int_{t_0}^T \left(x'(t) Q_i x(t) + \sum_{j=1}^n u_j'(t, x) R_{ij} u_j(t, x) \right) dt, \quad i \in N, \quad (4.48)$$

где матрицы Q_i и R_{ij} являются симметричными, матрица R_{ii} положительно определена, а $(\cdot)'$ обозначает транспонирование.

4.2.4.2 Модель автономной линейно-квадратичной игры с непрерывным обновлением информации

Рассмотрим дифференциальную игру n лиц $\Gamma(x, t, \bar{T})$, $t \in [t_0, +\infty)$, определенную на временном интервале $[t, t + \bar{T}]$, где $0 < \bar{T} < +\infty$.

Уравнения движения в игре $\Gamma(x, t, \bar{T})$ имеют вид

$$\begin{aligned} \dot{x}^t(s) &= Ax^t(s) + B_1 u_1^t(s, x^t) + \dots + B_n u_n^t(s, x^t), \\ x^t(t) &= x, \\ x^t \in \mathbb{R}^l, \quad u^t &= (u_1^t, \dots, u_n^t), \quad u_i^t = u_i^t(s, x^t) \in U_i \subset \text{comp} \mathbb{R}^k, \quad t \in [t_0, +\infty). \end{aligned} \quad (4.49)$$

Определим функцию выигрыша игрока $i \in N$ в игре $\Gamma(x, t, \bar{T})$ следующим образом:

$$K_i^t(x^t, t, \bar{T}; u^t) = \int_t^{t+\bar{T}} \left((x^t(s))' Q_i x^t(s) + \sum_{j=1}^n (u_j^t(s, x^t))' R_{ij} u_j^t(s, x^t) \right) ds, \quad (4.50)$$

где $x^t(s)$ и $u^t(s, x)$ — траектория и стратегии в игре $\Gamma(x, t, \bar{T})$. Матрицы Q_i и R_{ij} удовлетворяют тем же условиям, что и в исходной модели автономной линейно-квадратичной дифференциальной игры.

4.2.4.3 Теорема о существовании равновесия по Нэшу с непрерывным обновлением информации

Ниже представлена явная форма равновесия по Нэшу для дифференциальной игры 2-х лиц с непрерывным обновлением информации.

Теорема 4.2.3 *При любом начальном состоянии линейно-квадратичная дифференциальная игра 2-х лиц $\Gamma(x_0, t_0, \bar{T})$, $N = \{1, 2\}$, обладает равновесием по Нэшу с непрерывным обновлением информации в линейных позиционных стратегиях тогда и только тогда, когда система связанных дифференциальных*

уравнений Риккати

$$\begin{aligned} \dot{K}_i(\tau) = & -(A\bar{T} - S_j K_j(\tau))' K_i(\tau) - K_i(\tau)(A\bar{T} - S_j K_j(\tau)) + \\ & + K_i(\tau) S_i K_i(\tau) - Q_i - K_j(\tau) S_{ji} K_j(\tau), \\ & K_i(1) = 0, \quad i \neq j \in N, \end{aligned} \quad (4.51)$$

где

$$S_i = \bar{T}^2 B_i R_{ii}^{-1} B_i', \quad S_{ij} = \bar{T}^2 B_i R_{ii}^{-1} R_{ji} R_{ii}^{-1} B_i', \quad i \neq j \in N, \quad (4.52)$$

имеет набор симметричных решений K_1, K_2 на временном интервале $[0, 1]$. В этом случае существует единственное равновесие по Нэшу с непрерывным обновлением информации в позиционных стратегиях, которое имеет вид

$$u_i^{NE}(t, x) = -R_{ii}^{-1} B_i' K_i(0) \bar{T} x, \quad i \in N. \quad (4.53)$$

Доказательство Для доказательства теоремы введем замену переменных

$$\begin{aligned} s &= t + \bar{T}\tau, \\ y(\tau) &= x(t + \bar{T}\tau), \\ v_i(\tau, y) &= u_i(t + \bar{T}\tau, x), \quad i \in N. \end{aligned} \quad (4.54)$$

Подставляя (4.54) в уравнения движения (4.5) и функцию выигрыша (4.50), получаем:

$$\dot{y}(\tau) = \bar{T} A y(s) + \bar{T} B_1 v_1(\tau, y) + \bar{T} B_2 v_2(\tau, y) \quad (4.55)$$

и

$$K_i(y, \tau; v) = \int_0^1 y'(s) Q_i y(s) + \sum_{j=1}^2 (v_j(s, y))' R_{ij} v_j(s, y) ds, \quad i \in N. \quad (4.56)$$

Известно [224], что критерием существования равновесия по Нэшу в позиционных стратегиях является наличие симметричных решений для системы дифференциальных уравнений (4.51). Согласно [224], равновесные по Нэшу позиционные стратегии имеют вид

$$v_i^{NE}(\tau, y) = -R_{ii}^{-1} B_i' K_i(\tau) \bar{T} y. \quad (4.57)$$

Из (4.54) находим

$$\tau = \frac{s-t}{\bar{T}}.$$

Возвращаясь к исходным переменным, получаем следующие стратегии:

$$u_i^t(s, x) = -R_{ii}^{-1} B_i' K_i \left(\frac{s-t}{\bar{T}} \right) \bar{T}x.$$

По построению они образуют равновесие по Нэшу в позиционных стратегиях в подыгре $\Gamma(x, t, \bar{T})$.

Игра (4.55), (4.56) и решение (4.57) совпадают по форме для всех значений t в исходной игре с непрерывным обновлением информации. Обобщенное равновесие по Нэшу с непрерывным обновлением информации имеет вид

$$\tilde{u}_i^{NE}(t, s, x) = -R_{ii}^{-1} B_i' K_i \left(\frac{s-t}{\bar{T}} \right) \bar{T}x. \quad (4.58)$$

Наконец, применим процедуру (4.11) для определения равновесия по Нэшу с непрерывным обновлением информации с помощью обобщенного равновесия по Нэшу с непрерывным обновлением информации (5.16), $s = t$:

$$u_i^{NE}(t, x) = -R_{ii}^{-1} B_i' K_i(0) \bar{T}x, \quad t \in [t_0, +\infty), \quad i \in N. \quad (4.59)$$

Теорема доказана. □

4.2.4.4 Линейно-квадратичная игра с динамическим обновлением информации

В этом разделе определим модель игры с динамическим обновлением информации. Это необходимо для того, чтобы ниже продемонстрировать сходимость равновесных по Нэшу стратегий и соответствующих траекторий для случая динамического и непрерывного обновления информации.

В статьях [76; 77; 108; 110; 111; 112; 116] был описан метод построения модели дифференциальной игры с динамическим обновлением информации, в предположении, что игроки имеют информацию о структуре игры только на усеченном временном интервале и принимают решения на основании этой информации.

Чтобы смоделировать поведение игроков в случае динамического обновления информации, будем считать, что информация обновляется с периодом $\Delta t > 0$, а поведение игроков на каждом временном отрезке $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, $j = 0, 1, 2, \dots$ описывается с помощью понятия усеченной подыгры.

Определение 4.2.3 Пусть $j = 0, 1, 2, \dots$. Усеченной подыгрой $\bar{\Gamma}_j(x_0^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ называется игра, определенная на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ следующим образом. На временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ функция выигрыша и уравнения движения в усеченной подыгре и исходной игре $\Gamma(x_0, T - t_0)$ совпадают:

$$\begin{aligned} \dot{x}^j(s) &= Ax^j(s) + B_1 u_1^j(s, x^j) + \dots + B_n u_n^j(s, x^j), \\ x^j(t_0 + j\Delta t) &= x_0^j, \\ x^j \in \mathbb{R}^n, \quad u^j &= (u_1^j, \dots, u_n^j), \quad u_i^j = u_i^j(s, x^j) \in U_i \subset \text{comp} \mathbb{R}^k, \quad t \in [t_0, +\infty); \end{aligned} \quad (4.60)$$

$$\begin{aligned} K_i^j(x^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; u^j) &= \int_{t_0 + j\Delta t}^{t_0 + j\Delta t + \bar{T}} (x^j(s))' Q_i x^j(s) \\ &+ \sum_{k=1}^n (u^k(s, x^j))' R_{ik} u^k(s, x^j) ds, \quad i \in N. \end{aligned} \quad (4.61)$$

В любой момент времени $t = t_0 + j\Delta t$ информация о структуре игры обновляется, а игроки адаптируются к ней. В качестве концепции решения дифференциальной игры с динамическим обновлением информации будем использовать равновесие по Нэшу в позиционных стратегиях с динамическим обновлением информации. Так же, как и в разделе 4.2.4.3, необходимо определить его особую форму. Ранее в разделе 3 понятие равновесия по Нэшу не вводилось напрямую, поскольку в основном рассматривался кооперативный случай игры с динамическим обновлением информации. Для определения решения игр с динамическим обновлением информации, введем понятие результирующего равновесия по Нэшу в позиционных стратегиях:

Определение 4.2.4 Результирующее равновесие по Нэшу в позиционных стратегиях

$$\hat{u}^{NE}(t, x) = (\hat{u}_1^{NE}(t, x), \dots, \hat{u}_n^{NE}(t, x))$$

в игре с динамическим обновлением информации имеет вид

$$\{\hat{u}^{NE}(t, x)\}_{t=t_0}^{\infty} = \begin{cases} u_0^{NE}(t, x), & t \in [t_0, t_0 + \Delta t], \\ \dots \\ u_j^{NE}(t, x), & t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t], \\ \dots \end{cases} \quad (4.62)$$

где $u_j^{NE}(t, x) = (u_1^{j,NE}(t, x), \dots, u_n^{j,NE}(t, x))$ – некоторое фиксированное равновесие по Нэшу в позиционных стратегиях в усеченной подыгре $\bar{\Gamma}_j(x_0^{j,NE}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, $j = 0, 1, 2, \dots$, начинающейся вдоль равновесной траектории предыдущей усеченной подыгры: $x_0^{j,NE} = x^{j-1,NE}(t_0 + j\Delta t)$.

Обозначим через $\hat{x}^{NE}(t)$ траекторию, полученную с помощью исходных линейно-квадратичных уравнений движения (4.47) и результирующего равновесия по Нэшу в позиционных стратегиях $\hat{u}^{NE}(t, x) = (\hat{u}_1^{NE}(t, x), \dots, \hat{u}_n^{NE}(t, x))$. Будем называть эту траекторию результирующей равновесной траекторией.

4.2.4.5 Результирующее равновесие по Нэшу в позиционных стратегиях с динамическим обновлением информации

Следующая теорема представляет необходимые и достаточные условия равновесия по Нэшу в линейно-квадратичной дифференциальной игре с непрерывным обновлением информации.

Теорема 4.2.4 При любом начальном состоянии линейно-квадратичная дифференциальная игра 2-х лиц $\bar{\Gamma}_j(x_0^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ обладает равновесием по Нэшу в линейных позиционных стратегиях тогда и только тогда, когда система связанных дифференциальных уравнений Риккати

$$\begin{aligned} \dot{K}_i(\tau) = & -(A\bar{T} - S_j K_j(\tau))' K_i(\tau) - K_i(\tau)(A\bar{T} - S_j K_j(\tau)) + \\ & + K_i(\tau) S_i K_i(\tau) - Q_i - K_j(\tau) S_{ji} K_j(\tau), \\ & K_i(1) = 0, \quad i \neq j, \end{aligned} \quad (4.63)$$

где

$$S_i = \bar{T}^2 B_i R_{ii}^{-1} B_i', \quad S_{ij} = \bar{T}^2 B_i R_{ii}^{-1} R_{ji} R_{ii}^{-1} B_i', \quad i \neq j \in N, \quad (4.64)$$

имеет набор симметричных решений K_1, K_2 на временном интервале $[0, 1]$. В этом случае существует единственное равновесие по Нэшу в позиционных стратегиях с непрерывным обновлением информации, которое имеет вид

$$u_i^{j,NE}(t, x) = -R_{ii}^{-1} B_i' K_i \left(\frac{t - (t_0 + j\Delta t)}{\bar{T}} \right) \bar{T}x. \quad (4.65)$$

Доказательство По аналогии с (4.54), введем следующую замену переменных для каждой усеченной подыгры:

$$\tau = \frac{t - (t_0 + j\Delta t)}{\bar{T}}. \quad (4.66)$$

Далее воспользуемся теми же рассуждениями, что и при доказательстве Теоремы 5.2.1. □

Согласно (4.62), равновесие по Нэшу $\hat{u}_i^{NE}(t, x)$ для игры с динамическим обновлением информации может быть построено с использованием равновесия по Нэшу $u_i^{j,NE}(t, x)$, определенного в каждой усеченной подыгре. Соответствующая траектория $\hat{x}^{NE}(t)$ строится с использованием $\hat{u}_i^{NE}(t, x)$ и (4.91).

4.2.4.6 Сходимость результирующих равновесных по Нэшу стратегий и траекторий

Теорема 4.2.5 При $\Delta t \rightarrow 0$ и $x \in X$, где множество X ограничено, результирующие равновесные по Нэшу позиционные стратегии $\hat{u}_i^{NE}(t, x)$ в игре с динамическим обновлением информации равномерно сходятся к равновесным по Нэшу позиционным стратегиям $\tilde{u}_i^{NE}(t, x)$ в игре с непрерывным обновлением информации:

$$\hat{u}_i^{NE}(t, x) \underset{[t_0, +\infty)}{\rightrightarrows} \tilde{u}_i^{NE}(t, x), \quad i \in N. \quad (4.67)$$

Доказательство Введем обозначение $t_j \stackrel{\text{def}}{=} t_0 + j\Delta t$ и положим $t \in [t_j, t_{j+1}]$ для некоторого j . Согласно определению $\hat{u}_i^{NE}(t, x)$ (4.62), необходимо показать,

что $\|\tilde{u}_i^{NE}(t, x) - u_i^{j, NE}(t, x)\| \rightarrow 0$ при $\Delta t \rightarrow 0$. Рассмотрим выражения для \tilde{u}_i^{NE} и $u_i^{j, NE}$:

$$\begin{aligned}\tilde{u}_i^{NE}(t, x) &= -R_{ii}^{-1} B_i' K_i(0) \bar{T} x, \\ u_i^{j, NE}(t, x) &= -R_{ii}^{-1} B_i' K_i\left(\frac{t-t_j}{\bar{T}}\right) \bar{T} x.\end{aligned}$$

Запишем неравенство для нормы разности:

$$\|\tilde{u}_i^{NE}(t, x) - u_i^{j, NE}(t, x)\| \leq \|R_{ii}^{-1} B_i'\| \|x\| \left(\left\| \dot{K}(0) \right\| \frac{\Delta t}{\bar{T}} + o(\Delta t) \right). \quad (4.68)$$

При $\Delta t \rightarrow 0$ правая часть неравенства (4.68) сходится к нулю; в результате, и левая часть (4.68) также сходится к нулю. Теорема доказана. \square

Теорема 4.2.6 *При $\Delta t \rightarrow 0$ равновесная траектория $\hat{x}^{NE}(t)$ в игре с динамическим обновлением информации поточечно сходится к равновесной траектории $\tilde{x}^{NE}(t)$ в игре с непрерывным обновлением информации:*

$$\hat{x}^{NE}(t) \xrightarrow{[t_0, +\infty)} \tilde{x}^{NE}(t). \quad (4.69)$$

Доказательство Пусть $t \in [t_j, t_{j+1}]$ для некоторого j . Согласно определению $\hat{x}^{NE}(t)$, необходимо показать, что $\|\tilde{x}^{NE}(t) - x_j^{NE}(t)\| \rightarrow 0$ при $\Delta t \rightarrow 0$.

Траектории $\tilde{x}^{NE}(t)$ и $x_j^{NE}(t)$ удовлетворяют соответственно дифференциальным уравнениям

$$\begin{aligned}\dot{\tilde{x}}(t) &= (A - B_1 R_{11}^{-1} B_1' K_1(0) \bar{T} - B_2 R_{22}^{-1} B_2' K_2(0) \bar{T}) \tilde{x}(t), \\ \dot{x}_j(t) &= \left(A - B_1 R_{11}^{-1} B_1' K_1\left(\frac{t-t_j}{\bar{T}}\right) \bar{T} - B_2 R_{22}^{-1} B_2' K_2\left(\frac{t-t_j}{\bar{T}}\right) \bar{T} \right) x_j(t).\end{aligned}$$

Отметим, что

$$K_i(0) \tilde{x} - K_i\left(\frac{t-t_j}{\bar{T}}\right) x_j = K_i(0) (\tilde{x} - x_j) + \left(K_i(0) - K_i\left(\frac{t-t_j}{\bar{T}}\right) \right) x_j.$$

Пусть $y_j^{NE}(t) = \tilde{x}^{NE}(t) - x_j^{NE}(t)$, $\tilde{A} = A - B_1 R_{11}^{-1} B_1' K_1(0) \bar{T} - B_2 R_{22}^{-1} B_2' K_2(0) \bar{T}$ и

$$f_j(t) = -B_1 R_{11}^{-1} B_1' \left[K_1(0) - K_1 \left(\frac{t - t_j}{\bar{T}} \right) \right] \bar{T} x_j(t) - \\ - B_2 R_{22}^{-1} B_2' \left[K_2(0) - K_2 \left(\frac{t - t_j}{\bar{T}} \right) \right] \bar{T} x_j(t).$$

Тогда $y_j^{NE}(t)$ удовлетворяет дифференциальному уравнению

$$\dot{y}_j(t) = \tilde{A} y_j(t) + f_j(t).$$

Рассмотрим

$$y(t) = \begin{cases} y_0(t), & t \in [t_0, t_0 + \Delta t], \\ \dots \\ y_j(t), & t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t], \\ \dots \end{cases} \quad (4.70)$$

и

$$f(t) = \begin{cases} f_0(t), & t \in [t_0, t_0 + \Delta t], \\ \dots \\ f_j(t), & t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t], \\ \dots \end{cases}$$

Тогда вектор (4.70) удовлетворяет дифференциальному уравнению

$$\dot{y}(t) = \tilde{A} y(t) + f(t)$$

с начальным условием $y(t_0) = 0$, поскольку $\hat{x}^{NE}(t_0) = \tilde{x}^{NE}(t_0) = x_0$. Используя формулу Коши, для любого $t \geq t_0$ получаем:

$$y(t) = \int_{t_0}^t e^{\tilde{A}(t-s)} f(s) ds.$$

С учетом этого, при фиксированном t можно записать

$$\lim_{\Delta t \rightarrow 0} \|y(t)\| \leq \lim_{\Delta t \rightarrow 0} \|e^{\tilde{A}(t-t_0)}\| (t-t_0)\beta \left(\frac{\Delta t}{T} + o(\Delta t) \right) = 0, \quad (4.71)$$

где

$$\beta = \left(\|B_1 R_{11}^{-1} B_1'\| \|\dot{K}_1(0)\| + \|B_2 R_{22}^{-1} B_2'\| \|\dot{K}_2(0)\| \right) \bar{T} M(t),$$

$$M(t) = \max_{\tau \in [t_0, t]} \|\hat{x}^{NE}(\tau)\|.$$

В силу (4.71) $y(t) \xrightarrow{[t_0, +\infty)} 0$ при $\Delta t \rightarrow 0$. Теорема доказана. \square

4.2.4.7 Общественный фонд знаний в дифференциальной игре с непрерывным обновлением информации

4.2.4.7.1 Модель исходной игры

Рассмотрим модель, в которой два индивидуума инвестируют в общественный фонд знаний (см. также статью Докнера и соавторов [50]). Пусть $x(t)$ — запас знаний в момент времени t , а $u_i(t)$ — инвестиция игрока i в общественное знание в момент времени t . Предположим, что запас знаний накапливается согласно дифференциальному уравнению

$$\dot{x}(t) = -\beta x(t) + u_1(t, x) + u_2(t, x), \quad x(0) = x_0, \quad (4.72)$$

где β — скорость устаревания знаний. Предположим, что каждый игрок получает квадратичную полезность от потребления запасов знаний, а инвестиционные затраты являются квадратичной возрастающей функцией инвестиционных усилий. То есть функция затрат обоих игроков имеет вид

$$K_i(x_0, t_0, T; u) = \int_0^T (-q_i x^2(t) + r_i u_i^2(t, x)) dt, \quad i = 1, 2. \quad (4.73)$$

Рассмотрим исходную игру (4.72), (4.73) в терминах теории линейно-квадратичных игр [224]. Для нахождения равновесия по Нэшу в позиционных стратегиях, необходимо решить следующую систему связанных дифференциальных урав-

нений Риккати:

$$\begin{cases} \dot{k}_1(t) = -2(-\beta - \frac{1}{r_2}k_2(t))k_1(t) + \frac{1}{r_1}k_1^2(t) + q_1, \\ \dot{k}_2(t) = -2(-\beta - \frac{1}{r_1}k_1(t))k_2(t) + \frac{1}{r_2}k_2^2(t) + q_2, \\ k_1(T) = 0, \\ k_2(T) = 0. \end{cases} \quad (4.74)$$

В качестве примера, рассмотрим симметричный случай $r_1 = r_2 = r$, $q_1 = q_2 = q$. Пусть $k(t) = k_1(t) = k_2(t)$. Получаем дифференциальное уравнение вида

$$\begin{cases} \dot{k}(t) = 2\beta k(t) + \frac{3k^2(t)}{r} + q, \\ k(T) = 0. \end{cases} \quad (4.75)$$

Решением задачи Коши (4.75) является функция

$$k(t) = \frac{\beta r + v}{3} \left(\frac{2v}{(v - \beta r)e^{\frac{2v}{r}(t-T)} + v + \beta r} - 1 \right),$$

где $v = \sqrt{\beta^2 r^2 - 3qr}$. Согласно [224], равновесие по Нэшу в позиционных стратегиях для исходной игры имеет вид

$$u_i^{NE}(t, x) = -\frac{k(t)x}{r}, \quad i = 1, 2. \quad (4.76)$$

Подставляя выражение для $k(t)$ в формулу (4.76), находим

$$u_i^{NE}(t, x) = \frac{\beta r + v}{3r} \left(1 - \frac{2v}{(v - \beta r)e^{\frac{2v}{r}(t-T)} + v + \beta r} \right) x(t).$$

4.2.4.7.2 Модель игры с непрерывным обновлением информации

Теперь рассмотрим случай с непрерывным обновлением информации. Предположим, что в каждый момент времени $t \in [t_0, +\infty)$ два индивидуума используют информацию об уравнениях движения и функциях выигрыша на временном интервале $[t, t + \bar{T}]$. По мере продвижения текущего времени t , временной интервал, на котором доступна информация, также смещается. Уравнения

движения для модели игры с непрерывным обновлением информации имеют вид

$$\dot{x}^t(s) = -\beta x^t(s) + u_1^t(s, x) + u_2^t(s, x), \quad x^t(t) = x, \quad t \in [t_0, +\infty). \quad (4.77)$$

Определим функцию выигрыша игрока $i \in N$ в игре с непрерывным обновлением информации следующим образом:

$$K_i^t(x^t, t, \bar{T}; u^t) = \int_t^{t+\bar{T}} \left(- (x^t(s))^2 q_i + (u_i^t(s, x))^2 r_i \right) ds, \quad i = 1, 2. \quad (4.78)$$

Согласно Теореме 4.2.4, определяющей явный вид равновесия по Нэшу в позиционных стратегиях с непрерывным обновлением информации, на первом шаге необходимо решить дифференциальное уравнение

$$\begin{cases} \dot{k}(\tau) = 2\beta\bar{T}k(\tau) + \frac{3\bar{T}k^2(\tau)}{r} + \bar{T}q, \\ k(1) = 0. \end{cases} \quad (4.79)$$

Решением уравнения (4.79) является функция

$$k(\tau) = \frac{\beta r + v}{3} \left(\frac{2v}{(v - \beta r)e^{\frac{2v\bar{T}}{r}(\tau-1)} + v + \beta r} - 1 \right), \quad (4.80)$$

где $v = \sqrt{\beta^2 r^2 - 3qr}$. Согласно (4.97), равновесие по Нэшу в позиционных стратегиях с непрерывным обновлением информации имеет вид

$$\tilde{u}_i^{NE}(t, x) = -\frac{k(0)x\bar{T}}{r}. \quad (4.81)$$

При подстановке выражения (4.80) в формулу (4.81), получаем:

$$\tilde{u}_i^{NE}(t, x) = \frac{\beta r + v}{3r} \left(1 - \frac{2v}{(v - \beta r)e^{-\frac{2v\bar{T}}{r}} + v + \beta r} \right) \bar{T}x. \quad (4.82)$$

Наконец, подставляя выражение (4.82) в уравнение (4.72), устанавливаем, что функция $\tilde{x}^{NE}(t)$ — решение дифференциального уравнения

$$\dot{\tilde{x}}^{NE}(t) = -\beta\tilde{x}^{NE}(t) + \tilde{u}_1^{NE}(t, x) + \tilde{u}_2^{NE}(t, x), \quad \tilde{x}^{NE}(0) = x_0. \quad (4.83)$$

4.2.4.7.3 Модель игры с динамическим обновлением информации

На основе подхода, описанного выше в Разделе 4.2.5.4.1, и вычислений для исходной игры, выполним аналогичные вычисления для результирующего равновесия по Нэшу в игре с динамическим обновлением информации. В результате получаем:

$$\tilde{u}_i^{NE}(t, x) = -\frac{k \left(\frac{t-t_i}{T} \right) x \bar{T}}{r}, \quad t \in [t_i, t_{i+1}]. \quad (4.84)$$

При подстановке выражения (4.80) в формулу (4.84) находим:

$$\hat{u}_i^{NE}(t, x) = \frac{\beta r + v}{3r} \left(1 - \frac{2v}{(v - \beta r) e^{\frac{2v(t-t_i-\bar{T})}{r}} + v + \beta r} \right) \bar{T} x, \quad t \in [t_i, t_{i+1}]. \quad (4.85)$$

Наконец, подставляя (4.85) в (4.72), устанавливаем, что функция $\hat{x}^{NE}(t)$ — решение дифференциального уравнения

$$\dot{\hat{x}}^{NE}(t) = -\beta\hat{x}^{NE}(t) + \hat{u}_1^{NE}(t, x) + \hat{u}_2^{NE}(t, x), \quad \hat{x}^{NE}(0) = x_0. \quad (4.86)$$

4.2.4.7.4 Модель игры на бесконечном временном интервале

Рассмотрим классический подход для равновесия по Нэшу в игре на бесконечном временном интервале $[0, +\infty)$. Уравнения движения имеют вид

$$\dot{x}(t) = -\beta x(t) + u_1(t, x) + u_2(t, x), \quad x(0) = x_0. \quad (4.87)$$

Определим функцию выигрыша игрока $i \in N$ следующим образом:

$$K_i(x_0; u) = \lim_{T \rightarrow \infty} \int_0^T (-q_i x^2(t) + r_i u_i^2(t, x)) dt, \quad i = 1, 2. \quad (4.88)$$

Согласно [224], в симметричном случае ($r_1 = r_2 = r$, $q_1 = q_2 = q$) равновесные по Нэшу позиционные стратегии имеют вид

$$u^{NE}(t, x) = -\frac{kx}{r}, \quad (4.89)$$

где k — решение квадратного уравнения

$$\frac{3k^2}{r} + 2\beta k + q = 0.$$

Подставляя выражение (4.89) в уравнения (4.87), устанавливаем, что $x^{NE}(t)$ — решение дифференциального уравнения

$$\dot{x}^{NE}(t) = \left(-\beta - \frac{2k}{r}\right) x^{NE}(t), \quad x^{NE}(0) = x_0. \quad (4.90)$$

4.2.4.7.5 Численное моделирование

Рассмотрим результаты численного моделирования для представленной выше игровой модели на временном интервале $[0, 8]$, т.е. $t_0 = 0$ и $T = 8$. В начальный момент времени $t_0 = 0$ запас знаний равен 100, т.е. $x_0 = 100$. Остальные параметры модели таковы: $\beta = 0.9$, $r = 6$, и $q = 1$. Предположим, что при динамическом обновлении информации (синие сплошные и пунктирные линии, см. рис.4.5-4.6) интервалы между моментами обновления информации равны $\Delta t = 2$, и следовательно $l = 4$. На рисунке 4.5 представлено сравнение результирующего равновесия по Нэшу в игре с динамическим обновлением информации (синяя линия) и равновесия по Нэшу с непрерывным обновлением информации (красные линии). На рисунке 4.6 представлены аналогичные результаты для соответствующих стратегий.

Для иллюстрации Теорем 4.2.5 и 4.2.6 о сходимости результирующих равновесных стратегий и соответствующей траектории к равновесным стратегиям и траектории с непрерывным обновлением информации, рассмотрим результаты моделирования для случая частого обновления информации, а именно $l = 20$. На рисунках 4.7–4.8 показаны те же решения, что и на рисунках 4.5–4.6, но для случая $\Delta t = 0,4$. Таким образом, сходимость стратегий и траектории

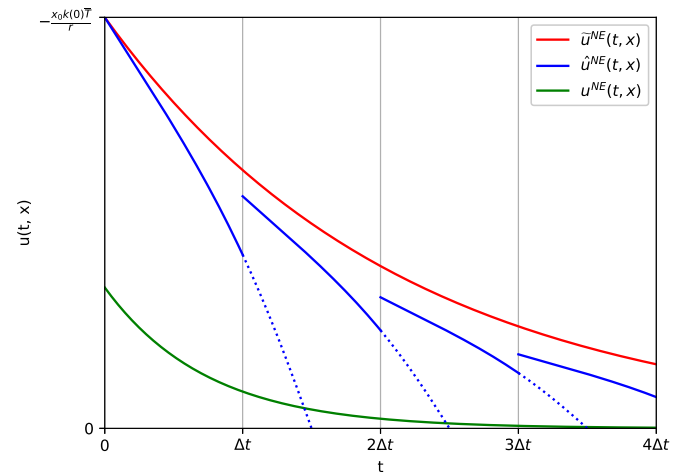
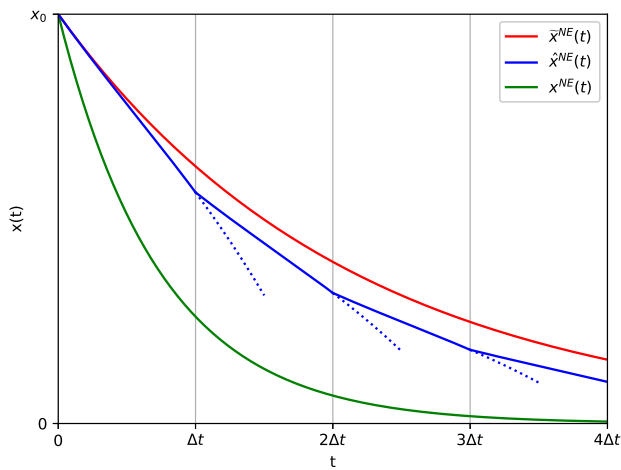


Рис. 4.5: $\tilde{x}^{NE}(t)$ (4.83) — красная линия сверху, $\hat{x}^{NE}(t)$ (4.86) — синяя пунктир-пунктирная линия, и $x^{NE}(t)$ (4.90) — зеленая линия снизу.

Рис. 4.6: $\tilde{u}^{NE}(t, x)$ (4.82) — красная линия сверху, $\hat{u}^{NE}(t, x)$ (4.85) — синяя пунктир-пунктирная линия, и $u^{NE}(t, x)$ (4.89) — зеленая линия снизу.

подтверждается численными экспериментами, представленными ниже.

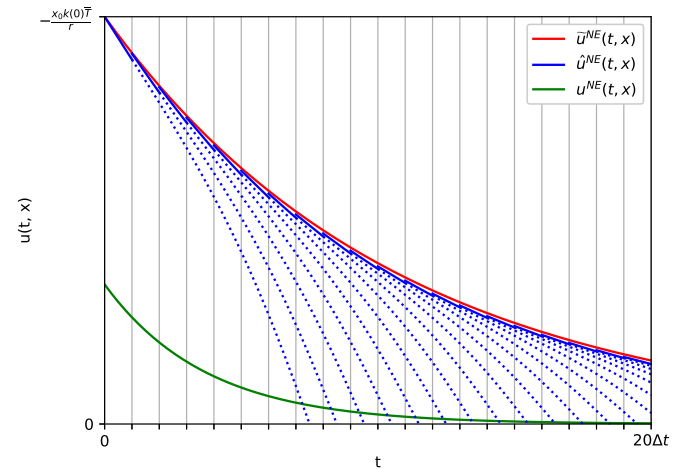
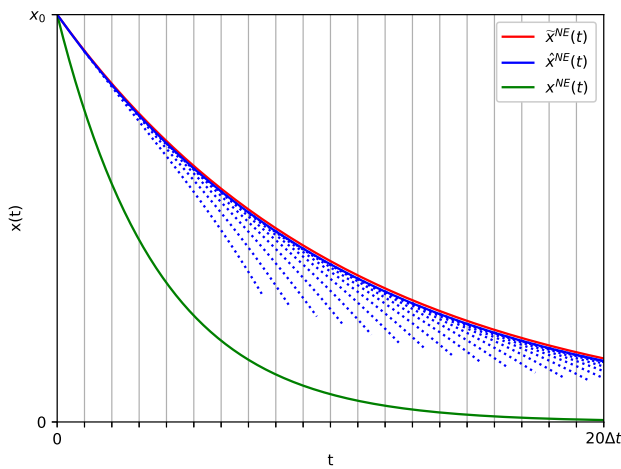


Рис. 4.7: $\tilde{x}^{NE}(t)$ (4.83) — красная линия сверху, $\hat{x}^{NE}(t)$ (4.86) — синяя пунктир-пунктирная линия, $x^{NE}(t)$ (4.90) — тирная линия, $u^{NE}(t)$ (4.89) — зеленая линия снизу.

Рис. 4.8: $\tilde{u}^{NE}(t, x)$ (4.82) — красная линия сверху, $\hat{u}^{NE}(t, x)$ (4.85) — синяя пунктир-пунктирная линия, $u^{NE}(t, x)$ (4.89) — зеленая линия снизу.

4.2.5 Неавтономный случай линейно-квадратичной игры с непрерывным обновлением информации

Здесь рассматривается особый тип неавтономной линейно-квадратичной дифференциальной модели игры с непрерывным обновлением информации. Это важный класс игровых постановок с непрерывным обновлением информации, позволяющий моделировать обновление уравнений движения и функций выигрыша через зависимость от текущего времени t .

4.2.5.1 Исходная неавтономная линейно-квадратичная игра

Рассмотрим неавтономную линейно-квадратичную дифференциальную игру n лиц $\Gamma(x_0, t_0, T)$, определенную на временном интервале $[t_0, T]$. В этой игре $|N| = n$.

Уравнения движения имеют вид

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B_1(t)u_1(t, x) + \dots + B_n(t)u_n(t, x), \\ x(t_0) &= x_0, \\ x \in \mathbb{R}^l, \quad u &= (u_1, \dots, u_n), \quad u_i = u_i(t, x) \in U_i \subset \text{comp} \mathbb{R}^k, \quad t \in [t_0, T]. \end{aligned} \quad (4.91)$$

Определим функцию выигрыша игрока $i \in N$ следующим образом:

$$K_i(x_0, t_0, T; u) = \int_{t_0}^T \left(x'(t)Q_i(t)x(t) + \sum_{j=1}^n u_j'(t, x)R_{ij}(t)u_j(t, x) \right) dt, \quad i \in N, \quad (4.92)$$

где матрицы $Q_i(t)$ и $R_{ij}(t)$ являются симметричными при $t \in [t_0, T]$, матрица R_{ii} положительно определена при $t \in [t_0, T]$, а $(\cdot)'$ обозначает транспонирование.

4.2.5.2 Неавтономная линейно-квадратичная игра с непрерывным обновлением информации

Для класса игр с непрерывным обновлением информации, отличие неавтономной модели от автономной состоит в том, что правая часть уравнений движения

и подынтегральная функция в функции выигрыша явно зависят от текущего времени t . Для случая линейно-квадратичной модели зависимость от текущего времени t в (4.91), (4.92) вводится с помощью матриц

$$A(t), B_i(t), Q_i(t), R_{ij}(t), i, j = 1, \dots, n. \quad (4.93)$$

Данная зависимость имеет особое значение. В начале игры, в момент времени t_0 , игроки обладают информацией только об уравнениях движения и функциях выигрыша на временном интервале $[t_0, t_0 + \bar{T}]$. Таким образом, игроки знают элементы матриц (4.93) лишь на временном интервале $[t_0, t_0 + \bar{T}]$, а на временном интервале $[t_0 + \bar{T}, +\infty]$ элементы матриц (4.93) могут измениться. В случае автономной системы, матрицы (4.93) являются константами, и можно сказать, что они известны на всем интервале определения игры.

Рассмотрим дифференциальную игру n лиц $\Gamma(x, t, t + \bar{T})$, $t \in [t_0, +\infty)$, определенную на временном интервале $[t, t + \bar{T}]$, где $0 < \bar{T} < +\infty$.

Уравнения движения в игре $\Gamma(x, t, t + \bar{T})$ имеют вид

$$\begin{aligned} \dot{x}^t(s) &= A(s)x^t(s) + B_1(s)u_1^t(s, x^t) + \dots + B_n(s)u_n^t(s, x^t), \\ x^t(t) &= x, \\ x^t \in \mathbb{R}^l, u^t &= (u_1^t, \dots, u_n^t), u_i^t = u_i^t(s, x^t) \in U_i \subset \text{comp} \mathbb{R}^k, t \in [t_0, +\infty). \end{aligned} \quad (4.94)$$

Определим функцию выигрыша игрока $i \in N$ в игре $\Gamma(x, t, t + \bar{T})$ следующим образом:

$$K_i^t(x, t, \bar{T}; u^t) = \int_t^{t+\bar{T}} \left((x^t(s))' Q_i(s) x^t(s) + \sum_{j=1}^n (u_j^t(s, x^t))' R_{ij}(s) u_j^t(s, x^t) \right) ds, \quad (4.95)$$

где $x^t(s)$ и $u^t(s, x^t)$ — соответственно траектория и стратегии в игре $\Gamma(x, t, t + \bar{T})$.

4.2.5.3 Теоремы о существовании равновесия по Нэшу с непрерывным обновлением информации

Одним из основных результатов этого раздела является установление достаточных условий для существования равновесия по Нэшу в позиционных стратегиях с непрерывным обновлением информации для неавтономного случая. Связанные результаты можно найти в статьях [79] и [117], где были представлены достаточные условия для равновесия по Нэшу с непрерывным обновлением информации для автономного случая.

Теорема 4.2.7 *Рассмотрим линейно-квадратичную дифференциальную игру N лиц $\Gamma(x_0, t_0, t_0 + \bar{T})$ с непрерывным обновлением информации и матрицами $Q_i(\cdot) \geq 0$ и $R_{ij}(\cdot) \geq 0$ ($i, j \in N, i \neq j$). Предположим, что система N связанных матричных дифференциальных уравнений Риккати*

$$\begin{aligned} \frac{dZ_i^t(\tau)}{d\tau} + Z_i^t(\tau)F^t(\tau) + (F^t(\tau))'Z_i^t(\tau) + Q_i(t + \bar{T}\tau) + \\ + \bar{T}^2 \sum_{j \in N} Z_j^t(\tau)B_j(t + \bar{T}\tau)R_{jj}^{-1}(t + \bar{T}\tau)R_{ij}(t + \bar{T}\tau) \times \\ \times R_{jj}^{-1}(t + \bar{T}\tau)B_j'(t + \bar{T}\tau)Z_j^t(\tau) = 0, \quad \tau \in [0, 1], \\ Z_i^t(1) = 0, \quad i \in N, \end{aligned} \quad (4.96)$$

где

$$F^t(\tau) = \bar{T}A(t + \bar{T}\tau) - \bar{T}^2 \sum_{i \in N} B_i(t + \bar{T}\tau)R_{ii}^{-1}(t + \bar{T}\tau)B_i'(t + \bar{T}\tau)Z_i^t(\tau),$$

имеет набор решений $Z_i^t(\cdot) \geq 0$, $i \in N$ при $t \geq t_0$. Тогда данная игра имеет равновесие по Нэшу в линейных позиционных стратегиях с непрерывным обновлением информации вида

$$u_{fb,i}^{NE}(t, x) = -R_{ii}^{-1}(t)B_i'(t)Z_i^t(0)\bar{T}x, \quad i \in N. \quad (4.97)$$

Доказательство Для доказательства теоремы введем замену переменных

$$\begin{aligned} s &= t + \bar{T}\tau, \\ y^t(\tau) &= x_{fb}^t(t + \bar{T}\tau), \\ v_i^t(\tau, y) &= u_{fb,i}(t + \bar{T}\tau, x), \quad i \in N. \end{aligned} \quad (4.98)$$

Подставляя (4.98) в уравнения движения (4.94) и функцию выигрыша (4.95), получаем:

$$\dot{y}^t(\tau) = \bar{T}A(t + \bar{T}\tau)y^t(\tau) + \sum_{i=1}^N \bar{T}B_i(t + \bar{T}\tau)v_i^t(\tau, y) \quad (4.99)$$

и

$$\begin{aligned} K_i^t(y^t, \tau; v^t) &= \int_0^1 (y^t(\tau))' Q_i(t + \bar{T}\tau) y^t(\tau) \\ &+ \sum_{j=1}^N (v_j^t(\tau, y))' R_{ij}(t + \bar{T}\tau) v_j^t(\tau, y) d\tau, \quad i \in N. \end{aligned} \quad (4.100)$$

Поскольку система дифференциальных уравнений (4.96) имеет набор решений $Z_i^t(\tau) \geq 0$, можно применить Следствие 6.5 из [224] (достаточные условия существования равновесия по Нэшу в аффинно-квадратичной игре). В результате приходим к равновесным по Нэшу позиционным стратегиям в подыгре $\Gamma(x, t, t + \bar{T})$, которые имеют вид

$$v_i^{t,NE}(\tau, y) = -R_{ii}^{-1}(t + \bar{T}\tau)B_i'(t + \bar{T}\tau)Z_i^t(\tau)\bar{T}y. \quad (4.101)$$

Из (4.98) находим

$$\tau = \frac{s - t}{\bar{T}}.$$

Возвращаясь к исходным переменным, получаем следующие стратегии:

$$u_{fb,i}^t(s, x) = -R_{ii}^{-1}(s)B_i'(s)Z_i^t\left(\frac{s - t}{\bar{T}}\right)\bar{T}x. \quad (4.102)$$

По построению они образуют равновесие по Нэшу в позиционных стратегиях в подыгре $\Gamma(x, t, t + \bar{T})$.

Задача (4.99), (4.100) и решение (4.101) совпадают по форме для всех значений t в исходной игре с непрерывным обновлением информации. Тогда

обобщенное равновесие по Нэшу в позиционных стратегиях с непрерывным обновлением информации имеет вид

$$\tilde{u}_{fb,i}^{NE}(t, s, x) = -R_{ii}^{-1}(s)B'_i(s)Z_i\left(\frac{s-t}{\bar{T}}\right)\bar{T}x. \quad (4.103)$$

Наконец, применим процедуру (4.11) для определения равновесия по Нэшу с непрерывным обновлением информации с помощью обобщенного равновесия по Нэшу с непрерывным обновлением информации (4.103), $s = t$:

$$u_{fb,i}^{NE}(t, x) = -R_{ii}^{-1}(t)B'_i(t)Z_i(0)\bar{T}x, \quad t \in [t_0, +\infty), \quad i \in N. \quad (4.104)$$

Теорема доказана. □

Явный вид равновесия по Нэшу в программных стратегиях с непрерывным обновлением информации был представлен в [117]. Эти результаты будут использованы ниже для анализа сходимости; см раздел 4.2.5.4.

Теорема 4.2.8 *Рассмотрим линейно-квадратичную дифференциальную игру N лиц с непрерывным обновлением информации и матрицами $Q_i(\cdot) \geq 0$, $R_{ij}(\cdot) \geq 0$ ($i, j \in N, i \neq j$). Предположим, что система N связанных матричных дифференциальных уравнений Риккати*

$$\begin{aligned} \frac{dM_i^t(\tau)}{d\tau} + \bar{T}M_i^t(\tau)A(t + \bar{T}\tau) + \bar{T}A'(t + \bar{T}\tau)M_i^t(\tau) + Q_i(t + \bar{T}\tau) - \\ - \bar{T}^2 M_i^t(\tau) \sum_{j \in N} B_j(t + \bar{T}\tau) (R_{jj}(t + \bar{T}\tau))^{-1} B_j'(t + \bar{T}\tau) M_j^t(\tau) = 0, \end{aligned} \quad (4.105)$$

$$M_i^t(1) = 0, \quad \tau \in [0, 1], \quad i \in N.$$

имеет набор решений $\{M_i^t, i \in N, t \geq t_0\}$. Тогда данная игра имеет равновесие по Нэшу в программных стратегиях с непрерывным обновлением информации вида

$$u_{ol,i}^{NE}(t, x) = -R_{ii}^{-1}(t)B'_i(t)M_i^t(0)\bar{T}x, \quad i \in N.$$

Замечание 2 *Обращаем внимание, что решение в программных стратегиях с непрерывным обновлением информации имеет форму обратной связи.*

Другими словами, равновесие по Нэшу в программных стратегиях с непрерывным обновлением информации явно зависит от текущего состояния. Это происходит из-за способа построения решения: в каждый текущий момент времени t игроки пересматривают свои решения в соответствии с непрерывно обновляемой информацией.

Пример 4.2.0.1 Рассмотрим следующую модель автономной дифференциальной игры 2-х лиц, чтобы сравнить равновесия по Нэшу в программных и позиционных стратегиях с непрерывным обновлением информации друг с другом. Пусть уравнения движения в подыгре имеют вид

$$\begin{aligned} \dot{x}^t(s) &= -\beta x^t(s) + u_1^t(s, x) + u_2^t(s, x), \\ x^t(t) &= x, \quad x^t \in \mathbb{R}^1, \end{aligned} \quad (4.106)$$

а функция выигрыша игрока $i \in \{1, 2\}$ определена следующим образом:

$$K_i(x, t, \bar{T}, u^t) = \int_t^{t+\bar{T}} \left(q (x^t(s))^2 + r_1 (u_i^t(s, x^t))^2 + r_2 (u_j^t(s, x^t))^2 \right) ds, \quad (4.107)$$

где $i, j \in \{1, 2\}$, $i \neq j$, и $q, r_1, r_2 > 0$.

В силу симметрии игроков, можно рассматривать только одно уравнение. Для равновесия по Нэшу в позиционных стратегиях с непрерывным обновлением информации, дифференциальное уравнение Риккати имеет вид

$$\begin{aligned} \dot{z}(\tau) &= 2\bar{T}\beta z(\tau) + 2\bar{T}^2 z^2(\tau) \left(\frac{2}{r_1} - \frac{r_2}{r_1^2} \right) - q, \quad \tau \in [0, 1], \\ z(1) &= 0. \end{aligned}$$

Для равновесия по Нэшу в программных стратегиях с непрерывным обновлением информации, получаем:

$$\begin{aligned} \dot{m}(\tau) &= 2\bar{T}\beta m(\tau) + 2\bar{T}^2 \frac{m^2(\tau)}{r_1} - q, \quad \tau \in [0, 1], \\ m(1) &= 0. \end{aligned}$$

Отметим важное обстоятельство: поскольку в этом примере мы рассмат-

риваем случай двух игроков, условия Теорем 4.2.7 и 4.2.8 также становятся необходимыми. Таким образом, можно сказать, что равновесия по Нэшу в позиционных и программных стратегиях с непрерывным обновлением информации являются единственными для этой модели. Эти равновесия были смоделированы численно с помощью приведенных выше уравнений Риккати для параметров $\beta = 0,01$, $r_1 = 1$, $r_2 = 0,1$, $q = 5$, $\bar{T} = 1$, $t \in [0, 2]$, и $x(0) = 100$. Графики полученных равновесий в программных и позиционных стратегиях, а также соответствующие равновесные траектории представлены соответственно на рис. 4.9 и рис. 4.10. Очевидно, что решения явно отличаются, несмотря на то, что оба фактически имеют обратную связь (зависят от состояния x).

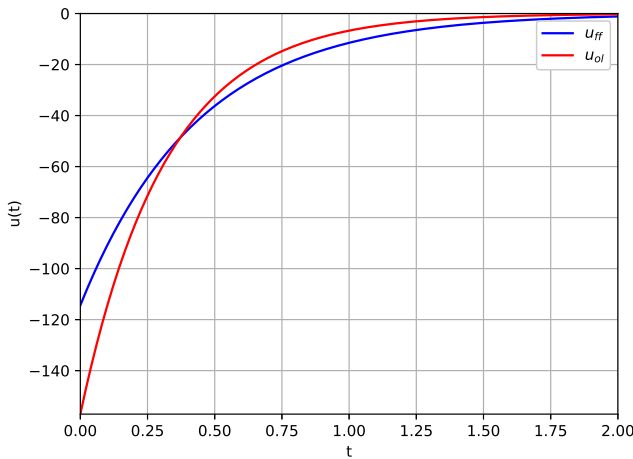


Рис. 4.9: $u_{fb}^{NE}(t)$ — равновесие по Нэшу в позиционных стратегиях с непрерывным обновлением информации для (4.106), (4.107), $u_{ol}^{NE}(t)$ — равновесие по Нэшу в программных стратегиях с непрерывным обновлением информации для (4.106), (4.107).

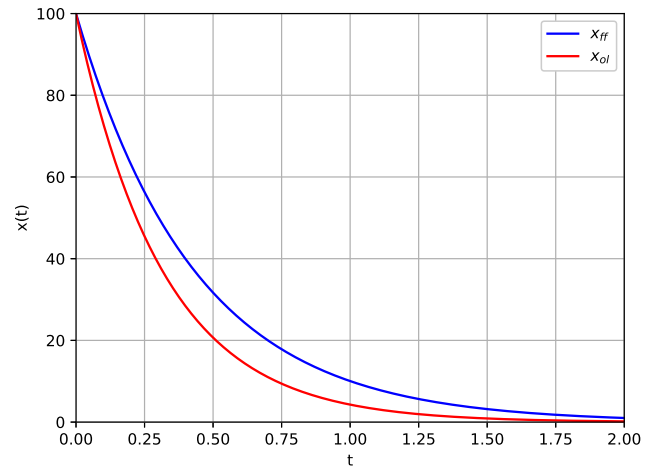


Рис. 4.10: $x_{fb}^{NE}(t)$ — равновесная по Нэшу траектория в позиционных стратегиях с непрерывным обновлением информации для (4.106), (4.107), $x_{ol}^{NE}(t)$ — равновесная по Нэшу траектория в программных стратегиях с непрерывным обновлением информации для (4.106), (4.107).

4.2.5.3.1 Существование, единственность и непрерывность

Важным предположением, сделанным в предыдущем разделе, является предположение о существовании равновесия по Нэшу с непрерывным обновлением информации. Согласно определению, для построения равновесия по Нэшу с

непрерывным обновлением информации $u^{NE}(t, x)$ необходимо иметь равновесие по Нэшу $\tilde{u}^{NE}(t, s, x_t)$ в каждой подыгре $\Gamma(x, t, t + \bar{T})$, $t \geq t_0$.

Проблема единственности равновесия по Нэшу в дифференциальных играх нетривиальна. Известно [225], что для простейшей автономной линейно-квадратичной игры равновесие по Нэшу в программных стратегиях может быть неединственным даже при выполнении достаточных условий, аналогичных приведенным выше теоремам.

Предположим, что в каждой подыгре $\Gamma(x, t, t + \bar{T})$, $t \geq t_0$, игры с непрерывным обновлением информации существует не менее двух равновесий по Нэшу на временном интервале $[t, t + \bar{T}]$. Тогда множество равновесий по Нэшу с непрерывным обновлением информации будет несчетным, поскольку игроки могут переключаться с одного равновесия по Нэшу на другое в каждый текущий момент времени t .

Пример 4.2.0.2 Рассмотрим игру с непрерывным обновлением информации, заданную уравнениями движения

$$\dot{x}^t(s) = \begin{bmatrix} u_1(s) & 0 \\ 0 & u_2(s) \end{bmatrix},$$

$$x^t(t) = x, \quad x^t \in \mathbb{R}^2,$$

с программными стратегиями игроков и функцией выигрыша игрока $i \in \{1, 2\}$ вида

$$K_i^t \left(x, t, \frac{\pi}{2}, u^t \right) = \frac{1}{2} \int_t^{t+\frac{\pi}{2}} \left((x^t(s))' Q_i x^t(s) + (u_i^t(s))^2 \right) ds,$$

где

$$Q_1 = \begin{bmatrix} \frac{1}{2} & -1 \\ -1 & 4 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 15 & -2 \\ -2 & \frac{1}{3} \end{bmatrix}.$$

При $x = (0, 0)'$, согласно Примеру 4.1 из [225], эта дифференциальная подыгра имеет тривиальное решение $u_1^t = u_2^t \equiv 0$ и множество нетривиальных решений $u_1^t = \frac{2}{3}$, $u_2^t \equiv \alpha$, $\alpha \in \mathbb{R}$. Таким образом, множество равновесий по Нэшу в программных стратегиях с непрерывным обновлением информации

образуют кусочно-непрерывные функции вида

$$u_{ol,i}(t) = \begin{cases} 0, & t \in [t_0, t_1), \\ u_i^*(t), & t \in [t_1, \infty), \end{cases}$$

где t_1 — точка переключения с тривиального решения на нетривиальное, а $u_i^*(t)$ — некоторая нетривиальная стратегия игрока i после этой точки переключения. Очевидно, что множество точек переключения является несчетным. Следовательно, множество равновесий по Нэшу в программных стратегиях с непрерывным обновлением информации также несчетно.

Более того, равновесие по Нэшу в программных стратегиях с непрерывным обновлением информации может быть непрерывным или иметь точку разрыва в некоторый момент времени t_1 , когда игроки переключают свои стратегии с тривиальных на нетривиальные. Таким образом, если имеется несколько равновесий по Нэшу в подыграх, то в игре с непрерывным обновлением информации равновесие по Нэшу с непрерывным обновлением информации может быть разрывной функцией.

Лемма 4.2.1 *Предположим, что дифференциальная игра автономна, а равновесие по Нэшу в каждой подыгре $\Gamma(x, t, t + \bar{T})$ единственно и непрерывно. Тогда равновесие по Нэшу с непрерывным обновлением информации также единственно и непрерывно.*

Доказательство Единственность равновесия по Нэшу с непрерывным обновлением информации следует из его определения и единственности равновесия по Нэшу в каждой подыгре $\Gamma(x, t, t + \bar{T})$.

В автономном случае игроки рассматривают одну и ту же подыгру $\Gamma(x, t, t + \bar{T})$ в каждый момент времени t . Без ограничения общности, пусть $u^*(s, x)$ — равновесие по Нэшу в этой общей подыгре $\Gamma(x, t_0, t_0 + \bar{T})$. Тогда, согласно Определению 4.2.2, равновесие по Нэшу с непрерывным обновлением информации имеет вид $u^{NE}(t, x) = u^*(t_0, x)$. Таким образом, равновесие по Нэшу с непрерывным обновлением информации не зависит явно от времени, а его непрерывность следует из непрерывности равновесия по Нэшу в подыгре $\Gamma(x, t, t + \bar{T})$. \square

4.2.5.4 Сходимость стратегий и траекторий

Понятно, что в реальных задачах невозможно обрабатывать новую информацию об игре непрерывно и мгновенно. По крайней мере, требуется время, чтобы получить новую информацию, которая к этому моменту времени потеряет свою актуальность из-за непрерывного обновления игры.

Ниже приведены результаты относительно сходимости равновесных по Нэшу стратегий и траекторий с непрерывным обновлением информации. Для этого представлена концепция динамического обновления информации, которая помогает моделировать некооперативное поведение игроков при обновлении информации в дискретные моменты времени. Доказана сходимость стратегий и траекторий с динамическим обновлением информации к соответствующим стратегиям и траекториям с непрерывным обновлением, как в случае программных стратегий, так и в случае позиционных стратегий.

4.2.5.4.1 Модель линейно-квадратичной игры с динамическим обновлением информации

В статьях [76; 77; 108; 110; 111; 112; 116] был описан метод построения модели дифференциальной игры с динамическим обновлением информации, в предположении, что игроки имеют информацию о структуре игры только на усеченном временном интервале и принимают решения на основании этой информации. Чтобы смоделировать поведение игроков в случае динамического обновления информации, будем считать, что информация обновляется с периодом $\Delta t > 0$, а поведение игроков на каждом временном отрезке $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, $j = 0, 1, 2, \dots$ описывается с помощью понятия усеченной подыгры.

В этом разделе рассмотрим позиционные и программные стратегии для случая динамического обновления информации. Поэтому будем использовать для них соответственно обозначения $v_{fb}^{NE}(t, x)$ и $v_{ol}^{NE}(t, x)$.

Определение 4.2.5 Пусть $j = 0, 1, 2, \dots$. Усеченной подыгрой $\bar{\Gamma}_j(x_0^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ называется игра, определенная на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ следующим образом. На временном интервале $[t_0 +$

$j\Delta t, t_0 + j\Delta t + \bar{T}]$ функция выигрыша и уравнения движения в усеченной подыгре и исходной игре $\Gamma(x_0, t_0, T)$ совпадают:

$$\begin{aligned} \dot{x}^j(s) &= A(s)x^j(s) + B_1(s)v_1^j(s, x^j) + \dots + B_n(s)v_n^j(s, x^j), \\ x^j(t_0 + j\Delta t) &= x_0^j, \\ x^j \in \mathbb{R}^n, v^j &= (v_1^j, \dots, v_n^j), v_i^j = v_i^j(s, x^j) \in U_i \subset \text{comp}\mathbb{R}^k, t \in [t_0, +\infty); \end{aligned} \quad (4.108)$$

$$\begin{aligned} K_i^j(x^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; v^j) &= \int_{t_0 + j\Delta t}^{t_0 + j\Delta t + \bar{T}} (x^j(s))' Q_i(s) x^j(s) \\ &+ \sum_{k=1}^n (v^k(s, x^j))' R_{ik}(s) v^k(s, x^j) ds, \quad i \in N. \end{aligned} \quad (4.109)$$

В любой момент времени $t = t_0 + j\Delta t$ информация о структуре игры обновляется, а игроки адаптируются к ней. Данный класс игровых моделей будем называть дифференциальными играми с динамическим обновлением информации.

Так же, как и в разделе 4.2.5.3, необходимо определить особую форму равновесия по Нэшу. Согласно описанному выше подходу, в любой момент времени $t \in [t_0, +\infty)$ игроки обладают или используют усеченную информацию о структуре игры. Поэтому классические подходы для определения оптимальных стратегий напрямую не применимы. Для решения игр с динамическим обновлением информации введем понятие *равновесия по Нэшу в позиционных (программных) стратегиях с динамическим обновлением информации*:

Определение 4.2.6 *Равновесие по Нэшу в позиционных (программных) стратегиях с динамическим обновлением информации*

$$\begin{aligned} v_{fb}^{NE}(t, x) &= (v_{fb,1}^{NE}(t, x), \dots, v_{fb,n}^{NE}(t, x)) \\ (v_{ol}^{NE}(t, x) &= (v_{ol,1}^{NE}(t, x), \dots, v_{ol,n}^{NE}(t, x))) \end{aligned}$$

в игре с динамическим обновлением информации имеет вид

$$\begin{aligned} \{v_{fb}^{NE}(t, x)\}_{t=t_0}^{\infty} &= \tilde{v}_{fb,j}^{NE}(t, x), \quad t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t], \quad j = 0, 1, 2, \dots \\ (\{v_{ol}^{NE}(t, x)\}_{t=t_0}^{\infty} &= \tilde{v}_{ol,j}^{NE}(t, x), \quad t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t], \quad j = 0, 1, 2, \dots) \end{aligned} \quad (4.110)$$

где $\tilde{v}_{fb,j}^{NE}(t, x) = (\tilde{v}_{fb,1}^{j,NE}(t, x), \dots, \tilde{v}_{fb,n}^{j,NE}(t, x))$ ($\tilde{v}_{ol,j}^{NE}(t, x) = (\tilde{v}_{ol,1}^{j,NE}(t, x), \dots, \tilde{v}_{ol,n}^{j,NE}(t, x))$) — некоторое фиксированное равновесие по Нэшу в позиционных (программных) стратегиях в усеченной подыгре $\bar{\Gamma}_j(x_0^{j,NE}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, $j = 0, 1, 2, \dots$, начинающейся вдоль равновесной траектории предыдущей усеченной подыгры: $x_0^{j,NE} = x^{j-1,NE}(t_0 + j\Delta t)$.

Важно отметить, что равновесие по Нэшу с динамическим обновлением информации $v^{NE}(t, x)$ не является равновесием по Нэшу в классическом смысле, но может использоваться как концепция решения, связанная с равновесием по Нэшу, для класса игр с динамическим обновлением информации. Соответствующая траектория $\hat{x}_{fb}^{NE}(t)$ ($\hat{x}_{ol}^{NE}(t)$) получается с помощью уравнений движения (4.91) и равновесия по Нэшу в позиционных (программных) стратегиях с динамическим обновлением информации $v^{NE}(t, x) = (v_1^{NE}(t, x), \dots, v_n^{NE}(t, x))$.

4.2.5.4.2 Результирующее равновесие по Нэшу с динамическим обновлением информации

Другими важными результатами диссертации являются достаточные условия для существования равновесия по Нэшу в позиционных и программных стратегиях с динамическим обновлением информации для неавтономного случая.

Теорема 4.2.9 *Рассмотрим линейно-квадратичную дифференциальную игру N лиц $\bar{\Gamma}_j(x_0^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ с динамическим обновлением информации и матрицами $Q_i(\cdot) \geq 0$, $R_{ik}(\cdot) \geq 0$ ($i, k \in N, i \neq k$). Предположим, что система*

N связанных матричных дифференциальных уравнений Риккати

$$\begin{aligned} & \frac{dZ_i^{t_j}(\tau)}{d\tau} + Z_i^{t_j}(\tau)F^{t_j}(\tau) + (F^{t_j}(\tau))' Z_i^{t_j}(\tau) + Q_i(t_j + \bar{T}\tau) + \\ & + \bar{T}^2 \sum_{k \in N} Z_k^{t_j}(\tau) B_k(t_j + \bar{T}\tau) R_{kk}^{-1}(t_j + \bar{T}\tau) R_{ik}(t_j + \bar{T}\tau) \times \\ & \times R_{kk}^{-1}(t_j + \bar{T}\tau) B_k'(t_j + \bar{T}\tau) Z_k^{t_j}(\tau) = 0, \\ & Z_i^{t_j}(1) = 0, \quad \tau \in [0, 1], \quad i \in N, \end{aligned} \quad (4.111)$$

где

$$F^{t_j}(\tau) = \bar{T}A(t_j + \bar{T}\tau) - \bar{T}^2 \sum_{i \in N} B_i(t_j + \bar{T}\tau) R_{ii}^{-1}(t_j + \bar{T}\tau) B_i'(t_j + \bar{T}\tau) Z_i^{t_j}(\tau),$$

имеет набор решений $Z_i^{t_j}(\cdot) \geq 0, i \in N$, при $t \geq t_0 + j\Delta t$.

Тогда данная игра имеет равновесие по Нэшу в позиционных стратегиях с динамическим обновлением информации вида

$$v_{fb,i}^{NE}(t, x) = -R_{ii}^{-1}(t) B_i'(t) Z_i^{t_j} \left(\frac{t - (t_0 + j\Delta t)}{\bar{T}} \right) \bar{T}x, \quad (4.112)$$

где $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, 1, 2, \dots$, $i \in N$.

Доказательство Доказательство этого результата аналогично доказательству Теоремы 4.2.7. Существенным отличием является временной интервал совпадения решений во всей игре и в подыгре. С одной стороны, равновесие по Нэшу с непрерывным обновлением информации совпадает с равновесием по Нэшу в подыгре $\Gamma(x, t, t + \bar{T})$ только в одной точке t . С другой стороны, равновесие по Нэшу с динамическим обновлением информации совпадает с равновесием по Нэшу в подыгре $\bar{\Gamma}_j(x_0^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ на временном интервале $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$. Таким образом, доказательство этой теоремы повторяет доказательство Теоремы 4.2.7 вплоть до (4.102).

В игре с динамическим обновлением информации, мы обновляем информацию только в дискретные моменты времени $t_j = t_0 + j\Delta t$, $j \in \mathbb{N}$, и указываем

(4.102) для игры с динамическим обновлением информации

$$v_{fb,i}^{t_j}(s, x^t) = -R_{ii}^{-1}(s)B'_i(s)Z_i^{t_j} \left(\frac{s - t_j}{\bar{T}} \right) \bar{T}x^t(s),$$

а это выражение, в свою очередь, совпадает с (4.112). \square

Теорема 4.2.10 *Рассмотрим линейно-квадратичную дифференциальную игру N лиц $\bar{\Gamma}_j(x_0^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ с динамическим обновлением информации и матрицами $Q_i(\cdot) \geq 0$, $R_{ik}(\cdot) \geq 0$ ($i, k \in N, i \neq k$). Предположим, что существует набор решений $\{M_i^{t_j}, i \in N, t_j \geq t_0\}$ для системы N связанных матричных дифференциальных уравнений Риккати*

$$\begin{aligned} & \frac{dM_i^{t_j}(\tau)}{d\tau} + \bar{T}M_i^{t_j}(\tau)A(t_j + \bar{T}\tau) + \bar{T}A'(t_j + \bar{T}\tau)M_i^{t_j}(\tau) + Q_i(t_j + \bar{T}\tau) - \\ & - \bar{T}^2 M_i^{t_j}(\tau) \sum_{k \in N} B_k(t_j + \bar{T}\tau) (R_{kk}(t_j + \bar{T}\tau))^{-1} B'_k(t_j + \bar{T}\tau) M_k^{t_j}(\tau) = 0, \end{aligned} \quad (4.113)$$

$$M_i^{t_j}(1) = 0, \quad \tau \in [0, 1], \quad i \in N.$$

Тогда данная игра имеет равновесие по Нэшу в программных стратегиях с динамическим обновлением информации вида

$$\begin{aligned} v_{ol,i}^{NE}(t, x) = & -R_{ii}^{-1}(t)B'_i(t)M_i^{t_j} \left(\frac{t - (t_0 + j\Delta t)}{\bar{T}} \right) \times \\ & \times \Phi^{t_j} \left(\frac{t - (t_0 + j\Delta t)}{\bar{T}} \right) \bar{T}x(t_0 + j\Delta t), \quad j = 0, 1, 2, \dots, \quad i \in N, \end{aligned}$$

где $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$ и $i \in N$, а Φ^{t_j} — решение дифференциального уравнения

$$\frac{d\Phi^t}{d\tau} = \left(A(t + \bar{T}\tau) - \sum_{i \in N} B_i(t + \bar{T}\tau)R_{ii}^{-1}(t + \bar{T}\tau)B'_i(t + \bar{T}\tau) \right) \Phi^t(\tau),$$

$$\Phi^t(0) = E.$$

Доказательство Этот результат может быть доказан с помощью Теоремы 4.2.8, по аналогии с доказательством Теоремы 4.2.9 с помощью Теоремы 4.2.7; см выше. \square

4.2.5.4.3 Сходимость результирующих равновесных по Нэшу стратегий и траектории

Теперь покажем сходимость равновесия по Нэшу и соответствующей равновесной траектории в дифференциальной игре с динамическим обновлением информации к равновесию по Нэшу и соответствующей траектории в дифференциальной игре с непрерывным обновлением информации.

Лемма 4.2.2 Пусть некоторая матричная функция $U(t)$ равномерно ограничена при $t \geq t_0$, некоторая параметрическая матричная функция $P^t(\tau)$ непрерывно дифференцируема как по t , так и по τ , а производные функции $\frac{dP^t(\tau)}{dt}$ и $\frac{dP^t(\tau)}{d\tau}$ равномерно ограничены при $t \geq t_0$ и $\tau \in [0, 1]$.

Тогда при $x \in X$, где X — ограниченное множество, $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$ и $\Delta t \leq \bar{T}$, имеет место следующая сходимость при $\Delta t \rightarrow 0$:

$$U(t)P^t(0)x \underset{[t_0, +\infty)}{\rightrightarrows} U(t)P^{t_0+j\Delta t} \left(\frac{t-t_0+j\Delta t}{\bar{T}} \right) x.$$

Доказательство Введем обозначение $t_j \stackrel{\text{def}}{=} t_0 + j\Delta t$. Тогда $t \in [t_j, t_{j+1}]$, и необходимо показать, что $\left\| U(t)P^t(0)x - U(t)P^{t_j} \left(\frac{t-t_j}{\bar{T}} \right) x \right\| \rightarrow 0$ при $\Delta t \rightarrow 0$.

Разлагая функцию $P^t(\tau)$ в ряд Тейлора в окрестности точки $t = t_j$, имеем:

$$P^t(\tau) = P^{t_j}(\tau) + \left. \frac{dP^t(\tau)}{dt} \right|_{t=t_j} (t - t_j) + o(t - t_j).$$

Аналогично, при разложении функции $P^{t_j}(\tau)$ в ряд Тейлора в окрестности точки $\tau = 0$ получаем:

$$P^{t_j}(\tau) = P^{t_j}(0) + \left. \frac{dP^{t_j}(\tau)}{d\tau} \right|_{\tau=0} \tau + o(\tau).$$

В результате, справедливо неравенство

$$\begin{aligned} & \left\| U(t)P^t(0)x - U(t)P^{t_j} \left(\frac{t-t_j}{\bar{T}} \right) x \right\| \leq \\ & \leq \|U(t)\| \|x\| \left(\left\| \frac{dP^t(\tau)}{dt} \Big|_{t=t_j} \right\| \Delta t + \left\| \frac{P^{t_j}(\tau)}{d\tau} \Big|_{\tau=0} \right\| \Delta t + o(\Delta t) \right), \end{aligned} \quad (4.114)$$

где $\tau = \frac{t-t_j}{\bar{T}}$.

При $\Delta t \rightarrow 0$ правая часть выражения (4.114) стремится к нулю. Следовательно, и левая часть (4.114) также стремится к нулю. Теорема доказана. \square

Теорема 4.2.11 В условиях Теоремы 4.2.7, пусть игра $\Gamma(x, t, t + \bar{T}) \forall t \geq t_0$ имеет единственное равновесие по Нэшу с непрерывным обновлением информации, матрица $R_{ii}^{-1}(t)B'_i(t)$ равномерно ограничена при $t \geq t_0$, решение $Z_i^t(\tau)$ уравнения Риккати (4.96) непрерывно дифференцируемо по t и τ , а производные функции $\frac{dZ_i^t(\tau)}{dt}$ и $\frac{dZ_i^t(\tau)}{d\tau}$ равномерно ограничены при $t \geq t_0$ и $\tau \in [0, 1]$.

Тогда при $x \in X$, где X — ограниченное множество, равновесие по Нэшу в позиционных стратегиях с динамическим обновлением информации $v_{fb,i}^{NE}(t, x)$ равномерно сходится при $\Delta t \rightarrow 0$ к равновесию по Нэшу в позиционных стратегиях с непрерывным обновлением информации $u_{fb,i}^{NE}(t, x)$:

$$v_{fb,i}^{NE}(t, x) \underset{[t_0, +\infty)}{\rightrightarrows} u_{fb,i}^{NE}(t, x), \quad i \in N. \quad (4.115)$$

Доказательство Вводя обозначение $t_j \stackrel{\text{def}}{=} t_0 + j\Delta t$, рассмотрим следующие выражения для $u_{fb,i}^{NE}$ и $v_{fb,i}^{NE}$:

$$\begin{aligned} u_{fb,i}^{NE}(t, x) &= -R_{ii}^{-1}(t)B'_i(t)Z_i^t(0)\bar{T}x, \\ v_{fb,i}^{NE}(t, x) &= -R_{ii}^{-1}(t)B'_i(t)Z_i^{t_j} \left(\frac{t-t_j}{\bar{T}} \right) \bar{T}x, \quad t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t], \end{aligned}$$

где $t \in [t_j, t_{j+1}]$, а $Z_i^t(\tau)$ — решение уравнения (4.96). Положим $U(t) = -\bar{T}R_{ii}^{-1}(t)B'_i(t)$ и $P^t(\tau) = Z_i^t(\tau)$. Непосредственно применяя Лемму 4.2.2, получаем требуемый результат. Теорема доказана. \square

Теорема 4.2.12 В условиях Теоремы 4.2.8, пусть игра $\Gamma(x, t, t + \bar{T}) \forall t \geq t_0$ имеет единственное равновесие по Нэшу с непрерывным обновлением информации, матрица $R_{ii}^{-1}(t)B'_i(t)$ равномерно ограничена при $t \geq t_0$, произведение решения $M_i^t(\tau)$ уравнения Риккати (4.111) на функцию $\Phi^t(\tau)$ непрерывно дифференцируемо по t и τ , а производные функции $\frac{d(M_i^t(\tau)\Phi^t(\tau))}{dt}$ и $\frac{d(M_i^t(\tau)\Phi^t(\tau))}{d\tau}$ равномерно ограничены при $t \geq t_0$ и $\tau \in [0, 1]$, где

$$\frac{d\Phi^t}{d\tau} = \left(A(t + \bar{T}\tau) - \sum_{i \in N} B_i(t + \bar{T}\tau)R_{ii}^{-1}(t + \bar{T}\tau)B'_i(t + \bar{T}\tau) \right) \Phi^t(\tau),$$

$$\Phi^t(0) = E.$$

Тогда при $x \in X$, где X — ограниченное множество, равновесие по Нэшу в программных стратегиях с динамическим обновлением информации $v_{ol,i}^{NE}(t, x)$ равномерно сходится при $\Delta t \rightarrow 0$ к равновесию по Нэшу в программных стратегиях с непрерывным обновлением информации $u_{ol,i}^{NE}(t, x)$:

$$v_{ol,i}^{NE}(t, x) \underset{[t_0, +\infty)}{\rightrightarrows} u_{ol,i}^{NE}(t, x), \quad i \in N. \quad (4.116)$$

Доказательство Вводя обозначение $t_j \stackrel{\text{def}}{=} t_0 + j\Delta t$, рассмотрим следующие выражения для $u_{ol,i}^{NE}$ и $v_{ol,i}^{NE}$:

$$u_{ol,i}^{NE}(t, x) = -R_{ii}^{-1}(t)B'_i(t)M_i^t(0)\bar{T}x,$$

$$\begin{aligned} v_{ol,i}^{NE}(t, x) &= -R_{ii}^{-1}(t)B'_i(t)M_i^{t_j} \left(\frac{t - (t_0 + j\Delta t)}{\bar{T}} \right) \times \\ &\times \Phi^{t_j} \left(\frac{t - (t_0 + j\Delta t)}{\bar{T}} \right) \bar{T}x, \quad t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t], \end{aligned}$$

где $t \in [t_j, t_{j+1}]$, а $M_i^t(\tau)$ — решение (4.111).

Положим $U(t) = -\bar{T}R_{ii}^{-1}(t)B'_i(t)$ и $P^t(\tau) = M_i^t(\tau)\Phi^t(\tau)$. Непосредственно применяя Лемму 4.2.2, получаем требуемый результат. Теорема доказана. \square

Лемма 4.2.3 Пусть некоторые матричные функции $V_1(t)$ и $V_2(t)$ равномерно ограничены при $t \geq t_0$, некоторая параметрическая матричная функция $P^t(\tau)$

непрерывно дифференцируема как по t , так и по τ , а производные функции $\frac{dP^t(\tau)}{t}$ и $\frac{dP^t(\tau)}{\tau}$ равномерно ограничены при $t \geq t_0$ и $\tau \in [0, 1]$.

Рассмотрим дифференциальное уравнение для $y(t)$, $t \in [t_0, +\infty]$, вида

$$\frac{dy(t)}{dt} = (V_1(t) + V_2(t)P^t(0))y, \quad (4.117)$$

где $y(t_0) = y_0$.

Кроме того, рассмотрим дифференциальное уравнение для $z_j(t)$, $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, вида

$$\frac{dz_j(t)}{dt} = \left(V_1(t) + V_2(t)P^{t_0+j\Delta t} \left(\frac{t - (t_0 + j\Delta t)}{\bar{T}} \right) \right) z_j, \quad (4.118)$$

где $z_j(t_0 + j\Delta t) = z_{j-1}(t_0 + j\Delta t)$ при $j \geq 1$, и $z_0(t_0) = y_0$.

Пусть $y^*(t)$ является решением уравнения (4.117) и

$$z^*(t) = \begin{cases} z_0(t), & t \in [t_0, t_0 + \Delta t], \\ \dots \\ z_j(t), & t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t], \\ \dots \end{cases}$$

где $z_j(t)$ удовлетворяет (4.118). Пусть $y^*(t)$ и $z^*(t)$ существуют при $t \geq t_0$. Тогда $z^*(t)$ поточечно сходится к $y^*(t)$ при $\Delta t \rightarrow 0$:

$$z^*(t) \xrightarrow{\Delta t \rightarrow 0} y^*(t).$$

Доказательство Введем обозначение $t_j \stackrel{\text{def}}{=} t_0 + j\Delta t$. Тогда $t \in [t_j, t_{j+1}]$, и необходимо показать, что $\|z^*(t) - y^*(t)\| \rightarrow 0$ при $\Delta t \rightarrow 0$.

Траектории $z^*(t)$ и $y^*(t)$ удовлетворяют дифференциальным уравнениям (4.118), (4.117) при $t \in [t_j, t_{j+1}]$. Заметим, что

$$P^t(0)y - P^{t_j} \left(\frac{t - t_j}{\bar{T}} \right) z = P^t(0)(y - z) + \left(P^t(0) - P^{t_j} \left(\frac{t - t_j}{\bar{T}} \right) \right) z.$$

Положим $w_j(t) = y(t) - z(t)$ при $t \in [t_j, t_{j+1}]$, $V_3(t) = V_1(t) + V_2(t)P^t(0)$ и

$$f_j(t) = V_2(t) \left(P^t(0) - P^{t_j} \left(\frac{t - t_j}{T} \right) \right) z(t).$$

Тогда $w(t)$ удовлетворяет дифференциальному уравнению

$$\frac{dw_j(t)}{dt} = V_3(t)w_j(t) + f_j(t).$$

Рассмотрим функции

$$w(t) = \begin{cases} w_0(t), & t \in [t_0, t_0 + \Delta t], \\ \dots \\ w_j(t), & t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t], \\ \dots \end{cases} \quad (4.119)$$

и

$$f(t) = \begin{cases} f_0(t), & t \in [t_0, t_0 + \Delta t], \\ \dots \\ f_j(t), & t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t]. \\ \dots \end{cases}$$

Тогда (4.119) удовлетворяет дифференциальному уравнению

$$\dot{w}(t) = V_2(t)w(t) + f(t)$$

с начальным условием $w(t_0) = 0$, поскольку $z(t_0) = y(t_0)$.

По формуле Коши при любом $t \geq t_0$ имеем:

$$w(t) = Y(t) \int_{t_0}^t Y^{-1}(\xi) f(\xi) d\xi,$$

где $Y(t)$ — фундаментальная матрица уравнения $\frac{dw}{dt} = V_2(t)w(t)$. С учетом

этого, при фиксированном t получаем:

$$\lim_{\Delta t \rightarrow 0} \|w(t)\| \leq \lim_{\Delta t \rightarrow 0} [\|Y(t)\|\hat{w}(t)\beta\Delta t(t-t_0) + o(\Delta t)] = 0, \quad (4.120)$$

где

$$\begin{aligned} \hat{w}(t) &= \max_{\tau \in [t_0, t]} \|Y^{-1}(\tau)\| \\ \beta &= \sum_{i=1}^N \|V_2(t)\| \left(\left\| \frac{dP^t(\tau)}{dt} \Big|_{t=t_j} \right\| + \left\| \frac{dP^{t_j}(\tau)}{d\tau} \Big|_{\tau=0} \right\| \right) M(t), \\ M(t) &= \max_{\tau \in [t_0, t]} \|z(\tau)\|. \end{aligned}$$

Согласно (4.120), $w(t) \xrightarrow{[t_0, +\infty)} 0$ при $\Delta t \rightarrow 0$. Лемма доказана. \square

Теорема 4.2.13 В условиях Теоремы 4.2.7, пусть игра $\Gamma(x, t, t + \bar{T}) \forall t \geq t_0$ имеет единственное равновесие по Нэшу с непрерывным обновлением информации, матрицы $A(t)$ и $B_i(t)R_{ii}^{-1}(t)B_i'(t)$ равномерно ограничены при $i \in N$ и $t \geq t_0$, решение $Z_i^t(\tau)$ уравнения Риккати (4.96) непрерывно дифференцируемо по t и τ , а производные функции $\frac{dZ_i^t(\tau)}{dt}$ и $\frac{dZ_i^t(\tau)}{d\tau}$ равномерно ограничены при $t \geq t_0$ и $\tau \in [0, 1]$.

Тогда при $x \in X$, где X — ограниченное множество, равновесная по Нэшу траектория в позиционных стратегиях $\hat{x}_{fb}^{NE}(t)$ в игре с динамическим обновлением информации поточечно сходится при $\Delta t \rightarrow 0$ к равновесной по Нэшу траектории в позиционных стратегиях $\tilde{x}_{fb}^{NE}(t)$ в игре с непрерывным обновлением информации:

$$\hat{x}_{fb}^{NE}(t) \xrightarrow{[t_0, +\infty)} \tilde{x}_{fb}^{NE}(t). \quad (4.121)$$

Доказательство Положим

$$\begin{aligned} V_1(t) &= A(t), \\ V_2(t) &= [-\bar{T}B_1(t)R_{11}^{-1}(t)B_1'(t), \dots, -\bar{T}B_N(t)R_{NN}^{-1}(t)B_N'(t)], \\ P^t(\tau) &= \begin{bmatrix} Z_1^t \\ \vdots \\ Z_N^t \end{bmatrix}. \end{aligned}$$

Непосредственно применяя Лемму 4.2.3, получаем требуемый результат. Теорема доказана. \square

Теорема 4.2.14 В условиях Теоремы 4.2.8, пусть игра $\Gamma(x, t, t + \bar{T}) \forall t \geq t_0$ имеет единственное равновесие по Нэшу с непрерывным обновлением информации, матрицы $A(t)$ и $B_i(t)R_{ii}^{-1}(t)B'_i(t)$ равномерно ограничены при $i \in N$ и $t \geq t_0$, произведение $M_i^t(\tau)\Phi^t(\tau)$ непрерывно дифференцируемо по t и τ , где $M_i^t(\tau)$ — решение уравнения Риккати (4.111), а $\Phi^t(\tau)$ — решение дифференциального уравнения

$$\frac{d\Phi^t}{d\tau} = \left(A(t + \bar{T}\tau) - \sum_{i \in N} B_i(t + \bar{T}\tau)R_{ii}^{-1}(t + \bar{T}\tau)B'_i(t + \bar{T}\tau) \right) \Phi^t(\tau),$$

$$\Phi^t(0) = E,$$

а производные функции $\frac{dM_i^t(\tau)\Phi^t(\tau)}{dt}$ и $\frac{dM_i^t(\tau)\Phi^t(\tau)}{d\tau}$ равномерно ограничены при $t \geq t_0$ и $\tau \in [0, 1]$.

Тогда при $x \in X$, где X — ограниченное множество, равновесная по Нэшу траектория в программных стратегиях $\hat{x}_{ol}^{NE}(t)$ в игре с динамическим обновлением информации поточечно сходится при $\Delta t \rightarrow 0$ к равновесной по Нэшу траектории в программных стратегиях $\tilde{x}_{ol}^{NE}(t)$ в игре с непрерывным обновлением информации:

$$\hat{x}_{ol}^{NE}(t) \xrightarrow{\Delta t \rightarrow 0} \tilde{x}_{ol}^{NE}(t). \quad (4.122)$$

Доказательство Положим

$$V_1(t) = A(t),$$

$$V_2(t) = [-\bar{T}B_1(t)R_{11}^{-1}(t)B'_1(t), \dots, -\bar{T}B_N(t)R_{NN}^{-1}(t)B'_N(t)],$$

$$P^t(\tau) = \begin{bmatrix} M_1^t(\tau)\Phi^t(\tau) \\ \vdots \\ M_N^t(\tau)\Phi^t(\tau) \end{bmatrix}.$$

Непосредственно применяя Лемму 4.2.3, получаем требуемый результат. Теорема доказана. \square

4.2.5.5 Дифференциальная игра стабилизации государственного долга с непрерывным обновлением информации

4.2.5.5.1 Общее описание

Рассмотрим следующую дифференциальную игру, основанную на модели стабилизации государственного долга [226]. Динамическое ограничение государственного бюджета показывает взаимосвязь между первичным бюджетным дефицитом, $f(t)$, монетизацией или сеньоражем, $m(t)$, выплатой процентов по государственному долгу, $rd(t)$, и накоплением государственного долга $\dot{d}(t)$:

$$\begin{aligned}\dot{d}(t) &= rd(t) + f(t) - m(t), \\ d(t_0) &= d_0, \quad x^t \in \mathbb{R}^1.\end{aligned}$$

Здесь $d(t)$, $f(t)$ и $m(t)$ выражены в долях ВВП, а $r > 0$ представляет собой процентную ставку по непогашенному государственному долгу за вычетом темпа роста производства. Предположим, что налогово-бюджетная и денежно-кредитная политика контролируется разными государственными органами с разными целями, описанными с помощью функций выигрыша. Целью налогово-бюджетного органа является минимизация суммы временных профилей первичного бюджетного дефицита, роста денежной базы и государственного долга:

$$J_1 = \int_{t_0}^{\infty} e^{-\delta t} (f^2(t) + \eta m^2(t) + \lambda d^2(t)) dt.$$

Параметры η и λ выражают относительный приоритет, придаваемый фискальным органом росту денежной базы и государственному долгу. Предполагается, что денежно-кредитный орган выбирает рост денежной базы таким образом, чтобы сумма временных профилей роста денежной базы и государственного долга была минимальной:

$$J_2 = \int_{t_0}^{\infty} e^{-\delta t} (m^2(t) + \kappa d^2(t)) dt.$$

Здесь $1/\kappa$ можно интерпретировать как меру консерватизма центрального банка в отношении роста денежной массы. Кроме того, все переменные нормированы таким образом, что их целевые значения равны нулю, а все параметры положительны.

4.2.5.5.2 Игра с непрерывным обновлением информации

Теперь рассмотрим неавтономный случай с непрерывным обновлением информации. Предположим, что оба органа в каждый момент времени $t \in [t_0, +\infty)$ используют информацию об уравнениях движения и функциях выигрыша на временном интервале $[t, t + \bar{T}]$. По мере продвижения текущего времени t смещается и временной интервал, на котором доступна информация. Уравнения движения для игры с непрерывным обновлением информации имеют вид

$$\dot{d}^t(s) = rd^t(s) + f^t(s, d) - m^t(s, d), \quad d^t(t) = d. \quad (4.123)$$

Функции выигрыша игроков в игре с непрерывным обновлением информации имеют вид

$$J_1^t = \int_t^{t+\bar{T}} e^{-\delta s} \left((f^t(s))^2 + \eta (m^t(s, d))^2 + \lambda (d^t(s, d))^2 \right) ds, \quad (4.124)$$

$$J_2^t = \int_t^{t+\bar{T}} e^{-\delta s} \left((m^t(s, d))^2 + \kappa (d^t(s, d))^2 \right) ds.$$

Важно отметить, что неавтономный характер этой игры выражается в дополнительном члене $e^{-\delta s}$, представляющем зависимость функции выигрыша от текущего времени s .

Для аналитической разрешимости предположим, что $\eta = 0$, $\lambda = \kappa = q$ и $\delta = 2r$. Тогда игроки симметричны. Согласно Теореме 4.2.7, для нахождения равновесия по Нэшу в позиционных стратегиях с непрерывным обновлением информации на первом шаге, нам необходимо решить следующее дифференци-

альное уравнение:

$$\begin{cases} \dot{k}^t(\tau) = 3\bar{T}^2 e^{2r(t+\bar{T}\tau)} (k^t(\tau))^2 - 2\bar{T}rk(\tau) - qe^{-2r(t+\bar{T}\tau)}, \\ k^t(1) = 0. \end{cases} \quad (4.125)$$

Как показано в [227], решением уравнения (4.125) является функция

$$k^t(\tau) = \gamma e^{-2r(t+\bar{T}\tau)} \tanh\left(3\bar{T}^2\gamma(1-\tau)\right), \quad (4.126)$$

где $\gamma = \sqrt{\frac{q}{3\bar{T}^2}}$. Согласно (4.97), равновесие по Нэшу в позиционных стратегиях с непрерывным обновлением информации имеет вид

$$\begin{aligned} \tilde{f}^{NE}(t, d) &= -\bar{T}k^t(0)d, \\ \tilde{m}^{NE}(t, d) &= \bar{T}k^t(0)d. \end{aligned} \quad (4.127)$$

При подстановке выражения (4.126) в формулу (4.127), получаем

$$\begin{aligned} \tilde{f}^{NE}(t, d) &= -\gamma\bar{T}de^{-2rt} \tanh\left(3\bar{T}^2\gamma\right), \\ \tilde{m}^{NE}(t, d) &= \gamma\bar{T}de^{-2rt} \tanh\left(3\bar{T}^2\gamma\right). \end{aligned} \quad (4.128)$$

Наконец, подставляя выражение (4.128) в уравнение (4.123), устанавливаем, что функция $\tilde{d}^{NE}(t)$ — решение дифференциального уравнения

$$\begin{aligned} \dot{d}(t) &= \left(r - 2\gamma\bar{T}e^{-2rt} \tanh\left(3\bar{T}^2\gamma\right)\right) d(t), \\ d(t_0) &= d_0, \end{aligned} \quad (4.129)$$

где $\gamma = \sqrt{\frac{q}{3\bar{T}^2}}$, $t \in [t_0, +\infty)$.

4.2.5.5.3 Игра с динамическим обновлением информации

На основе подхода, описанного выше в Разделе 4.2.5.4.1, и вычислений для исходной игры, выполним аналогичные вычисления для равновесия по Нэшу в

игре с динамическим обновлением информации. В результате получаем:

$$\begin{aligned}\hat{f}^{NE}(t, d) &= -\bar{T}k^{t_i} \left(\frac{t - t_i}{\bar{T}} \right) d, \\ \hat{m}^{NE}(t, d) &= \bar{T}k^{t_i} \left(\frac{t - t_i}{\bar{T}} \right) d,\end{aligned}\tag{4.130}$$

где $t \in [t_i, t_{i+1}]$. При подстановке выражения (4.126) в формулу (4.130) находим:

$$\begin{aligned}\hat{f}^{NE}(t, d) &= -\gamma\bar{T}de^{-2rt} \tanh \left(3\bar{T}^2 \gamma \left(1 - \frac{t - t_i}{\bar{T}} \right) \right), \\ \hat{m}^{NE}(t, d) &= \gamma\bar{T}de^{-2rt} \tanh \left(3\bar{T}^2 \gamma \left(1 - \frac{t - t_i}{\bar{T}} \right) \right).\end{aligned}\tag{4.131}$$

Наконец, подставляя (4.131) в (4.123), устанавливаем, что функция $\hat{d}^{NE}(t)$ — решение дифференциального уравнения

$$\begin{aligned}\dot{d}(t) &= \left(r - 2\gamma\bar{T}e^{-2rt} \tanh \left[3\bar{T}^2 \gamma \left(1 - \frac{t - t_i}{\bar{T}} \right) \right] \right) d(t), \\ d(t_0) &= d_0,\end{aligned}\tag{4.132}$$

где $\gamma = \sqrt{\frac{q}{3\bar{T}^2}}$.

4.2.5.5.4 Автономный и неавтономный случаи

Предположим, что информация не меняется с течением времени. Другими словами, рассмотрим автономный случай с непрерывным обновлением. Для этого зафиксируем $t = t_0$ в (4.123) и (4.124). В результате получаем следующие стратегии:

$$\begin{aligned}\tilde{f}_{t_0}^{NE}(t, d) &= -\gamma\bar{T}de^{-2rt_0} \tanh \left(3\bar{T}^2 \gamma \right), \\ \tilde{m}_{t_0}^{NE}(t, d) &= \gamma\bar{T}de^{-2rt_0} \tanh \left(3\bar{T}^2 \gamma \right).\end{aligned}\tag{4.133}$$

Подставляя выражение (4.133) в уравнения (4.123), устанавливаем, что $\tilde{d}_{t_0}^{NE}(t)$ — решение дифференциального уравнения

$$\begin{aligned} \dot{d}(t) &= \left(r - 2\gamma\bar{T}e^{-2rt_0} \tanh\left(3\bar{T}^2\gamma\right) \right) d(t), \\ d(t_0) &= d_0, \end{aligned} \quad (4.134)$$

где $\gamma = \sqrt{\frac{q}{3\bar{T}^2}}$, $t \in [t_0, +\infty)$.

4.2.5.5.5 Решение без обновления информации

Рассмотрим следующую замену переменных, превращающую данную модель в классическую игру из [224]:

$$\bar{d}(t) = e^{-\frac{1}{2}\delta t} d(t), \quad \bar{m}(t) = e^{-\frac{1}{2}\delta t} m(t), \quad \bar{f}(t) = e^{-\frac{1}{2}\delta t} f(t).$$

Тогда динамика системы описывается дифференциальными уравнениями

$$\begin{aligned} \dot{\bar{d}}(t) &= \left(r - \frac{1}{2}\delta \right) \bar{d} + \bar{f} - \bar{m}, \\ \bar{d}(t_0) &= e^{-\frac{1}{2}\delta t_0} d_0, \end{aligned}$$

и функции затрат игроков имеют вид

$$\begin{aligned} \bar{J}_1 &= \int_{t_0}^{\infty} (\bar{f}^2(t) + \eta\bar{m}^2(t) + \lambda\bar{d}^2(t)) dt, \\ \bar{J}_2 &= \int_{t_0}^{\infty} (\bar{m}^2(t) + \kappa\bar{d}^2(t)) dt. \end{aligned}$$

Для корректного сравнения со случаем непрерывного обновления информации, положим $\eta = 0$, $\lambda = \kappa = q$ и $\delta = 2r$. Согласно [224], равновесием по Нэшу в позиционных стратегиях являются

$$\bar{f}^{NE}(t, \bar{d}) = -\sqrt{\frac{q}{2}}\bar{d}, \quad \bar{m}^{NE}(t, \bar{d}) = \sqrt{\frac{q}{2}}\bar{d}.$$

Получаем следующее равновесие по Нэшу в позиционных стратегиях в исходных переменных:

$$\begin{aligned} f^{NE}(t, d) &= -e^{-2rt} \sqrt{\frac{q}{2}} d, \\ m^{NE}(t, d) &= e^{-2rt} \sqrt{\frac{q}{2}} d. \end{aligned} \quad (4.135)$$

Подставляя выражение (4.135) в уравнения (4.123), устанавливаем, что $d^{NE}(t)$ — решение дифференциального уравнения

$$\begin{aligned} \dot{d}(t) &= \left(r - e^{-2rt} \sqrt{2q} \right) d(t), \\ d(t_0) &= d_0. \end{aligned} \quad (4.136)$$

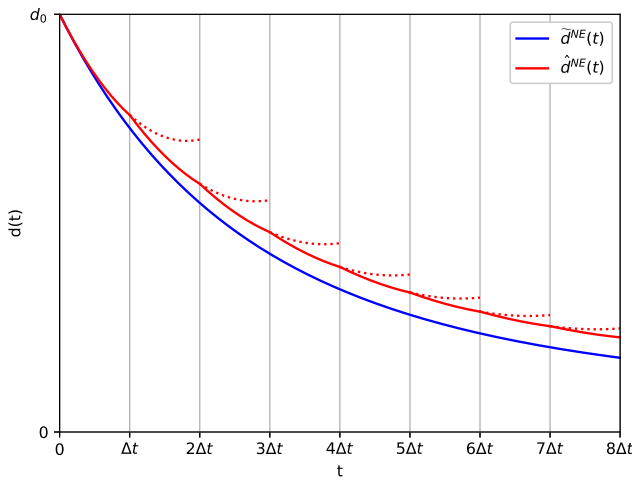


Рис. 4.11: $\tilde{d}^{NE}(t)$ (4.129) — синяя линия, $\hat{d}^{NE}(t)$ (4.132) — красная пунктирная линия.

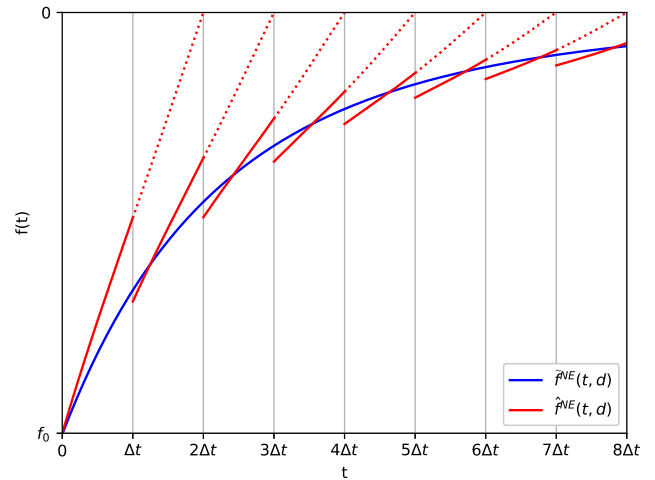


Рис. 4.12: $\tilde{f}^{NE}(t)$ (4.128) — синяя линия, $\hat{f}^{NE}(t)$ (4.131) — красная пунктирная линия.

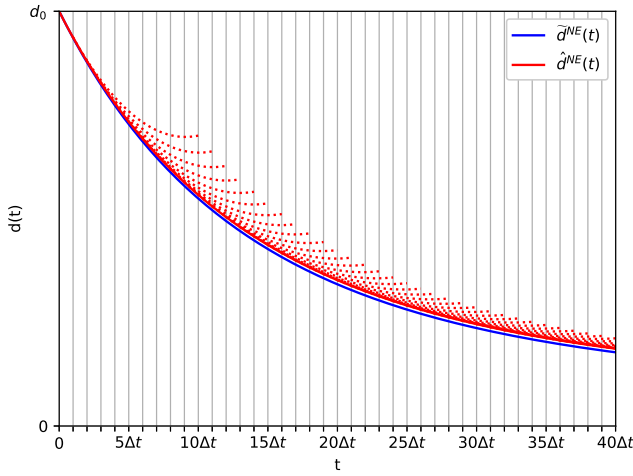


Рис. 4.13: $\tilde{d}^{NE}(t)$ (4.129) — синяя линия, $\hat{d}^{NE}(t)$ (4.132) — красная пунктирная линия.

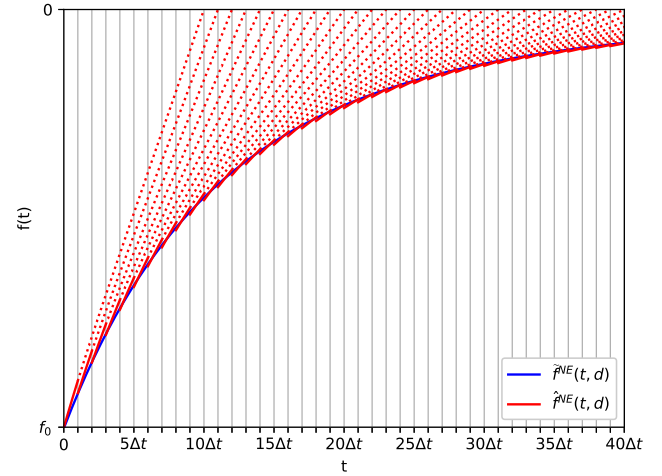


Рис. 4.14: $\tilde{f}^{NE}(t)$ (4.128) — синяя линия, $\hat{f}^{NE}(t)$ (4.131) — красная пунктирная линия.

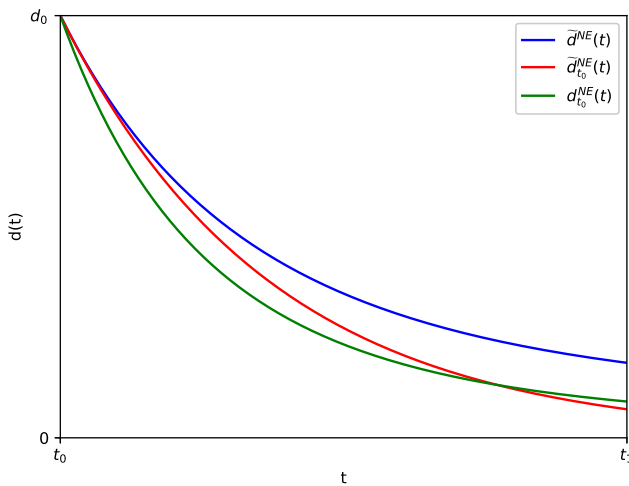


Рис. 4.15: $\tilde{d}^{NE}(t)$ (4.129) — синяя линия, $\tilde{d}_{t_0}^{NE}(t)$ (4.134) — красная линия, $d^{NE}(t)$ (4.136) — зеленая линия.

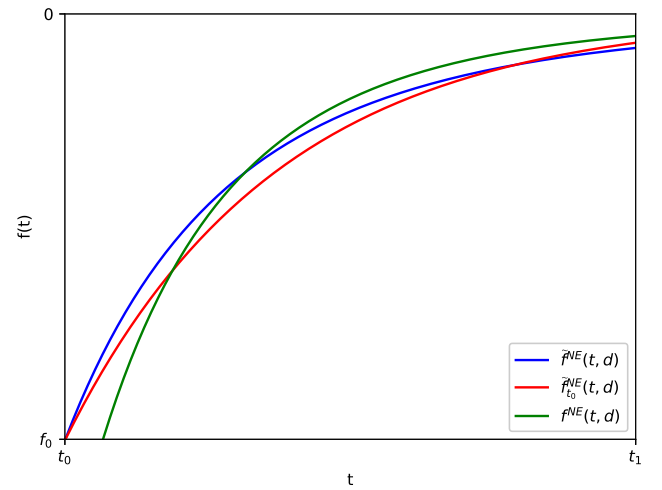


Рис. 4.16: $\tilde{f}^{NE}(t, x)$ (4.128) — синяя линия, $\tilde{f}_{t_0}^{NE}(t, x)$ (4.133) — красная линия, $f^{NE}(t, x)$ (4.135) — зеленая линия.

4.2.5.5.6 Численное моделирование

Рассмотрим результаты численного моделирования для представленной выше игровой модели на временном интервале $[1, 5]$, т.е. $t_0 = 1$ и $T = 5$. В начальный момент времени $t_0 = 1$ запас знаний равен 1, т.е. $x_0 = 1$. Остальные параметры

модели таковы: $r = 0.1$, $q = 0.8$, и $\bar{T} = 1$. Предположим, что при динамическом обновлении информации (красные сплошные и пунктирные линии, см. рис. 4.25–4.26) интервалы между моментами обновления информации равны $\Delta t = 2$. Следовательно, число интервалов в игре с динамическим обновлением информации равно $l = \frac{T-t_0}{\Delta t} = 8$. На рисунке 4.25 представлено сравнение равновесия по Нэшу с динамическим обновлением информации (красная линия) и равновесия по Нэшу с непрерывным обновлением информации (синяя линия). На рисунке 4.26 представлены аналогичные результаты для соответствующих стратегий.

Для иллюстрации Теорем 4.2.11 и 4.2.13 о сходимости равновесных стратегий с динамическим обновлением информации и соответствующей траектории к равновесным стратегиям и траектории с непрерывным обновлением информации, рассмотрим результаты моделирования для случая частого обновления информации, а именно $l = \frac{T-t_0}{\Delta t} = 40$. На рисунках 4.27–4.28 показаны те же решения, что и на рисунках 4.25–4.26, но для случая $\Delta t = 0,1$. Таким образом, сходимость стратегий и траектории подтверждается численными экспериментами, представленными ниже.

Для сравнения автономного и неавтономного случаев с непрерывным обновлением информации, рассмотрим результаты моделирования при $t \in [1, 5]$, с параметрами $r = 0.1$, $q = 0.8$, $\bar{T} = 1$, и $x_0 = 1$. Стратегии $\tilde{f}^{NE}(t, x)$ и $\tilde{f}_{t_0}^{NE}(t, x)$, полученные соответственно в неавтономном и автономном случаях, представлены на рисунке 4.16. Дополнительные траектории $\tilde{d}^{NE}(t)$ и $\tilde{d}_{t_0}^{NE}(t)$ представлены на рисунке 4.15. На этих рисунках видны различия между стратегиями с непрерывным обновлением информации и соответствующими стратегиями в исходной модели дифференциальной игры без обновления информации $f^{NE}(t)$, а также между соответствующими траекториями $d^{NE}(t)$.

4.3 Модель кооперативной дифференциальной игры с непрерывным обновлением информации

В модели кооперативной дифференциальной игры с трансферабельной полезностью также существуют две проблемы: как можно определить кооперативное

поведение, используя кооперативные стратегии, и как можно определить кооперативные решения? Следует обратиться к обеим проблемам, используя подход с непрерывным обновлением информации.

4.3.1 Кооперативная игра с непрерывным обновлением информации

4.3.1.1 Кооперативные стратегии в игре с непрерывным обновлением информации

При непрерывно обновляемой информации важно моделировать поведение игроков. В этом разделе рассмотрим случай кооперативных стратегий. Для класса дифференциальных игр с непрерывным обновлением информации, придадим ему следующий вид:

- для любого фиксированного $t \in [t_0, +\infty)$, $u^*(t, x) = (u_1^*(t, x), \dots, u_n^*(t, x))$ совпадает с кооперативными стратегиями в игре (4.5), (4.6), определенной на временном интервале $[t, t + \bar{T}]$ в момент времени t .

Однако прямое применение классических подходов для определения кооперативных стратегий невозможно. Рассмотрим два временных интервала, $[t, t + \bar{T}]$ и $[t + \epsilon, t + \bar{T} + \epsilon]$, где $\epsilon \ll \bar{T}$. Тогда согласно постановке задачи:

- $u^*(t, x)$ в момент времени t совпадает с кооперативными стратегиями в игре, определенной на временном интервале $[t, t + \bar{T}]$,
- $u^*(t + \epsilon, x)$ в момент времени $t + \epsilon$ должна совпадать с кооперативными стратегиями в игре, определенной на временном интервале $[t + \epsilon, t + \epsilon + \bar{T}]$.

Для построения таких стратегий рассмотрим концепцию обобщенного равновесия по Нэшу в *позиционных* (и *программных*) стратегиях:

$$\tilde{u}_{fb}^*(t, x; s, x_t) = (\tilde{u}_{fb,1}^*(t, x; s, x_t), \dots, \tilde{u}_{fb,n}^*(t, x; s, x_t)), \quad t \in [t_0, T], \quad s \in [t, t + \bar{T}],$$

$$(\tilde{u}_{ol}^*(t, x; s) = (\tilde{u}_{ol,1}^*(t, x; s), \dots, \tilde{u}_{ol,n}^*(t, x; s)), \quad t \in [t_0, T], \quad s \in [t, t + \bar{T}]),$$

которое будет использовано в дальнейшем для построения стратегий $u_{fb}^*(t, x)$ ($u_{ol}^*(t, x)$). Для каждого фиксированного t стратегии $\tilde{u}_{fb}^*(t, x; s, x_t) = (\tilde{u}_{fb,1}^*(t, x; s, x_t), \dots, \tilde{u}_{fb,n}^*(t, x; s, x_t))$ ($\tilde{u}_{ol}^*(t, x; s) = (\tilde{u}_{ol,1}^*(t, x; s), \dots, \tilde{u}_{ol,n}^*(t, x; s))$) определяются следующим образом:

$$\tilde{u}_{fb}^*(t, x; s, x_t) = (\tilde{u}_{fb,1}^*(t, x; s, x_t), \dots, \tilde{u}_{fb,n}^*(t, x; s, x_t)) = \arg \max_{\tilde{u}_{fb,1}, \dots, \tilde{u}_{fb,n}} \sum_{i=1}^n K_i^t(x, t; \tilde{u}(t)), \quad (4.137)$$

$$\left(\tilde{u}_{ol}^*(t, x; s) = (\tilde{u}_{ol,1}^*(t, x; s), \dots, \tilde{u}_{ol,n}^*(t, x; s)) = \arg \max_{\tilde{u}_{ol,1}, \dots, \tilde{u}_{ol,n}} \sum_{i=1}^n K_i^t(x, t; \tilde{u}(t)) \right), \quad (4.138)$$

где для фиксированного текущего момент времени t задача оптимизации (4.137) (задача оптимизации (4.138)) построена на основе информации, доступной игрокам в момент времени t , т.е. на временном интервале $[t, t + \bar{T}]$. Предположим, что максимум в выражении (4.137) (в выражении (4.138)) достигается на множестве допустимых стратегий $\tilde{u}_{fb}^*(t, x; s, x_t)$ ($\tilde{u}_{ol}^*(t, x; s)$) с фиксированным t .

Определение 4.3.1 Стратегии из ситуации $\tilde{u}_{fb}^*(t, x; s, x_t) = (\tilde{u}_{fb,1}^*(t, x; s, x_t), \dots, \tilde{u}_{fb,n}^*(t, x; s, x_t))$ ($\tilde{u}_{ol}^*(t, x; s) = (\tilde{u}_{ol,1}^*(t, x; s), \dots, \tilde{u}_{ol,n}^*(t, x; s))$) называются обобщенными позиционными (программными) кооперативными стратегиями в игре с непрерывным обновлением информации, если для любого фиксированного $t \in [t_0, +\infty)$ стратегии $\tilde{u}_{fb}^*(t, x; s, x_t)$ ($\tilde{u}_{ol}^*(t, x; s)$) являются позиционными (программными) кооперативными стратегиями в игре $\Gamma(x, t, t + \bar{T})$.

Важно отметить, что обобщенные позиционные (программные) кооперативные стратегии $\tilde{u}_{fb}^*(t, x; s, x_t)$ ($\tilde{u}_{ol}^*(t, x; s)$) при фиксированном t являются функцией s и x , где определены на временном интервале $[t, t + \bar{T}]$. Используя обобщенные кооперативные стратегии, можно определить кооперативные стратегии, которые будут реализованы в игре с непрерывным обновлением информации.

Определение 4.3.2 Стратегии из ситуации $u_{fb}^*(t, x)$ ($u_{ol}^*(t, x)$) называются позиционными (программными) кооперативными стратегиями с непрерывным обновлением информации, если

$$u^*(t, x) = \tilde{u}^*(t, s, x)|_{s=t} = (\tilde{u}_1^*(t, s, x)|_{s=t}, \dots, \tilde{u}_n^*(t, s, x)|_{s=t}), \quad (4.139)$$

$$t \in [t_0, +\infty),$$

где $\tilde{u}_{fb}^*(t, x; s, x_t)$ ($\tilde{u}_{ol}^*(t, x; s)$) являются обобщенными позиционными (программными) кооперативными стратегиями в смысле Определения 4.3.1.

В отличие от обобщенных кооперативных стратегий, $u_{fb}^*(t, x)$ ($u_{ol}^*(t, x)$) не содержит кооперативных стратегий ни для каких $s \in [t, t + \bar{T}]$. Ситуация $u_{fb}^*(t, x)$ ($u_{ol}^*(t, x)$) учитывает только стратегии игроков, действующих в соответствии с процедурой непрерывного обновления информации с $s = t$ (см. описание в Разделе 4.2). Стратегии из ситуации $u_{fb}^*(t, x)$ ($u_{ol}^*(t, x)$) будут использоваться в качестве кооперативных в игровой модели с непрерывным обновлением информации.

4.3.1.2 Кооперативная траектория в игре с непрерывным обновлением информации

Траектория $x^*(t)$, соответствующая кооперативным стратегиям с непрерывным обновлением информации $u^*(t, x)$, может быть получена из системы

$$\begin{aligned} \dot{x}(t) &= f(t, x, u^*), \\ x(t_0) &= x_0, \\ x &\in \mathbb{R}^l. \end{aligned} \tag{4.140}$$

4.3.1.3 Кооперативное решение с непрерывным обновлением информации

Для распределения кооперативного выигрыша между игроками, необходимо определить, как протекает игра в целом, как игроки прогнозируют свое поведение в каждый текущий момент времени $t \in [t_0, +\infty]$ на будущем временном интервале $[t, t + \bar{T}]$, задавая:

- 1 траекторию игры (Раздел 4.3.1.3.1),
- 2 характеристическую функцию вдоль прогнозируемой траектории (Раздел 4.3.1.3.2),

3 дележ и кооперативное решение вдоль прогнозируемой траектории (Раздел 4.3.1.3.3),

и, наконец, как строится кооперативное решение с непрерывным обновлением информации (Раздел 4.3.1.3.4), а также выяснить его свойства (Раздел 4.3.1.4).

4.3.1.3.1 Кооперативные траектории в игре на временном интервале $[t, t + \bar{T}]$

Для каждой модели дифференциальной игры $\Gamma(x, s, t + \bar{T})$, определенной на временном интервале $[t, t + \bar{T}]$, определим кооперативную траекторию, которая должна быть реализована в игре в момент времени t в позиции x на основе информации, доступной в момент времени t . Пусть $\tilde{x}_t^*(s)$ — соответствующая кооперативная траектория, являющаяся решением следующей системы:

$$\begin{aligned} \dot{x}(s) &= f(s, x, \tilde{u}^*(t, s, x)), \\ x(t) &= x \quad (x(t_0) = x_0), \\ x &\in \mathbb{R}^l \end{aligned} \tag{4.141}$$

где функции $\tilde{u}^*(t, s, x)$ взяты при фиксированном текущем времени t .

4.3.1.3.2 Характеристическая функция на временном интервале $[t, t + \bar{T}]$

Поскольку игра непрерывно развивается во времени, процедура вычисления характеристической функции также должна быть определена нестандартным образом. Но на данном этапе сосредоточимся на рассмотрении решения в текущий момент времени t , т.е. решения для игры, определенной на временном интервале $[t, t + \bar{T}]$. Понятие характеристической функции $\tilde{V}_t(S; x, s, t + \bar{T})$, $S \subseteq N$, определено для каждой подыгры $\Gamma(x, t, t + \bar{T})$ при фиксированном текущем моменте времени $t \in [t_0, +\infty)$ и моменте времени $s \in [t, t + \bar{T}]$.

Предположим, что $\tilde{V}_t(S; x, s, t + \bar{T})$, $S \subseteq N$, построено, как в работе [228]:

$$\tilde{V}_t(S; x, s, t + \bar{T}) = \begin{cases} \max_{u^t} \sum_{i \in N} K_i^t(x, s, t + \bar{T}; u^t), & S = N, \\ \max_{u_i^t, i \in S} \sum_{i \in S} K_i^t(x, s, t + \bar{T}; u_S^t, u_{N \setminus S}^{t, NE}), & S \subset N, \\ 0, & S = \emptyset, \end{cases}$$

где $\max_{u_i^t, i \in S} \sum_{i \in S} K_i^t(x, s, t + \bar{T}; u_S^t, u_{N \setminus S}^{t, NE})$ — максимальный кооперативный выигрыш коалиции S при условии, что игроки коалиции $N \setminus S$ используют стратегии $u_{N \setminus S}^{t, NE} = \{u_j^{t, NE}\}_{j \in N \setminus S}$ из фиксированного равновесия по Нэшу с непрерывным обновлением информации $u^{t, NE} = (u_1^{t, NE}, \dots, u_n^{t, NE})$ в некооперативной игре $\Gamma(x, t, t + \bar{T})$.

4.3.1.3.3 Кооперативное решение на временном интервале $[t, t + \bar{T}]$

Чтобы построить результирующее кооперативное решение игры с непрерывным обновлением информации, рассмотрим развитие игры во времени. В этом разделе построим кооперативное решение для любого фиксированного текущего момента времени $t \in [t_0, +\infty)$ вдоль локальной кооперативной траектории $\tilde{x}_t^*(s)$, $t \in [t, t + \bar{T}]$. Траектория $\tilde{x}_t^*(s)$ является кооперативной в игре (4.5), (4.6), определенной на временном интервале $[t, t + \bar{T}]$.

Для каждой кооперативной подыгры $\Gamma(x, t, t + \bar{T})$ дележ $\tilde{\xi}_t(\tilde{x}_t^*(s), s, t + \bar{T})$ определяется как произвольный вектор, удовлетворяющий условиям

$$\begin{aligned} \tilde{\xi}_i^t(\tilde{x}_t^*(s), s, t + \bar{T}) &\geq \tilde{V}_t(\{i\}; \tilde{x}_t^*(s), s, t + \bar{T}), \quad i \in N, \\ \sum_{i \in N} \tilde{\xi}_i^t(\tilde{x}_t^*(s), s, t + \bar{T}) &= \tilde{V}_t(N; \tilde{x}_t^*(s), s, t + \bar{T}). \end{aligned} \quad (4.142)$$

Обозначим через $\tilde{L}_t(\tilde{x}_t^*(s), s, t + \bar{T})$ множество всех возможных дележей для каждой игры, определенной на временном интервале $[t, t + \bar{T}]$. Пусть для каждой игры, определенной на временном интервале $[t, t + \bar{T}]$, непустое решение задается следующим образом:

$$\tilde{M}_t(\tilde{x}_t^*(s), s, t + \bar{T}) \subset \tilde{L}_t(\tilde{x}_t^*(s), s, t + \bar{T}). \quad (4.143)$$

Здесь $\widetilde{M}_t(\tilde{x}_t^*(s), s, t + \bar{T})$ может быть С-ядром, решением Неймана–Моргенштерна, N-ядром, или вектором Шепли.

4.3.1.3.4 Построение кооперативного решения в игре с непрерывным обновлением информации

Чтобы построить кооперативное решение в игре с непрерывным обновлением информации, будем использовать комбинацию дележей и соответствующих ПРД для каждого текущего момента времени $t \in [t_0, T]$.

Пусть в каждой игре, определенной на временном интервале $[t, t + \bar{T}]$, выбрано решение $\widetilde{M}_t(\tilde{x}_t^*(s), s, t + \bar{T}) \neq \emptyset$ вдоль кооперативной траектории $\tilde{x}_t^*(s)$, $s \in [t, t + \bar{T}]$. Кроме того, пусть в начальной позиции $\tilde{x}_t^*(s)$ каждой игры $\Gamma(x, t, t + \bar{T})$ выбраны дележ

$$\tilde{\xi}_t(\tilde{x}_t^*(s), s, t + \bar{T}) \in \widetilde{M}_t(\tilde{x}_t^*(s), s, t + \bar{T}), \quad s \in [t, t + \bar{T}]$$

и соответствующие ПРД

$$\tilde{\beta}_t(s, \tilde{x}_t^*) = [\tilde{\beta}_1^t(s, \tilde{x}_t^*), \dots, \tilde{\beta}_n^t(s, \tilde{x}_t^*)], \quad s \in (t, t + \bar{T}),$$

гарантирующие динамическую устойчивость этого дележа (см. Определение 1.1.0.3):

$$\tilde{\xi}_t(x_t^*(s), s, t + \bar{T}) = \int_s^{t+\bar{T}} \tilde{\beta}_t(\tau, \tilde{x}_t^*) d\tau, \quad s \in (t, t + \bar{T}). \quad (4.144)$$

Отметим, что условие (4.144) эквивалентно условию, описанному в (1.13):

$$\tilde{\beta}_t(s, \tilde{x}_t^*) = -\frac{d}{ds} \tilde{\xi}_t(x_t^*(s), s, t + \bar{T}), \quad s \in (t, t + \bar{T}). \quad (4.145)$$

Новая концепция решения для игры $\Gamma(x_0, t_0, T)$ с непрерывным обновлением информации состоит из комбинации решений $\widetilde{M}_t(\tilde{x}_t^*(t), t, t + \bar{T})$ (соответствующих ПРД) для игр $\Gamma(x, t, t + \bar{T})$, $t \in [t_0, T]$.

Определение 4.3.3 Для каждого множества $\tilde{\xi}_t(\tilde{x}_t^*(s), s, t + \bar{T}) \in \widetilde{M}_t(\tilde{x}_t^*(s), s, t + \bar{T})$

\bar{T}), $t \in [t_0, T]$, ПРД с непрерывным обновлением информации $\beta(s, x^*)$ определяется с помощью соответствующей ПРД $\tilde{\beta}_t(s, \tilde{x}_t^*)$ следующим образом:

$$\beta(t, x^*) = \tilde{\beta}_t(t, \tilde{x}_t^*(t)), \quad t \in [t_0, T]. \quad (4.146)$$

Важно отметить, что траектория с непрерывным обновлением информации $x^*(t)$, определенная в (4.140) в момент времени t , совпадает с $\tilde{x}_t^*(t)$. Это следует из определения траектории $\tilde{x}_t^*(s)$ (см. Раздел 4.3.1.2). Используя ПРД с непрерывным обновлением информации $\beta(t, x^*)$, построим следующий вектор.

Определение 4.3.4 *Дележом с непрерывным обновлением информации $\xi(x^*(t), t, T)$ называется вектор, заданный с помощью ПРД $\beta(t, x^*)$ следующим образом:*

$$\xi(x^*(t), t, T) = \int_t^T \beta(\tau, x^*(\tau)) d\tau, \quad t \in [t_0, T]. \quad (4.147)$$

Используя дележ с непрерывным обновлением информации, введем кооперативное решение с непрерывным обновлением информации.

Определение 4.3.5 *Кооперативным решением с непрерывным обновлением информации $W(x^*(t), t, T)$ называется множество дележей $\xi(x^*(t), t, T)$, (4.147) для всех возможных ПРД $\beta(t, x^*)$ (4.146).*

Представленный в этом разделе подход является естественным способом моделирования кооперативного решения в модели игры с непрерывным обновлением информации. Здесь важно, чтобы игроки в каждый текущий момент времени t располагали только информацией об игре, определенной на интервале $[t, t + \bar{T}]$, и на основе этой информации они могли построить кооперативное решение в каждый момент t (4.144), а позднее — составить результирующий дележ (4.147), используя ПРД с непрерывным обновлением информации (4.146) в качестве мгновенного дележа в каждый момент t . Некоторые свойства построенного решения обсуждаются в разделах 4.3.1.4, 4.3.1.5.

4.3.1.4 Динамическая устойчивость кооперативного решения в игре с непрерывным обновлением информации

Кооперативное решение с непрерывным обновлением информации $W(x_0, t_0, T)$ является динамически устойчивым по построению. Но можно доказать, что это решение также обладает свойством сильной динамической устойчивости [159].

Теорема 4.3.1 *Произвольное кооперативное решение с непрерывным обновлением информации $W(x_0, t_0, T)$ является сильным динамически устойчивым в игре $\Gamma(x_0, t_0, T)$ с непрерывным обновлением информации.*

Произвольное кооперативное решение — любое кооперативное решение, которое может быть получено при выборе различных $\tilde{\xi}_t(\tilde{x}_t^*(s), s, t + \bar{T}) \in \tilde{M}_t(\tilde{x}_t^*(s), s, t + \bar{T})$ в каждой игре, определенной на временном интервале $[t, t + \bar{T}]$, т.е. в каждый момент времени $t \in [t_0, T]$.

Доказательство Согласно определению сильной динамической устойчивости, необходимо показать, что для каждого $\xi(x_0, t_0, T) \in W(x_0, t_0, T)$ существует ПРД $\beta(t, x^*)$, удовлетворяющая при любом $\xi'(x^*(t), t, T) \in W(x^*(t), t, T)$ условию

$$\int_{t_0}^t \beta(\tau, x^*(\tau)) d\tau + \xi'(x^*(t), t, T) \in W(x_0, t_0, T). \quad (4.148)$$

Зафиксируем дележ $\bar{\xi}(x_0, t_0, T) \in W(x_0, t_0, T)$. Тогда необходимо показать, что для этого дележа существует ПРД $\bar{\beta}(t, x^*)$ такая, что

$$\int_{t_0}^t \bar{\beta}(\tau, x^*(\tau)) d\tau + \xi'(x^*(t), t, T) \in W(x_0, t_0, T) \quad (4.149)$$

при любом $\xi'(x^*(t), t, T) \in W(x^*(t), t, T)$. Выберем $\bar{\beta}$ в соответствие со следующим правилом (процедурой непрерывного обновления):

$$\bar{\beta}(\tau, x^*(\tau)) = \tilde{\beta}_\tau(\tau, \tilde{x}_\tau^*(\tau)), \quad (4.150)$$

где $\tilde{\beta}_\tau(\tau, \tilde{x}_\tau^*(\tau))$ выбрано исходя из динамической устойчивости дележа $\tilde{\xi}_\tau(x_\tau^*(s), s, \tau + \bar{T})$ с помощью (4.145); $\xi'(x^*(t), t, T)$ — дележ с непрерывным обновлением

информации в подыгре, начинающейся в момент времени t из позиции $x^*(t)$. Согласно Определению 4.3.4, $\xi'(x^*(t), t, T)$ — вектор, заданный с помощью ПРД:

$$\xi'(x^*(t), t, T) = \int_t^T \beta'(\tau, x^*(\tau)) d\tau, \quad (4.151)$$

где $\beta'(\tau, x^*(\tau))$ построена по процедуре (4.146). Это означает, что $\beta'(\tau, x^*(\tau)) = \tilde{\beta}'_\tau(\tau, \tilde{x}_\tau^*(\tau))$, где $\tilde{\beta}'_\tau(s, \tilde{x}_\tau^*(s))$ — ПРД для некоторого $\tilde{\xi}'_t(\tilde{x}_t^*(s), s, t + \bar{T}) \in \tilde{M}_t(\tilde{x}_t^*(s), s, t + \bar{T})$. С учетом (4.151), формулу (4.149) можно записать в виде

$$\int_{t_0}^t \bar{\beta}(\tau, x^*(\tau)) d\tau + \xi'(x^*(t), t, T) = \int_{t_0}^t \bar{\beta}(\tau, x^*(\tau)) d\tau + \int_t^T \beta'(\tau, x^*(\tau)) d\tau. \quad (4.152)$$

Далее, рассмотрим ПРД

$$\hat{\beta}(\tau, x^*(\tau)) = \begin{cases} \bar{\beta}(\tau, x^*(\tau)), & \tau \in [t_0, t), \\ \beta'(\tau, x^*(\tau)), & \tau \in [t, T). \end{cases} \quad (4.153)$$

В результате получаем:

$$\int_{t_0}^t \bar{\beta}(\tau, x^*(\tau)) d\tau + \xi'(x^*(t), t, T) = \int_{t_0}^T \hat{\beta}(\tau, x^*(\tau)) d\tau. \quad (4.154)$$

Однако, $\hat{\beta}(\tau, x^*(\tau))$ построена согласно формуле (4.146) подхода с непрерывным обновлением информации. Следовательно, $\int_{t_0}^t \bar{\beta}(\tau, x^*(\tau)) d\tau + \xi'(x^*(t), t, T)$ является дележом с непрерывным обновлением информации, и $\int_{t_0}^t \bar{\beta}(\tau, x^*(\tau)) d\tau + \xi'(x^*(t), t, T)$ принадлежит $W(x_0, t_0, T)$. \square

4.3.1.5 Вид характеристической функции в игре с непрерывным обновлением информации

Предположим, что функция $\tilde{V}_t(S; \tilde{x}_t^*(s), s, t + \bar{T})$, $S \subseteq N$, непрерывно дифференцируема по $s \in [t, t + \bar{T}]$ и интегрируема по $t \in [t_0, +\infty)$. Определим характеристическую функцию $V(S; x^*(t), t)$ в игре с непрерывным обновлением информации следующим образом:

Определение 4.3.6 Функция $V(S; x^*(t), t)$, $t \in [t_0, +\infty)$, $S \subseteq N$, является характеристической функцией с непрерывным обновлением информации, если она представляет собой интеграл

$$V(S; x^*(t), t) = \int_t^{+\infty} -\frac{d}{ds} \tilde{V}_\tau(S; \tilde{x}_\tau^*(s), s, \tau + \bar{T})|_{s=\tau} d\tau, \quad t \in [t_0, +\infty), \quad S \subseteq N, \quad (4.155)$$

где $\tilde{V}_\tau(S; \tilde{x}_\tau^*(s), s, t + \bar{T})$, $s \in [\tau, \tau + \bar{T}]$, $\tau \in [t, +\infty)$, $S \subseteq N$, — характеристическая функция в игре $\Gamma(\tilde{x}_\tau^*(s), s, \tau + \bar{T})$, определенной на временном интервале $[s, t + \bar{T}]$.

Обращаем внимание, что интеграл в (4.155) может быть бесконечным. Следовательно, необходимо определить кооперативную игру для конечного интервала времени, т.е. при $t \in [t_0, T]$:

$$V(S; x^*(t), t, T) = \int_t^T -\frac{d}{ds} \tilde{V}_\tau(S; \tilde{x}_\tau^*(s), s, \tau + \bar{T})|_{s=\tau} d\tau, \quad t \in [t_0, T], \quad S \subseteq N. \quad (4.156)$$

Чтобы доказать, что построенный дележ с непрерывным обновлением информации $\xi(x_0, t_0, T)$ (4.147) можно рассматривать как дележ в игре с непрерывным обновлением информации, покажем, что он представляет собой дележ для некоторой конкретной характеристической функции. Для этого используем характеристическую функцию с непрерывным обновлением информации $V(S; x_0, t_0, T)$ (4.156). Это обоснует применение дележа с непрерывным обновлением информации и характеристической функции с непрерывным обновлением информации одновременно.

Теорема 4.3.2 Дележ с непрерывным обновлением информации $\xi(x_0, t_0, T)$ является дележом в игре $\Gamma(x_0, t_0, T)$ с непрерывным обновлением информации и характеристической функцией $V(S; x^*(t), t, T)$ (4.156) если выполнено условие

$$\frac{d}{ds} \tilde{\xi}_i^t(x_t^*(s), s, t + \bar{T})|_{s=t} \leq \frac{d}{ds} \tilde{V}_t(\{i\}; x_t^*(s), s, t + \bar{T})|_{s=t}. \quad (4.157)$$

Доказательство Сначала покажем, что $\forall t \in [t_0, T]$ справедливы соотноше-

ния

$$\sum_{i=1}^n \xi_i(x^*(t), t, T) = V(N; x^*(t), t, T), \quad (4.158)$$

$$\xi_i(x^*(t), t, T) \geq V(\{i\}; x^*(t), t, T), \quad i \in N. \quad (4.159)$$

Согласно определению $\xi(x^*(t), t, T)$ и $V(S; x^*(t), t, T)$, правую часть (4.158) можно записать в виде

$$\begin{aligned} \sum_{i=1}^n \xi_i(x^*(t), t, T) &= \sum_{i=1}^n \int_t^T \beta_i(s, x^*(s)) ds = \\ &= \sum_{i=1}^n \int_t^T -\frac{d}{dl} \tilde{\xi}_i^s(x_s^*(l), l, s + \bar{T}) |_{l=s} ds. \end{aligned} \quad (4.160)$$

В силу определения характеристической функции, правую часть (4.158) можно записать в виде

$$V(N; x^*(t), t, T) = \int_t^T -\frac{d}{dl} \tilde{V}_s(N; x_s^*(l), l, s + \bar{T}) |_{l=s} ds. \quad (4.161)$$

Поскольку

$$\sum_{i=1}^n \tilde{\xi}_i^s(x_s^*(l), l, s + \bar{T}) = \tilde{V}_s(N; x_s^*(l), l, s + \bar{T}), \quad (4.162)$$

правые части выражений (4.160) и (4.161) являются эквивалентными, и условие (4.158) выполнено.

Теперь, выведем формулу (4.159), подставляя $\xi(x^*(t), t, T)$ и $V(\{i\}; x^*(t), t, T)$ в левую часть неравенства (4.159):

$$\xi_i(x^*(t), t, T) = \int_t^T \beta_i(s, x^*(s)) ds = \int_t^T -\frac{d}{dl} \tilde{\xi}_i^s(x_s^*(l), l, s + \bar{T}) |_{l=s} ds. \quad (4.163)$$

С учетом формулы (4.163) для дележа, формулы (4.161) для характеристической функции при $N = \{i\}$ и условия (4.157), приходим к формуле (4.159). Теорема доказана. \square

В этом разделе введено понятие характеристической функции с непрерывным обновлением информации $V(S; x_0, t_0, T)$. Показано, что дележ с непре-

рывным обновлением информации $\xi(x_0, t_0, T)$ действительно является дележом в классическом смысле, когда выполняются дополнительные условия (4.157). Стоит отметить, что условие (4.157) не всегда справедливо. В целом трудно сказать, когда условие (4.157) имеет место: это зависит как от типа дележа (вектор Шепли, пропорциональный дележ и т. д.), так и от типа характеристической функции ($\alpha, \beta, \gamma, \sigma$ и т. д.). Основная идея этого результата — показать, что в игре с непрерывным обновлением информации можно (при условиях (4.157)) использовать характеристическую функцию с непрерывным обновлением информации (4.156) для построения дележа с непрерывным обновлением информации без прямого применения подхода, представленного в Разделе 4.3.1.3.4.

4.3.1.6 Результирующее решение как функция результирующей характеристической функции

В этом разделе показано, что если в каждый момент времени $t \in [t_0, T]$ игроки выбирают дележ $\tilde{\xi}_t(\tilde{x}_t^*(s), s, t + \bar{T}) \in \tilde{M}_t(\tilde{x}_t^*(s), s, t + \bar{T})$ на основе $\tilde{V}_t(S; \tilde{x}_t^*(s), s, t + \bar{T})$, $s \in [t, t + \bar{T}]$, по одинаковому правилу, то соответствующий дележ с непрерывным обновлением информации $\xi(x^*(t), t, T)$ соответствует дележу, выбранному по тому же правилу с использованием характеристической функции с непрерывным обновлением информации $V(S; x^*(t), t, T)$. Ниже это свойство будет установлено для ряда принципов оптимальности.

Докажем его для всех типов дележей $\tilde{\xi}_t(\tilde{x}_t^*(s), s, t + \bar{T})$, представляющих собой конечно аддитивные функции от $\tilde{V}_t(S; \tilde{x}_t^*(s), s, t + \bar{T})$, $S \subseteq N$, $t \in [t_0, T]$. В самом деле, большинство классических дележей — конечно аддитивные функции значений характеристических функций. В частности, вектор Шепли является аддитивной функцией от характеристической функции $V(S)$, $S \subseteq N$, согласно первой аксиоме линейности:

$$Sh_i(V) = \sum_{S \subseteq N, i \in S} \frac{(s-1)!(n-s)!}{n!} [V(S) - V(S \setminus i)], \quad i \in N. \quad (4.164)$$

Обозначим через $\xi(\tilde{V}_t; \tilde{x}_t^*(s), s, T) = \tilde{\xi}_t(\tilde{x}_t^*(s), s, t + \bar{T})$ дележ, вычисленный с помощью значений характеристической функции $\tilde{V}_t(S; \tilde{x}_t^*(s), s, t + \bar{T})$, $S \subseteq N$, а

через $\xi(V; x^*(t), t, T)$ — тот же тип дележа (вектор Шепли и т.д.), вычисленный с помощью значений характеристической функции с непрерывным обновлением информации $V(S; x^*(t), t, T)$, $S \subseteq N$.

Приведенная ниже теорема для линейных аддитивных дележей (например, вектора Шепли в частном случае) показывает, что для расчета кооперативного решения в игре с непрерывным обновлением информации нет необходимости вычислять дележ для каждого текущего момента времени $t \in [t_0, T]$. Достаточно вычислить значения характеристической функции с непрерывным обновлением информации $V(S; x^*(t), t, T)$, $S \subseteq N$ и использовать их для вычисления соответствующего кооперативного решения как характеристической функции.

Теорема 4.3.3 *Предположим, что в каждой игре $\Gamma(x^*(t), t, t + \bar{T})$ в качестве дележа используется любая конечно аддитивная функция $\xi(\tilde{V}_t; \tilde{x}_t^*(s), s, t + \bar{T}) = \tilde{\xi}_t(\tilde{x}_t^*(s), s, t + \bar{T})$. Тогда соответствующий дележ $\xi(x^*(t), t, T)$ (4.147) совпадает с таким же дележом $\xi(V; x^*(t), t, T)$, вычисленным с помощью значений характеристической функции с непрерывным обновлением информации.*

Доказательство Необходимо показать, что

$$\xi(x^*(t), t, T) = \xi(V; x^*(t), t, T). \quad (4.165)$$

Запишем левую часть формулы (4.165) в виде

$$\xi(x^*(t), t, T) = \int_t^T \beta(x^*(s), s) ds = \int_t^T -\frac{d}{dl} \tilde{\xi}_s(\tilde{x}_s^*(l), l, s + \bar{T}) |_{l=s} ds. \quad (4.166)$$

Кроме того, запишем выражение для V в правой части формулы (4.165) в виде

$$\begin{aligned} V(S; x^*(t), t, T) &= \int_t^T -\frac{d}{dl} \tilde{V}_s(S; \tilde{x}_s^*(l), l, s + \bar{T}) |_{l=s} ds = \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n \left(-\frac{d}{dl} \tilde{V}_{y_i}(S; \tilde{x}_{y_i}^*(l), l, y_i + \bar{T}) |_{l=y_i} \right) (s_{i+1} - s_i), \end{aligned} \quad (4.167)$$

где $t = s_0 < s_1 < \dots < s_i < s_{i+1} < \dots < s_n = T$ и $y_i \in (s_i, s_{i+1})$.

С учетом свойства аддитивной функции, получаем:

$$\begin{aligned}
\xi(V; x^*(t), t, T) &= \\
&= \xi \left(\lim_{n \rightarrow \infty} \sum_{i=1}^n \left(-\frac{d}{dl} \tilde{V}_{y_i}(S; \tilde{x}_{y_i}^*(l), l, y_i + \bar{T}) \Big|_{l=y_i} \right) (s_{i+1} - s_i); x^*(t), t, T \right) = \\
&= \lim_{n \rightarrow \infty} \sum_{i=1}^n \left(-\frac{d}{dl} \xi(\tilde{V}_{y_i}; \tilde{x}_{y_i}^*(l), l, y_i + \bar{T}; x^*(t), t, T) \Big|_{l=y_i} \right) (s_{i+1} - s_i) = \\
&= \lim_{n \rightarrow \infty} \sum_{i=1}^n \left(-\frac{d}{dl} \tilde{\xi}_{y_i}(\tilde{x}_{y_i}^*(l), l, y_i + \bar{T}) \Big|_{l=y_i} \right) (s_{i+1} - s_i). \quad (4.168)
\end{aligned}$$

Теорема доказана. □

4.3.2 Уравнения Гамильтона–Якоби–Беллмана для кооперативной игры с непрерывным обновлением информации

В этом разделе рассмотрим только позиционные кооперативные стратегии с непрерывным обновлением информации $u_{fb}^*(t, x)$. Поэтому будем использовать упрощенное обозначение $u^*(t, x)$.

Чтобы построить уравнение Гамильтона–Якоби–Беллмана для кооперативных стратегий с непрерывным обновлением информации, необходимо определить обобщенные позиционные кооперативные стратегии $\tilde{u}^*(t, x; s, x_t)$ в игре с непрерывным обновлением информации. В рамках подхода с непрерывным обновлением информации, функция Беллмана $W(t; s, x)$ определяется как кооперативный выигрыш всех игроков в позиционных кооперативных стратегиях $\tilde{u}^*(t, x; s, x_t)$ в подыгре, начинающейся в момент времени s из состояния x в игре, определенной на временном интервале $[t, t + \bar{T}]$:

$$W(t; s, x) = \sum_{i \in N} \int_s^{t+\bar{T}} g^i[\tau, x_t(\tau), \tilde{u}^*(t, x; s, x_t(\tau))] d\tau \quad (4.169)$$

при условии

$$\begin{aligned}
\dot{x}_t(\tau) &= f(\tau, x_t, \tilde{u}^*), \\
x_t(s) &= x.
\end{aligned} \quad (4.170)$$

Имеет место следующая теорема.

Теорема 4.3.4 *Стратегии из ситуации $\tilde{u}^*(t, x; s, x_t)$ являются обобщенными кооперативными позиционными стратегиями в дифференциальной игре с непрерывным обновлением информации $\Gamma(x_0, t_0, \bar{T})$, если существует непрерывно дифференцируемая по s и x функция $W(t; s, x) : [t_0, +\infty) \times [t, t + \bar{T}] \times R \rightarrow R$, $i \in N$, которая удовлетворяет следующей системе уравнений в частных производных:*

$$\begin{aligned} -W_s(t; s, x) &= \max_{\phi_i} \left\{ \sum_{i \in N} g^i(s, x, u) + W_x(t; s, x) f(s, x, u) \right\} = \\ &= \sum_{i \in N} g^i(s, x, \tilde{u}^*) + W_x(t; s, x) f(s, x, \tilde{u}^*), \end{aligned} \quad (4.171)$$

$$W(t; t + \bar{T}, x) = 0.$$

Доказательство Согласно определению обобщенных кооперативных стратегий $\tilde{u}^*(t, x; s, x_t)$, стратегии из ситуации $\tilde{u}^*(t, x; s, x_t)$ должны быть кооперативными стратегиями для любого фиксированного t .

Зафиксировав t в формулировке Теоремы 4.3.4 и, в частности, в (4.171), получим классические достаточные условия для равновесия по Нэшу с непрерывным обновлением информации в позиционных стратегиях в дифференциальной игре с заданной продолжительностью $[t, t + \bar{T}]$ (см. [80]). Следовательно, для любого фиксированного t выполнены условия определения обобщенных кооперативных стратегий. Теорема доказана. \square

4.3.2.1 Модель добычи невозобновляемого ресурса

В качестве наглядного примера рассмотрим дифференциальную игру с непрерывным обновлением информации для добычи невозобновляемого ресурса (см. [50]). Предположим, что n игроков эксплуатируют природный ресурс, который находится в общей собственности и не восстанавливается с течением времени (например, природный газ или полезные ископаемые).

4.3.2.1.1 Исходная игра

Пусть вектор состояния $x(t)$ описывает запас ресурса в момент времени t , доступный для извлечения игроками. Обозначим через $u_i(t)$ скорость извлечения ресурса игроком i в момент времени t . Предположим, что $u_i(t) \geq 0$ и $u_i(t) = 0$ при $x(t) = 0$.

Динамика запаса задается дифференциальным уравнением

$$\dot{x}(t) = - \sum_{i=1}^n a_i u_i(t, x), \quad x(t_0) = x_0, \quad (4.172)$$

где $a_i > 0$ при всех $i = 1, \dots, n$, а $x_0 > 0$. Будем считать, что функции полезности игроков являются логарифмическими:

$$h_i(x(t), u_i(t)) = \ln u_i(t, x), \quad i = 1, \dots, n. \quad (4.173)$$

Функция выигрыша игрока i имеет вид

$$K_i(x_0, T - t_0) = \int_{t_0}^T \ln u_i(s, x) ds, \quad i = 1, \dots, n. \quad (4.174)$$

Предположим, что информация об уравнениях движения и функциях выигрыша непрерывно обновляется во времени. В каждый момент времени $t \in [0, T]$ игрокам доступна информация только на временном интервале $[t, t + \bar{T}]$. Это означает, что в каждый момент времени они могут рассчитывать на устойчивость данного процесса только в течение периода времени \bar{T} .

Предположим, что игра разыгрывается по кооперативному сценарию, т.е. у игроков есть возможность сотрудничать для получения максимального кооперативного выигрыша:

$$\max_{u_1, \dots, u_n} \sum_{i=1}^n K_i(x_0, T - t_0) = \max_{u_1, \dots, u_n} \sum_{i=1}^n \int_{t_0}^T \ln u_i(s, x) ds. \quad (4.175)$$

4.3.2.2 Оптимальные кооперативные стратегии с непрерывным обновлением информации

Согласно Разделу 4.3, для определения оптимальных позиционных кооперативных стратегий в игре с непрерывным обновлением информации рассмотрим семейство вспомогательных подыгр $\Gamma(x, t, t + \bar{T})$ с продолжительностью \bar{T} , начинающихся в момент времени t из состояния x .

Динамические ограничения игры $\Gamma(x, t, t + \bar{T})$ задаются следующим образом:

$$\begin{aligned} \dot{x}_t(s) &= - \sum_{i=1}^n a_i u_i^t(s, x_t), \\ x_t(t) &= x, \\ x_t, x &\in \mathbb{R}^l, \quad u^t = (u_1^t, \dots, u_n^t), \quad u_i^t = u_i^t(s, x_t) \in U_i \subset \text{comp} \mathbb{R}^k, \quad s \in [t, t + \bar{T}]. \end{aligned} \quad (4.176)$$

Функция выигрыша игрока i в $\Gamma(x, t, t + \bar{T})$ имеет вид

$$K_i^t(x, t; u^t) = \int_t^{t+\bar{T}} \ln u_i^t(s, x_t) ds, \quad i \in N. \quad (4.177)$$

Чтобы определить позиционные кооперативные стратегии $\tilde{u}^*(t, s, x)$ во вспомогательной подыгре $\Gamma(x, t, t + \bar{T})$, используем динамическое программирование. Обозначим через $W(t, s, x)$ функцию Беллмана в подыгре, начинающейся в момент времени s :

$$W(t; s, x) = \max_{u_1^t, \dots, u_n^t} \sum_{i=1}^n K_i^t(x, s; u^t). \quad (4.178)$$

Уравнение Гамильтона–Якоби–Беллмана (ГЯБ) имеет следующий вид:

$$\begin{aligned} -\frac{\partial W(t; s, x)}{\partial s} &= \max_{u_1^t, \dots, u_n^t} \left\{ \sum_{i=1}^n \ln u_i^t(s, x) - \frac{\partial W(t; s, x)}{\partial x} \sum_{i=1}^n a_i u_i^t(s, x) \right\}. \\ \lim_{s \rightarrow t + \bar{T}} W(t; s, x) &= 0. \end{aligned} \quad (4.179)$$

Найдем решение уравнения (4.179) в виде $W(t; s, x) = A(t, s) \ln x + B(t, s)$.

Частные производные этой функции равны

$$\frac{\partial W(t; s, x)}{\partial s} = \dot{A}(t, s) \ln x + \dot{B}(t, s); \quad \frac{\partial W(t; s, x)}{\partial x} = \frac{A(t, s)}{x}. \quad (4.180)$$

Максимизируя выражение в правой части (4.179) и подставляя результат в (4.180), получаем:

$$\tilde{u}^*(t, x; s, x_t) = \frac{x}{a_i A(t, s)}. \quad (4.181)$$

Подстановка формул (4.180) и (4.181) в уравнение (4.179) дает систему дифференциальных уравнений

$$\begin{aligned} \dot{A}(t, s) \ln x + \dot{B}(t, s) &= n \ln A(t, s) - n \ln x + \ln a^N + n, \\ \lim_{s \rightarrow t + \bar{T}} A(t, s) \ln x + B(t, s) &= 0, \end{aligned} \quad (4.182)$$

где $a^N = \prod_{i=1}^n a_i$. Решение системы (4.182) имеет вид

$$\begin{aligned} A(s, t) &= n(t + \bar{T} - s), \\ B(s, t) &= -(t + \bar{T} - s)(\ln a^N + n \ln n(t + \bar{T} - s)), \\ t &\in [t_0, T), \quad s \in (t, t + \bar{T}). \end{aligned} \quad (4.183)$$

Таким образом, приходим к следующим оптимальным позиционным кооперативным стратегиям во вспомогательной подыгре $\Gamma(x, t, t + \bar{T})$:

$$u_i^*(t, x; s, x_t) = \frac{1}{W_x} = \frac{x}{a_i n(t + \bar{T} - s)}, \quad s \in [t, t + \bar{T}]. \quad (4.184)$$

Кроме того,

$$W(t; s, x) = n(t + \bar{T} - s) \ln \frac{x}{n(t + \bar{T} - s)} - (t + \bar{T} - s) \ln a^N, \quad s \in [t, t + \bar{T}]. \quad (4.185)$$

Оптимальная кооперативная траектория во вспомогательной подыгре $\Gamma(x, t, t + \bar{T})$ имеет вид

$$\tilde{x}_t^*(s) = x \frac{(t + \bar{T} - s)}{\bar{T}}. \quad (4.186)$$

Следуя процедуре (4.139), строим оптимальные позиционные кооперативные стратегии с непрерывным обновлением информации:

$$u^*(t, x) = \tilde{u}^*(t, x; s, x_t)|_{s=t} = \frac{x}{a_i n \bar{T}}, \quad s \in [t, t + \bar{T}]. \quad (4.187)$$

Оптимальная кооперативная траектория $x^*(t)$ с непрерывным обновлением

информации равна

$$x^*(t) = x_0 e^{\frac{t_0-t}{\bar{T}}}.$$

Полученные результаты графически представлены на рисунках 4.17 и 4.18.

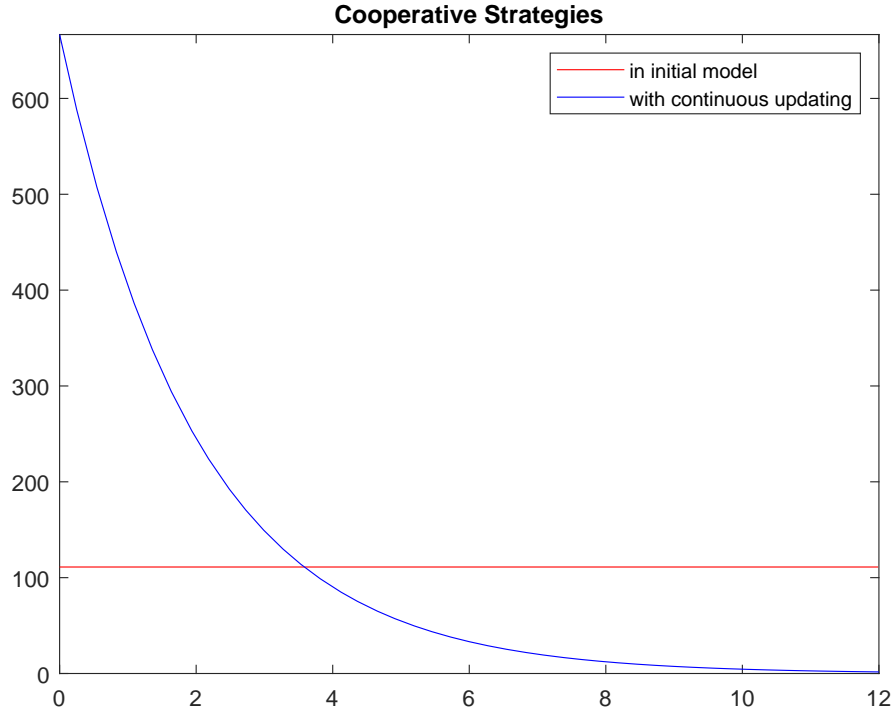


Рис. 4.17: Кооперативные стратегии в исходной игре (красная линия) и кооперативные стратегии с непрерывным обновлением информации $u^*(t, x)$ (синяя линия).

4.3.2.2.1 Характеристическая функция на временном интервале $[t, t + \bar{T}]$

Построим характеристическую функцию в подыграх $\Gamma(x, t, t + \bar{T})$, используя подход, предложенный в [228]:

$$\tilde{V}_t(S; x, s, t + \bar{T}) = \begin{cases} \sum_{i=1}^n K_i^t(x, s; \tilde{u}^*(t, x; s, x_t)), & S = N, \\ \max_{u_i, i \in S} \sum_{i \in S} K_i^t(x, s; \tilde{u}^{NE \setminus u^S}), & S \subset N, \\ 0, & S = \emptyset, \end{cases} \quad (4.188)$$

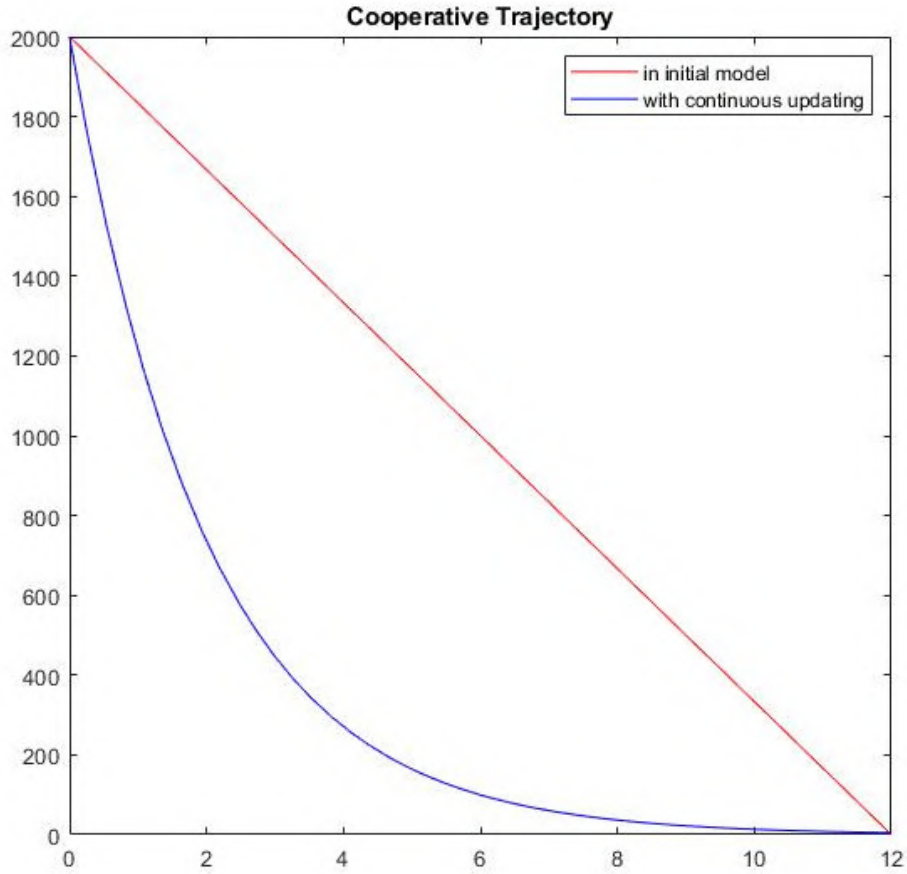


Рис. 4.18: Кооперативная траектория в исходной игре (красная линия) и кооперативная траектория с непрерывным обновлением информации $x^*(t)$ (синяя линия).

где $\max_{u_i, i \in S} \sum_{i \in S} K_i^t(x, s; \tilde{u}^{NE}/\tilde{u}^S)$ — максимальный кооперативный выигрыш коалиции S , при условии, что игроки из коалиции $N \setminus S$ используют стратегии некоторого фиксированного равновесия по Нэшу $\tilde{u}^{NE} = (\tilde{u}_1^{t, NE}, \dots, \tilde{u}_n^{t, NE})$ в подыгре $\Gamma(x, t, t + \bar{T})$.

Вообще говоря, такая характеристическая функция не является супераддитивной. Согласно (4.185), выигрыш максимальной коалиции равен

$$\tilde{V}_t(N; \tilde{x}_t^*(s), s, t + \bar{T}) = n(t + \bar{T} - s) \ln \frac{\tilde{x}_t^*(s)}{n(t + \bar{T} - s)} - (t + \bar{T} - s) \ln a^N.$$

Характеристическая функция $\tilde{V}_t(S; x, s, t + \bar{T})$ коалиции S строится в два этапа следующим образом. Во-первых, вычисляем равновесные по Нэшу

стратегии для всех игроков. Во-вторых, фиксируя эти стратегии для игроков из $N \setminus S$, считаем, что игроки из S стремятся максимизировать свой кооперативный выигрыш.

Для нахождения позиционных равновесных по Нэшу стратегий в подыгре $\Gamma(x, t, t + \bar{T})$ используется динамическое программирование. Обозначим через $W^i(t; s, x)$ функцию Беллмана, представляющую выигрыш игрока i в ситуации равновесия по Нэшу во вспомогательной подыгре игры $\Gamma(x, t, t + \bar{T})$, которая начинается в момент времени s .

Система уравнений Гамильтона–Якоби–Беллмана (ГЯБ) имеет следующий вид:

$$-\frac{\partial W^i(t; s, x)}{\partial s} = \max_{u_i} (\ln u_i^t(s, x) - \frac{\partial W^i(t; s, x)}{\partial x} (a_i u_i + \sum_{j \neq i} a_j \tilde{u}_j^{NE})), \quad i \neq j, \quad (4.189)$$

$$\lim_{s \rightarrow t + \bar{T}} W^i(t; s, x) = 0, \quad i = 1, \dots, n.$$

Функции Беллмана $W_i(t; s, x)$ будем искать в виде $W_i(t; s, x) = A_i(t, s)x + B_i(t, s)$. Максимизируя выражения в правой части (4.189) и решая соответствующие дифференциальные уравнения относительно $A_i(t, s)$ и $B_i(t, s)$, получаем:

$$A_i(t, s) = \bar{T} + t - s,$$

$$B_i(t, s) = (\bar{T} + t - s)(-\ln(\bar{T} + t - s) - \ln a_i - n - 1).$$

$$\tilde{u}_i^{NE}(t, x; s, x_t) = \frac{x}{a_i(\bar{T} + t - s)}.$$

Отметим, что характеристическая функция коалиции i совпадает с $W^i(t; s, x)$:

$$V_t(\{i\}; \tilde{x}_t^*(s), s, t + \bar{T}) = W^i(t; s, x) = (\bar{T} + t - s) \left(\ln \frac{\tilde{x}_t^*(s)}{\bar{T} + t - s} - \ln a_i - n - 1 \right).$$

Теперь построим $\tilde{V}_i(S; x, s, \bar{T})$. Обозначим через $W^S(t; s, x)$ функцию Беллмана как максимальный выигрыш коалиции S в подыгре, начинающейся в момент времени s при условии, что игроки из коалиции $N \setminus S$ используют стратегии $\tilde{u}_i^{NE}(t, x; s, x_t)$. Находим характеристическую функцию в виде $W^S(t; s, x) =$

$A^S(t, s) \ln x + B^S(t, s)$ как решение уравнения Гамильтона–Якоби–Беллмана:

$$-\frac{\partial W^S(t; s, x)}{\partial s} = \max_{u_i \in S} \left(\sum_{i \in S} \ln u_i^t(s, x) - \frac{\partial W^S(t; s, x)}{\partial x} \left(\sum_{i \in S} a_i u_i + \sum_{j \in N \setminus S} a_j \tilde{u}_j^{NE} \right) \right),$$

$$S \subset N,$$

$$\lim_{s \rightarrow t + \bar{T}} W^S(t; s, x) = 0. \quad (4.190)$$

Уравнения (4.191) дают стратегии для игроков $i \in S$, максимизирующие выражение в правой части (4.190):

$$\tilde{u}_i^S(t, x; s, x_t) = \frac{x}{a_i A^S(t, s)}. \quad (4.191)$$

Положим $|S| = k$. Подставляя стратегии (4.191) в (4.190), получаем следующую систему дифференциальных уравнений:

$$\begin{aligned} \dot{A}^S(t, s) &= -k, \\ \dot{B}^S(t, s) &= k \ln A^S(t, s) + \ln a^S + k + \frac{A^S(t, s)(n-k)}{T+t-s}, \\ \lim_{s \rightarrow t + \bar{T}} (A^S(t, s) \ln x + B^S(t, s)) &= 0, \end{aligned} \quad (4.192)$$

где $a^S = \prod_{i \in S} a_i$. Решением системы (4.192) является

$$\begin{aligned} A^S(s, t) &= k(t + \bar{T} - s), \\ B^S(s, t) &= -k(t + \bar{T} - s) \left(\ln k(t + \bar{T} - s) + \frac{\ln a^S}{k} + n - k \right), \\ t &\in [t_0, T), \quad s \in (t, t + \bar{T}). \end{aligned} \quad (4.193)$$

Стратегии игроков из коалиции S , максимизирующие кооперативный выигрыш этой коалиции, имеют следующий вид:

$$u_i^S(t, x; s, x_t) = \frac{x}{a_i k(t + \bar{T} - s)}, \quad s \in [t, t + \bar{T}). \quad (4.194)$$

Значение характеристической функции коалиции S равно

$$\begin{aligned} \tilde{V}_t(S; \tilde{x}_t^*(s), s, t + \bar{T}) &= W^S(t; s, x) = -k(t + \bar{T} - s) \\ &\quad \left(-\ln \tilde{x}_t^*(s) + \ln k(t + \bar{T} - s) + \frac{\ln a^S}{k} + n - k \right), \quad s \in [t, t + \bar{T}). \end{aligned} \quad (4.195)$$

Чтобы продемонстрировать супераддитивность построенной характеристической функции, положим $S \subset N$, $P \subset N$, и $S \cap P = \emptyset$. Тогда:

$$\begin{aligned} \tilde{V}_t(S \cup P; x, s, t + \bar{T}) - \tilde{V}_t(S; x, s, t + \bar{T}) - \tilde{V}_t(P; x, s, t + \bar{T}) &= \\ &= (t + \bar{T} - s) (k \ln k + p \ln p - k \ln(k + p) - p \ln(k + p) + 2pk) = \\ &= (t + \bar{T} - s) \left(k \left(\ln \frac{k}{k+p} + p \right) + p \left(\ln \frac{p}{k+p} + k \right) \right). \end{aligned} \quad (4.196)$$

Отметим, что при всех $k > 1$ и $p > 1$ справедливы неравенства $\frac{k+p}{k} < e^p$ и $\frac{k+p}{p} < e^k$. Следовательно, $(\ln \frac{k}{k+p} + p) > 0$ и $(\ln \frac{p}{k+p} + k) > 0$. Таким образом, мы доказали, что

$$\tilde{V}_t(S \cup P; x, s, t + \bar{T}) - \tilde{V}_t(S; x, s, t + \bar{T}) - \tilde{V}_t(P; x, s, t + \bar{T}) \geq 0 \quad (4.197)$$

при всех $s \in [t, t + \bar{T}]$. Это означает, что характеристическая функция супераддитивна.

4.3.2.2.2 Вектор Шепли с непрерывным обновлением информации

Для нахождения кооперативного решения можно применить любой из известных принципов оптимальности с помощью супераддитивной характеристической функции. Построим вектор Шепли [229] для этой игры.

$$\begin{aligned} \tilde{sh}_i^t(\tilde{x}_t^*(s), \bar{T} + t - s) &= \sum_{K \subseteq N, i \in K} \frac{(k-1)!(n-k)!}{n!} [(\tilde{V}_t(K; \tilde{x}_t^*(s), \bar{T} + t - s)) - \\ &\quad - (\tilde{V}_t(K \setminus i; \tilde{x}_t^*(s), \bar{T} + t - s))] = \\ &= (t + \bar{T} - s) \sum_{K \subseteq N, i \in K} \frac{(k-1)!(n-k)!}{n!} \left[\ln \frac{x}{t + \bar{T} - s} - \right. \\ &\quad \left. - k \ln k + (k-1) \ln(k-1) - \ln a_i + 2k - n - 1 \right] = \\ &= (t + \bar{T} - s) \left(\ln \frac{\tilde{x}_t^*(s)}{t + \bar{T} - s} - \ln a_i - \ln n \right). \end{aligned} \quad (4.198)$$

Теперь рассмотрим кооперативное решение дифференциальной игры с непрерывным обновлением информации. Согласно (4.147), построим вектор Шепли с непрерывным обновлением информации, используя вектор Шепли (4.198)

вспомогательной подыгры:

$$sh_i(x^*(t), T-t) = \int_t^T -\frac{d\widetilde{sh}_i^\tau(\widetilde{x}_\tau^*(s), \overline{T} + \tau - s)}{ds} \Big|_{s=\tau} d\tau =$$

$$= (T-t) \left(\ln \frac{x_0}{a_i n \overline{T}} + \frac{t_0}{\overline{T}} - \frac{t+T}{2\overline{T}} \right). \quad (4.199)$$

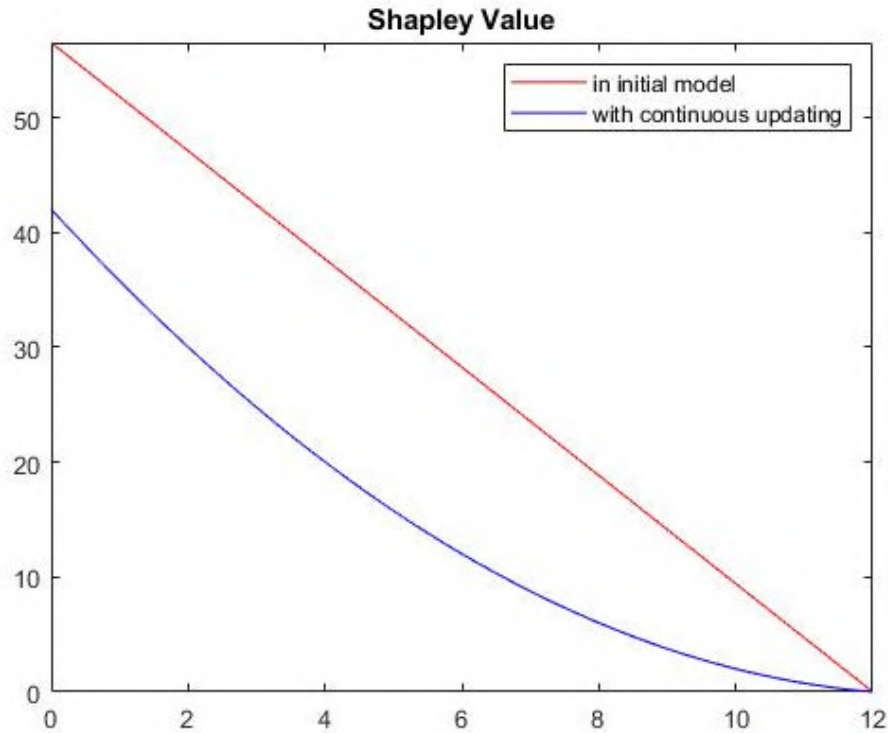


Рис. 4.19: Вектор Шепли игрока i в исходной игре (красная линия) и вектор Шепли с непрерывным обновлением информации $sh_i(x^*(t), T-t)$ (синяя линия).

Обращаем внимание, что условия (4.157) выполняются для построенного вектора Шепли с непрерывным обновлением информации. Таким образом, этот вектор представляет собой дележ в дифференциальной игре с непрерывным обновлением информации.

Такой же результат можно получить, применив характеристическую функцию дифференциальной игры с непрерывным обновлением информации, по-

строенную по правилу (4.155):

$$\begin{aligned}
 V(S, x^*(t), t, T - t) &= \int_t^T -\frac{d\tilde{V}_\tau(S; \tilde{x}^*(s), s, \tau + \bar{T})}{ds} \Big|_{s=\tau} d\tau = \\
 &= k(T - t) \left(\ln \frac{x_0}{k\bar{T}} + \frac{\ln a^S}{k} + \frac{t_0}{\bar{T}} - \frac{t + T}{2\bar{T}} - n + k \right). \quad (4.200)
 \end{aligned}$$

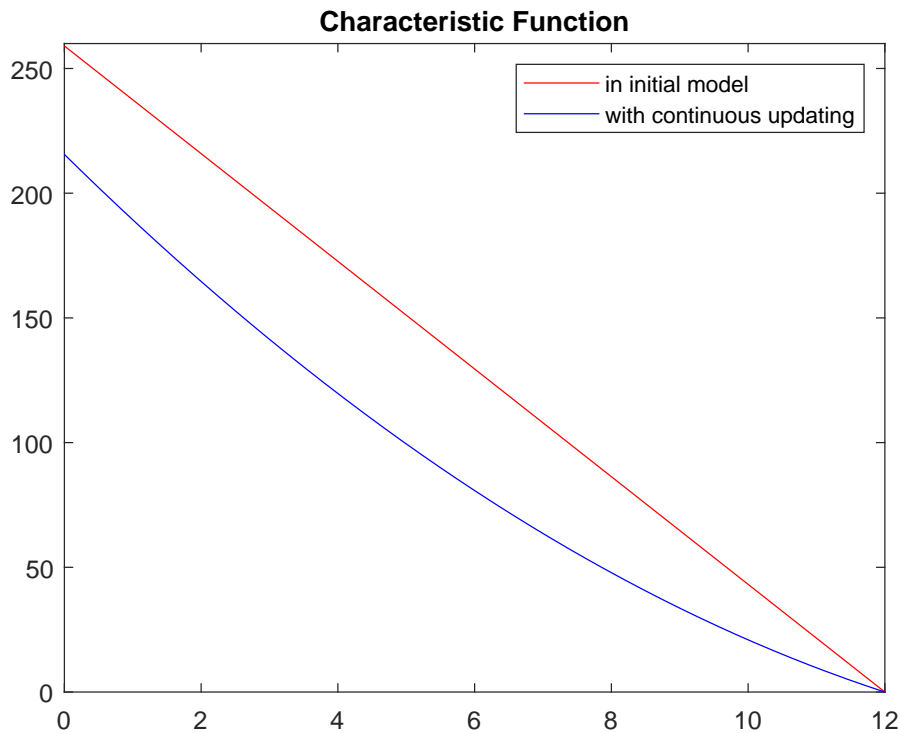


Рис. 4.20: Значение характеристической функции коалиции N в исходной игре (красная линия) и значение характеристической функции коалиции N с непрерывным обновлением информации $V(N, x^*(t), t, T - t)$ (синяя линия).

4.3.3 Принцип максимума Понтрягина для кооперативной дифференциальной игры с непрерывным обновлением информации

4.3.3.1 Кооперативная дифференциальная игра с непрерывным обновлением информации

В этом разделе принцип максимума Понтрягина применяется для модели кооперативной дифференциальной игры с непрерывным обновлением информации.

В кооперативной постановке, перед началом игры все игроки договариваются вести себя оптимальным образом сообща (сотрудничать). Но прежде чем перейти к этому шагу, представим тип характеристической функции, которая будет использоваться ниже. В данном разделе рассматриваем только программные кооперативные стратегии с непрерывным обновлением информации $u_{ol}^*(t, x)$, поэтому будем использовать упрощенное обозначение $u^*(t, x)$.

4.3.3.1.1 Подход к определению характеристической функции на временном интервале $[s, t + \bar{T}]$

Введем понятие характеристической функции $\tilde{V}_t(S; x, s, t + \bar{T})$, $\forall S \subseteq N$, определенной для каждой подыгры $\Gamma(x_{t,s}, s, t + \bar{T})$, где $s \in [t, t + \bar{T}]$ и $t \in [t_0, +\infty)$. Прежде чем вводить характеристическую функцию для подыгры $\Gamma(x_{t,s}, s, t + \bar{T})$, отметим, что $x_{t,s}$ зависит от начальной точки x . (Этот факт можно вывести из уравнения (4.5).) Следовательно, можем записать подыгру в виде $\Gamma(x, s, t + \bar{T})$, а функцию выигрыша для каждого игрока i в подыгре — в виде $K_i^t(x, s, t + \bar{T}; u)$. Для характеристической функции воспользуемся конструктивным подходом, предложенным Л. Петросяном и Г. Закуром [230] для построения δ -характеристической функции. Обозначим через $V(S; x^*(t), t, T)$ силу коалиции S в подыгре $\Gamma(x^*(t), t, T)$. Ее можно вычислить в два этапа: сначала необходимо найти равновесные по Нэшу стратегии $\{u_i^{NE}\}$ для всех игроков $i \in N$, а затем зафиксировать (заморозить) их для игроков из $N \setminus S$ при условии, что игроки из коалиции S стремятся максимизировать свой кооперативный выигрыш $\sum_{i \in S} K_i$ по $u_S = \{u_i\}_{i \in S}$. Следовательно, характеристическая функция имеет вид

$$\tilde{V}_t(S; x, s, t + \bar{T}) = \begin{cases} \max_{u_1, \dots, u_n} \sum_{i=1}^n K_i^t(x, s, t + \bar{T}; u_1, \dots, u_n), & S = N, \\ \max_{u_i, i \in S} \sum_{i \in S} K_i^t(x, s, t + \bar{T}; u_S, u_{N \setminus S}^{NE}), & S \subset N, \\ 0, & S = \emptyset, \end{cases}$$

где $x = x_t(t)$, $t \in [t_0, +\infty)$, и $s \in [t, t + \bar{T}]$ описываются (4.5). Кроме того, $u_S = \{u_i\}_{i \in S}$ — стратегии игроков из коалиции S .

Предположим, что характеристическая функция $\tilde{V}_t(S; x, s, t + \bar{T})$ удовлетво-

ряет условиям супераддитивности:

$$\begin{aligned} \tilde{V}_t(S_1 \cup S_2; x, s, t + \bar{T}) &\geq \tilde{V}_t(S_1; x, s, t + \bar{T}) + \tilde{V}_t(S_2; x, s, t + \bar{T}), \\ \forall S_1, S_2 \subseteq N, S_1 \cap S_2 &= \emptyset. \end{aligned} \quad (4.201)$$

4.3.3.1.2 Алгоритм вычисления характеристической функции с непрерывным обновлением информации и вектора Шепли

На шагах 1–3 находятся необходимые элементы для определения характеристической функции. На следующем шаге вычисляется вектор Шепли.

Шаг 1: Оптимизация общих выплат гранд коалиции с непрерывным обновлением информации

Обозначим через $\Gamma^c(x, t, t + \bar{T})$ описанную выше кооперативную дифференциальную игру. Продолжительность этой игры равна \bar{T} . Предположим, что нет никаких препятствий для сотрудничества между игроками и что их выигрыши являются трансферабельными. Точнее, будем считать, что до фактического начала игры игроки соглашаются играть сообща (кооперация).

Теорема 4.3.5 Пусть (i) $f(s, \cdot, u(t, s))$ непрерывно дифференцируема в пространстве \mathbb{R}^l , $\forall s \in [t, t + \bar{T}]$,

(ii) $g_i(\cdot, \cdot, u(t, s))$ непрерывно дифференцируема в пространстве $\mathbb{R} \times \mathbb{R}^l$, $\forall s \in [t, t + \bar{T}]$.

Если набор стратегий $\{\tilde{u}_i^*(t, s)\}_{i \in N}$ образует обобщенные программные кооперативные стратегии в дифференциальной игре с непрерывным обновлением информации, тогда при любом фиксированном $t \in [t_0, T]$ существует переменная сопряженного состояния $\psi^t(s)$ с $s \in [t, t + \bar{T}]$, удовлетворяющая следующим соотношениям:

$$(1) \dot{x}_t^*(s) = f(s, x_t^*(s), \tilde{u}^*(s)), x_t^*(t) = x^*(t), \forall s \in [t, t + \bar{T}],$$

$$(2) \tilde{u}_i^*(t, s) = \arg \max_{u_i \in U_i, i \in N} H^t(s, x_t^*(s), u(t, s), \psi^t(s)), \text{ где } s \in [t, t + \bar{T}], i \in N,$$

$$(3) \dot{\psi}^t(s) = -\frac{\partial}{\partial x_t} H^t(s, x_t(s), \tilde{u}^*(t, s), \psi^t(s)), \text{ где } s \in [t, t + \bar{T}],$$

$$\psi^t(t + \bar{T}) = 0.$$

Доказательство Зафиксируем $t \geq t_0$ и рассмотрим игру $\Gamma^c(x, t, t + \bar{T})$. Уравнения движения имеют вид

$$\dot{x}_t(s) = f(s, x_t(s), u(t, s)), \quad x_t(t) = x^*(t), \quad s \in [t, t + \bar{T}]. \quad (4.202)$$

Функция выигрыша гранд коалиции задается выражением

$$\sum_{i=1}^n K_i^t(x^*(t), t, t + \bar{T}; u_1(t, s), \dots, u_n(t, s)) = \sum_{i=1}^n \int_t^{t+\bar{T}} g_i[s, x_t(s), u(t, s)] ds. \quad (4.203)$$

Для задачи оптимизации (4.202)–(4.203), гамильтониан принимает вид

$$H^t(s, x_t(s), u(t, s), \psi^t(s)) = \sum_{i=1}^n g_i[s, x_t(s), u(t, s)] + \psi^t(s) f(s, x_t(s), u(t, s)), \\ s \in [t, t + \bar{T}]. \quad (4.204)$$

Если $\tilde{u}_i^*(t, s)$, $i \in N$, является обобщенной программной кооперативной стратегией в дифференциальной игре с непрерывным обновлением информации, то (см. Определение 4.3.1) при каждом фиксированном $t \geq t_0$, $\tilde{u}_i^*(t, s)$, $i \in N$, образует программную кооперативную стратегию в игре $\Gamma^c(x^*(t), t, t + \bar{T})$. Следовательно, при любом фиксированном $t \geq t_0$ условия (1)–(3) теоремы выполнены как необходимые условия для программных кооперативных стратегий (см. [231]). Теорема доказана. \square

С другой стороны, если при каждом $t \geq t_0$, гамильтониан H^t является вогнутым по $(x_t, u(t, s))$, то условия теоремы становятся и достаточными для кооперативного решения в программных стратегиях [232].

Тогда для любого фиксированного $t \in [t_0, +\infty)$ можно найти обобщенную программную кооперативную стратегию $\tilde{u}^*(t, s) = (\tilde{u}_1^*(t, s), \dots, \tilde{u}_n^*(t, s))$ и соответствующую траекторию состояния $x_t^*(s)$, $s \in [t, t + \bar{T}]$. Используя Определения 4.3.1 и 4.3.2, можно найти кооперативные стратегии с непрерывным обновлением информации $\{u_i^*(t)\}_{i \in N}$ и соответствующую кооперативную

траекторию с непрерывным обновлением информации $x^*(t)$, $t \in [t_0, +\infty)$.

Чтобы получить характеристическую функцию гранд коалиции в подыгре $\Gamma(x^*(t), s, t + \bar{T})$, подставим $\tilde{u}^*(t, \tau)$ и $x_t^*(\tau)$ в соответствующую функцию выигрыша и обозначим ее через $\tilde{V}_t(N; x^*(t), s, t + \bar{T})$ как характеристическую функцию коалиции N . Текущий максимальный кооперативный выигрыш $\tilde{V}_t(N; x^*(t), s, t + \bar{T})$ можно выразить в виде

$$\tilde{V}_t(N; x^*(t), s, t + \bar{T}) = \sum_{i=1}^n \int_s^{t+\bar{T}} g_i[\tau, x_t^*(\tau), \tilde{u}^*(t, \tau)] d\tau. \quad (4.205)$$

Шаг 2: Вычисление обобщенного равновесия по Нэшу в программных стратегиях с непрерывным обновлением информации

Некооперативную подыгру вдоль кооперативной траектории с непрерывным обновлением информации $\Gamma(x^*(t), s, t + \bar{T})$ можно сформулировать следующим образом:

$$\begin{aligned} & \max_{u_i \in U_i} K_i^t(x^*(t), s, t + \bar{T}; u_i(t, s), \tilde{u}_{-i}^{NE}(t, s)) = \\ & = \max_{u_i \in U_i} \int_s^{t+\bar{T}} g_i[\tau, x_t(\tau), u_i(t, \tau), \tilde{u}_{-i}^{NE}(t, \tau)] d\tau \quad \text{где } x_t(\tau) \text{ удовлетворяет (4.202)} \end{aligned} \quad (4.206)$$

где $\tilde{u}_{-i}^{NE}(t, \tau) = (\tilde{u}_1^{NE}(t, \tau), \dots, \tilde{u}_{i-1}^{NE}(t, \tau), \tilde{u}_{i+1}^{NE}(t, \tau), \dots, \tilde{u}_n^{NE}(t, \tau))$. В этой постановке, текущий гамильтониан можно записать в виде

$$H_i^t(\tau, x_t(\tau), u(t, \tau), \psi_i^t(\tau)) = g_i(\tau, x_t, u) + \psi_i^t(\tau) f(\tau, x_t, u), \tau \in [s, t + \bar{T}], i \in N, \quad (4.207)$$

где $\tau \in [s, t + \bar{T}]$ и $t \in [t_0, +\infty)$. Используя принцип максимума Понтрягина с непрерывным обновлением информации [120], можно получить равновесие по Нэшу в программных стратегиях $\{\tilde{u}_i^{NE}(t, \tau)\}_{i \in N}$, а также соответствующую траекторию $x_t^{NE}(\tau)$, $\forall \tau \in [s, t + \bar{T}]$, $t \in [t_0, +\infty)$. Тогда легко вывести характеристическую функцию коалиции из одного игрока для каждого $i =$

$1, 2, \dots, n :$

$$\tilde{V}_t(\{i\}; x^*(t), s, t + \bar{T}) = \int_s^{t+\bar{T}} g_i[\tau, x_t^{NE}(\tau), \tilde{u}^{NE}(t, \tau)] d\tau, \quad \forall s \in [t, t + \bar{T}], t \in [t_0, +\infty). \quad (4.208)$$

Шаг 3: Вычисление характеристической функции для всех возможных оставшихся коалиций с непрерывным обновлением информации

Здесь необходимо найти только коалиции, содержащие два и более игроков, за исключением гранд коалиции. В общей сложности будет $2^n - n - 2$ подмножеств, полученных следующим образом. Применим δ -характеристическую функцию, чтобы игроки из коалиции S максимизировали свой кооперативный выигрыш $\sum_{i \in S} K_i^t(x^*(t), s, t + \bar{T}; u_S, \tilde{u}_{N \setminus S}^{NE})$ вдоль кооперативной траектории с непрерывным обновлением информации $x^*(t)$ при условии, что остальные игроки (из $N \setminus S$) используют обобщенные программные равновесные по Нэшу стратегии

$$\tilde{u}_{N \setminus S}^{NE} = \{\tilde{u}_j^{NE}\}_{j \in N \setminus S}.$$

Таким образом, у нас есть двухэтапная процедура построения характеристической функции: (1) найти обобщенные программные равновесные по Нэшу стратегии с непрерывным обновлением информации $\tilde{u}_i^{NE}(t, \tau)$ для всех игроков $i \in N$ коалиции N , полученной на Шаге 2; (2) «заморозить» равновесные по Нэшу стратегии $\tilde{u}_j^{NE}(t, \tau)$ для игроков из $N \setminus S$ и, по аналогии с игроком из коалиции S , максимизировать их кооперативный выигрыш по $u_S = \{u_i\}_{i \in S}$. Чтобы проиллюстрировать функцию цены подыгры $\Gamma(x^*(t), s, t + \bar{T})$, $\forall t \in [t_0, +\infty), s \in [t, t + \bar{T}]$, введем следующее понятие.

Определение 4.3.7 *Набор стратегий $\tilde{u}_S^* = \{\tilde{u}_i^*(t, \tau)\}_{i \in S}$, $\tau \in [s, t + \bar{T}]$, образует обобщенные программные оптимальные стратегии коалиции $S \subset N$ в подыгре с непрерывным обновлением информации $\Gamma(x^*(t), s, t + \bar{T})$, если он является решением следующей задачи оптимизации, полученным с помощью*

принципа максимума Понтрягина:

$$\begin{aligned} & \max_{u_S \in U_S} \sum_{i \in S} K_i^t(x^*(t), s, t + \bar{T}; u_S, \tilde{u}_{N \setminus S}^{NE}) \\ & = \max_{u_S \in U_S} \sum_{i \in S} \int_s^{t + \bar{T}} g_i[\tau, x_t^S(\tau), u_S(t, \tau), \tilde{u}_{N \setminus S}^{NE}(t, \tau)] d\tau \end{aligned} \quad (4.209)$$

при условии $x_t^S(\tau) = f(\tau, x_t^S(\tau), u_S(t, \tau), \tilde{u}_{N \setminus S}^{NE}(t, \tau))$, $x_t^S(s) = x_{t,s}$.

При всех $S \subset N$ гамильтониан для задачи (4.209) имеет вид¹:

$$\begin{aligned} H_S^t(\tau, x_t^S(\tau), u_S(t, \tau), \tilde{u}_{N \setminus S}^{NE}(t, \tau), \psi_S^t) & = \sum_{i \in S} g_i(\tau, x_t^S, u_S, \tilde{u}_{N \setminus S}^{NE}) \\ & + \psi_S^t(\tau) f(\tau, x_t^S, u_S, \tilde{u}_{N \setminus S}^{NE}). \end{aligned} \quad (4.210)$$

Теорема 4.3.6 Пусть

(i) $f(\tau, \cdot, u(t, \tau))$ непрерывно дифференцируема в пространстве \mathbb{R}^l , $\forall \tau \in [s, t + \bar{T}]$,

(ii) $g_i(\cdot, \cdot, u(t, \tau))$ непрерывно дифференцируема в пространстве $\mathbb{R} \times \mathbb{R}^l$.

Набор стратегий $\tilde{u}_S^* = \{\tilde{u}_i^*(t, \tau)\}_{i \in S}$ образует обобщенные программные оптимальные стратегии коалиции S в подыгре с непрерывным обновлением информации $\Gamma(x^*(t), s, t + \bar{T})$ задачи (4.209) если существуют $2^n - n - 2$ функций сопряженного состояния $\psi_S^t(\tau)$, где $\tau \in [s, t + \bar{T}]$ и $S \subset N$, удовлетворяющих $\forall s \in [t, t + \bar{T}]$ и $t \in [t_0, +\infty)$ следующим соотношениям:

$$(1) \dot{x}_t^S(\tau) = f(\tau, x_t^S, \tilde{u}_S^*(t, \tau), \tilde{u}_{N \setminus S}^{NE}), \quad x_t^S(s) = x_{t,s}, \quad \forall \tau \in [s, t + \bar{T}];$$

$$(2) \tilde{u}_S^*(t, \tau) = \arg \max_{u_S \in U_S} H_S(x_t^S(\tau), u_S(t, \tau), \tilde{u}_{N \setminus S}^{NE}(t, \tau), \psi_S^t(\tau)), \quad \forall \tau \in [s, t + \bar{T}],$$

где $\tilde{u}_S^*(t, \tau) = \{\tilde{u}_i^S(t, \tau)\}_{i \in S}$;

$$(3) \dot{\psi}_S^t(\tau) = -\frac{\partial}{\partial x_t} H_S^t(\tau, x_t(\tau), \tilde{u}_S^*(t, \tau), \tilde{u}_{N \setminus S}^{NE}(t, \tau), \psi_S^t(\tau)), \quad \text{где } \tau \in [s, t + \bar{T}] \text{ и } S \subset N;$$

$$\psi_S^t(t + \bar{T}) = 0, \quad S \subset N.$$

¹Верхний индекс S указывает на коалицию S , например, x_t^S , ψ_S^t и u_S .

Доказательство Аналогично доказательству Теоремы 4.3.5. □

Следовательно, игроки в коалиции S выберут обобщенное программное оптимальное управление $\tilde{u}_S^*(t, \tau)$, описанное в Теореме 4.3.6. Обращаем внимание, что эти управления являются функциями фиксированного текущего времени $t \in [t_0, +\infty)$ и вспомогательного времени $\tau \in [s, t + \bar{T}]$.

Характеристическая функция для коалиции $S \subseteq N$ имеет вид

$$\tilde{V}_t(S; x^*(t), s, t + \bar{T}) = \sum_{i \in S} \int_s^{t + \bar{T}} g_i[\tau, x_t^S(\tau), \tilde{u}_S^*(t, \tau), \tilde{u}_{N \setminus S}^{NE}(t, \tau)] d\tau, \quad (4.211)$$

$$\forall s \in [t, t + \bar{T}], \quad t \in [t_0, +\infty),$$

где $x_t^S(\tau)$ — траектория в момент времени $\tau \in [s, t + \bar{T}]$, когда игроки из коалиции S используют обобщенные программные оптимальные стратегии $\tilde{u}_S^*(t, \tau)$ при условии, что игроки из $N \setminus S$ используют обобщенное равновесие по Нэшу в программных стратегиях с непрерывным обновлением информации $\tilde{u}_{N \setminus S}^{NE}(t, \tau)$, которое уже было получено на Шаге 2.

Вернемся к характеристической функции в игре с непрерывным обновлением информации. В предположении, что функция $\tilde{V}_t(S; x^*(t), s, t + \bar{T})$, $\forall S \subseteq N$, непрерывно дифференцируема по $s \in [t, t + \bar{T}]$ и интегрируема по $t \in [t_0, +\infty)$, определим характеристическую функцию $V(S; x^*(t), t, T)$ в игре с непрерывным обновлением информации следующим образом:

Определение 4.3.8 *Функция $V(S; x^*(t), t, T)$, $t \in [t_0, T]$, $S \subseteq N$, является характеристической функцией в дифференциальной игре с непрерывным обновлением информации $\Gamma(x^*(t), t, T)$, если она задается в виде интеграла*

$$V(S; x^*(t), t, T) = \int_t^T -\frac{d}{ds} \tilde{V}_{\tau'}(S; x^*(\tau'), s, \tau' + \bar{T})|_{s=\tau'} d\tau', \quad t \in [t_0, T], \quad S \subseteq N, \quad (4.212)$$

где $\tilde{V}_{\tau'}(S; x^*(\tau'), s, \tau' + \bar{T})$, $s \in [\tau', \tau' + \bar{T}]$, $\tau' \in [t, T]$, $S \subseteq N$, — характеристическая функция в игре $\Gamma(x^*(\tau'), s, \tau' + \bar{T})$, определенной на временном интервале $[s, \tau' + \bar{T}]$.

В выражении (4.212) предполагаем, что интеграл берется на конечном временном интервале: только в этом случае можно утверждать, что значения характеристической функции с непрерывным обновлением информации конечны. Позже мы рассмотрим модельный пример и вычислим характеристическую функцию и вектор Шепли, используя метод конечных интервалов. Считаем, что выполнены условия супераддитивности:

$$\begin{aligned} V(S_1 \cup S_2; x^*(t), t, T) &\geq V(S_1; x^*(t), t, T) + V(S_2; x^*(t), t, T), \\ \forall S_1, S_2 \subseteq N, S_1 \cap S_2 &= \emptyset. \end{aligned} \quad (4.213)$$

Шаг 4: Вычисление вектора Шепли на основе характеристической функции с непрерывным обновлением информации

Снова рассмотрим модель кооперативной игры $\Gamma^c(x^*(t), t, T)$ с непрерывным обновлением информации. Пусть игрокам разрешено образовывать различные коалиции, состоящие из подмножеств множества всех игроков $K \subseteq N$. В подмножестве K существует k игроков. Множество дележей кооперативной игры $\Gamma^c(x^*(t), t, T)$ — это множество $L(x^*(t), t, T) = \{\xi(x^*(t), t, T) = (\xi_1(x^*(t), t, T), \dots, \xi_n(x^*(t), t, T))\}$, удовлетворяющее условиям

$$\begin{aligned} \xi_i(x^*(t), t, T) &\geq V(\{i\}; x^*(t), t, T), \quad i \in N; \\ \sum_{i \in N} \xi_i(x^*(t), t, T) &= V(N; x^*(t), t, T). \end{aligned} \quad (4.214)$$

Кооперативное решение или оптимальный принцип представляет собой непустое подмножество множества дележей $L(x^*(t), t, T)$. В частности, вектор Шепли $sh(x^*(t), t, T) = (sh_1(x^*(t), t, T), \dots, sh_n(x^*(t), t, T))$ — дележ, состоящий из компонент

$$sh_i(x^*(t), t, T) = \sum_{K \subseteq N, i \in K} \frac{(k-1)!(n-k)!}{n!} [V(K; x^*(t), t, T) - V(K \setminus i; x^*(t), t, T)], \quad (4.215)$$

где $K \setminus i$ обозначает относительное дополнение i в K , а функция $V(K; x^*(t), t, T)$ задана Определением 4.3.8 и является прибылью коалиции K . Между тем, $[V(K; x^*(t), t, T) - V(K \setminus i; x^*(t), t, T)]$ — это маржинальный вклад игрока i в коалицию K .

Существуют и другие принципы оптимальности для кооперативных игр, например, решение фон Неймана–Моргенштерна, N-ядро, и C-ядро. Во всех случаях они включают некоторые подмножества множества дележей игры.

4.3.3.2 Кооперативная дифференциальная игра управления загрязнением атмосферы

Рассмотрим следующую игру, предложенную Лонгом [233]. Пронумеруем страны с помощью индекса $i \in N$ и обозначим $n = |N|$. У каждого игрока есть промышленная производственная площадка. Предполагается, что загрязнение u_i прямо пропорционально объему производства. Таким образом, стратегия игрока заключается в выборе объема загрязняющих веществ, выбрасываемых в атмосферу.

4.3.3.2.1 Исходная игра

Загрязнение атмосферы накапливается с течением времени. Обозначим через $x(t)$ уровень загрязнения в момент времени t и предположим, что каждая страна «вносит свой вклад» в общий уровень загрязнения. Для простоты будем считать, что уровень загрязнения $x(t)$ изменяется согласно линейному дифференциальному уравнению

$$\dot{x}(t) = \sum_{i=1}^n u_i(t) - \delta x(t), \quad x(t_0) = x_0,$$

где δ — постоянная скорость разложения загрязняющих веществ (скорость их поглощения окружающей средой).

Пусть скорость поглощения δ равна нулю:

$$\dot{x}(t) = \sum_{i=1}^n u_i(t), \quad x(t_0) = x_0. \quad (4.216)$$

Загрязнение отрицательно сказывается на здоровье, качестве жизни и продуктивности людей. Предположим, что эти неблагоприятные эффекты могут быть

описаны мгновенной функцией общественного благосостояния F_i , с переменной x в качестве аргумента и отрицательной производной:

$$F_i = F_i(x, t, u_i), \quad \frac{\partial F_i}{\partial x} < 0.$$

В каждой стране совокупное общественное благосостояние считается неотъемлемой частью мгновенного общественного благосостояния. Таким образом, выигрыш игрока i можно представить следующим образом:

$$K_i(x_0, t_0, T; u) = \int_{t_0}^T F_i(x, t, u_i) dt. \quad (4.217)$$

Для обеспечения управляемости часто предполагается, что функция F_i является сепарабельной:

$$F_i(x, t, u_i(t)) = R_i(u_i(t)) - D_i(x),$$

где $R_i(u_i)$ можно рассматривать как производственную полезность (выгоду от производства), а $D_i(x)$ — как «пагубность» (вред), вызванную загрязнением окружающей среды. Следуя стандартной практике, будем считать, что функция $R_i(u_i)$ строго вогнутая и возрастает по u_i , а функция $D_i(x)$ выпукла и возрастает по x . Не исключается и случай линейной функции D_i .

Предположим, что стоимость ущерба окружающей среды, вызванного уровнем загрязнения x со стороны игрока i , равна $D_i(x) = d_i x$, и функция $D_i(x)$ возрастает выпукло. Следующие допущения широко известны в литературе по экономике окружающей среды: функция производственной выгоды для игрока i может быть выражена как функция объема выбросов, то есть $R_i(u_i(t)) = b_i u_i - \frac{1}{2} u_i^2$, причем $R_i(0) = 0$, а b_i и d_i — положительные параметры. Чтобы функция производственной выгоды имела вогнутый рост по объему выбросов, введем ограничение $u_i(t) \in (0, b_i)$.

Предположим, что игра протекает по кооперативному сценарию, то есть у игроков есть возможность сотрудничать для достижения максимального

суммарного выигрыша:

$$\sum_{i=1}^n K_i(x_0, t_0, T; u) = \sum_{i=1}^n \int_{t_0}^T ((b_i - \frac{1}{2}u_i)u_i - d_i x) dt \rightarrow \max_{u_1, u_2, \dots, u_n}. \quad (4.218)$$

Для решения задачи оптимизации (4.218), (4.216), используем принцип максимума Понтрягина следующим образом. Очевидно, что возникает игра с линейным состоянием.² Класс игр с линейным состоянием обладает очень полезным свойством: из линейности по переменным состояния вместе с независимой структурой взаимодействия между переменными состояния и управляющими переменными следует, что равновесие в программных стратегиях является совершенным по Маркову, а функции значения игры — линейными по переменным состояния.

Легко проверить, что оптимальное управление выбросами игрока i для исходной модели дифференциальной игры дается выражением

$$u_i^*(t) = b_i - \sum_{i=1}^n d_i(T - t), \quad i \in N. \quad (4.219)$$

Чтобы получить кооперативную траекторию состояния для исходной дифференциальной игры, достаточно применить управление $u_i^*(t)$ в системе и решить соответствующее дифференциальное уравнение. В результате получаем:

$$x^*(t) = x_0 + \left(\sum_{i=1}^n b_i - n \sum_{i=1}^n d_i T \right) (t - t_0) + n \sum_{i=1}^n d_i \frac{t^2 - t_0^2}{2}. \quad (4.220)$$

²Игры, в которых динамика системы и функции полезности являются полиномами 1-й степени по переменным состояния и которые удовлетворяют определенному свойству взаимодействия между управляющими переменными и переменными состояния (см. ниже), называются играми с линейным состоянием.

4.3.3.2.2 Игра управления загрязнением с непрерывным обновлением информации

В игре $\Gamma(x, t, t + \bar{T})$ динамика общего уровня загрязнения $x_t(s)$ описывается дифференциальным уравнением

$$\dot{x}_t(s) = \sum_{i=1}^n u_i(t, s), \quad x_t(t) = x, \quad (4.221)$$

в предположении, что скорость поглощения, соответствующая естественному очищению атмосферы, равна нулю.

Мгновенный выигрыш игрока i имеет вид

$$R_i(u_i(t, s)) = b_i u_i(t, s) - \frac{1}{2} u_i^2(t, s), \quad i \in N.$$

Каждый игрок должен нести расходы по устранению загрязнений. Таким образом, мгновенный выигрыш (полезность) игрока i игрока равен $R_i(u_i(t, s)) - d_i x_t(s)$, где $d_i > 0$.

Таким образом, выигрыш игрока i определяется следующим образом:

$$K_i^t(x, t, t + \bar{T}; u) = \int_t^{t+\bar{T}} \left((b_i - \frac{1}{2} u_i) u_i - d_i x_t \right) ds, \quad (4.222)$$

где $u_i = u_i(t, s)$ — управление игрока i в момент времени $s \in [t, t + \bar{T}]$, а $x_t = x_t(s)$ — накопленный уровень загрязнения в момент времени s .

Следовательно, функция выигрыша игрока i в подыгре с непрерывным обновлением информации $\Gamma(x, s, t + \bar{T})$ имеет вид

$$K_i^t(x, s, t + \bar{T}; u) = \int_s^{t+\bar{T}} \left((b_i - \frac{1}{2} u_i(t, \tau)) u_i(t, \tau) - d_i x_t(\tau) \right) d\tau, \quad i \in N, \quad (4.223)$$

где $x_t(\tau)$ и $u(t, \tau)$, $\tau \in [s, t + \bar{T}]$, — соответственно траектория и стратегии в игре

$\Gamma(x, s, t + \bar{T})$. Динамика состояния описывается дифференциальным уравнением

$$\dot{x}_t(\tau) = \sum_{i=1}^n u_i(t, \tau), \quad x_t(s) = x_{t,s}. \quad (4.224)$$

Шаг 1: Оптимизация общих выплат гранд коалиции с непрерывным обновлением информации

Рассмотрим игру в кооперативной постановке. Это означает, что все игроки объединяют усилия, чтобы максимизировать свой суммарный выигрыш. Необходимо найти оптимальные стратегии $\tilde{u}^*(t, s) = (\tilde{u}_1^*(t, s), \dots, \tilde{u}_n^*(t, s))$ такие, что $\sum_{i=1}^n K_i^t \rightarrow \max_{u_1, u_2, \dots, u_n}$.

Задача оптимизации заключается в следующем:

$$\sum_{i=1}^n K_i^t(x, t, t + \bar{T}; u) = \sum_{i=1}^n \int_t^{t+\bar{T}} \left((b_i - \frac{1}{2}u_i(t, s))u_i(t, s) - d_i x_t(s) \right) ds \rightarrow \max_{u_1, u_2, \dots, u_n}$$

где $x_t(s)$ удовлетворяет (4.221). (4.225)

Для решения задачи (4.225) воспользуемся классическим принципом максимума Понтрягина. Соответствующий гамильтониан имеет вид

$$H^t(s, x_t(s), u(t, s), \psi^t(s)) = \sum_{i=1}^n (b_i - \frac{1}{2}u_i)u_i - \sum_{i=1}^n d_i x_t + \psi^t(s)(u_1 + u_2 + \dots + u_n). \quad (4.226)$$

Частные производные первого порядка по переменным u_i равны

$$\frac{\partial H^t}{\partial u_i}(s, x_t, u, \psi^t) = b_i - u_i + \psi^t = 0, \quad (4.227)$$

а матрица Гессе $\frac{\partial^2 H^t}{\partial u^2}(s, x_t, u, \psi^t)$ является отрицательно определенной. Следовательно, гамильтониан H^t обладает свойством вогнутости по переменным u_i , а кооперативные стратегии можно записать в виде

$$\tilde{u}_i^*(t, s) = b_i + \psi^t(s).$$

Рассмотрим принцип максимума Понтрягина применительно к системе с сопряженной переменной:

$$\dot{\psi}^t(s) = \sum_{i=1}^n d_i, \quad \psi^t(t + \bar{T}) = 0.$$

Задавая $\sum_{i=1}^n d_i = d_N$, получаем $\psi^t(s) = -d_N(t + \bar{T} - s)$. Таким образом, кооперативные стратегии имеют вид

$$\tilde{u}_i^*(t, s) = b_i - d_N(t + \bar{T} - s). \quad (4.228)$$

Из (4.221) получаем оптимальную кооперативную траекторию

$$x_t^*(s) = x + b_N(s - t) - nd_N(t + \bar{T})(s - t) + nd_N \frac{s^2 - t^2}{2}, \quad (4.229)$$

где $x = x_t(t)$, $d_N = \sum_{i=1}^n d_i$ и $b_N = \sum_{i=1}^n b_i$.

Следуя процедуре (4.139), строим программные оптимальные кооперативные стратегии с непрерывным обновлением информации:

$$u_i^*(t) = \tilde{u}_i^*(t, s)|_{s=t} = b_i - d_N \bar{T}. \quad (4.230)$$

После подстановки $u_i^*(t)$ в дифференциальное уравнение (4.216) приходим к оптимальной кооперативной траектории $x^*(t)$ с непрерывным обновлением информации:

$$x^*(t) = x_0 + b_N(t - t_0) - nd_N \bar{T}(t - t_0). \quad (4.231)$$

Сравнение кооперативных стратегий и соответствующих траекторий, полученных для исходной дифференциальной игры и дифференциальной игры с непрерывным обновлением информации, представлено на рисунках 4.21 и 4.22.

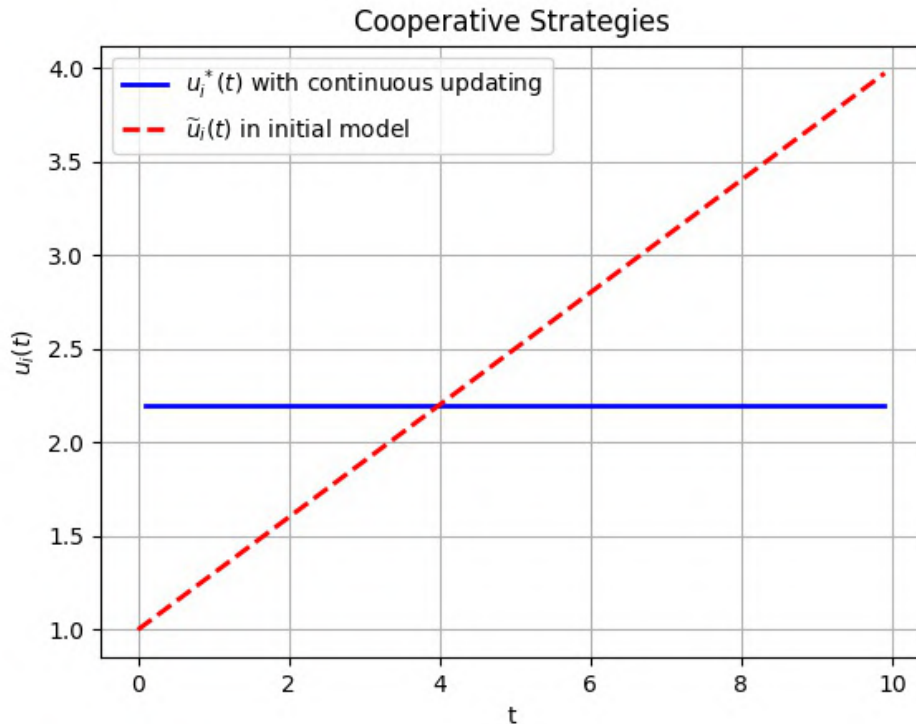


Рис. 4.21: Кооперативные стратегии в исходной игре $\tilde{u}(t)$ (красная линия) и кооперативные стратегии с непрерывным обновлением информации $u^*(t)$ (синяя линия).

Очевидно (см. рис. 4.21), что оптимальное управление с непрерывным обновлением информации более устойчиво, чем оптимальное управление в исходной игре. Кроме того, с момента времени $t = 4$ оптимальное управление в исходной игре сильно растет. Другими словами, игроки должны увеличить выбросы загрязняющих веществ в атмосферу в исходной дифференциальной игре, что является негативным результатом. С точки зрения окружающей среды и устойчивости производственной деятельности, лучше использовать оптимальные стратегии с непрерывным обновлением информации.

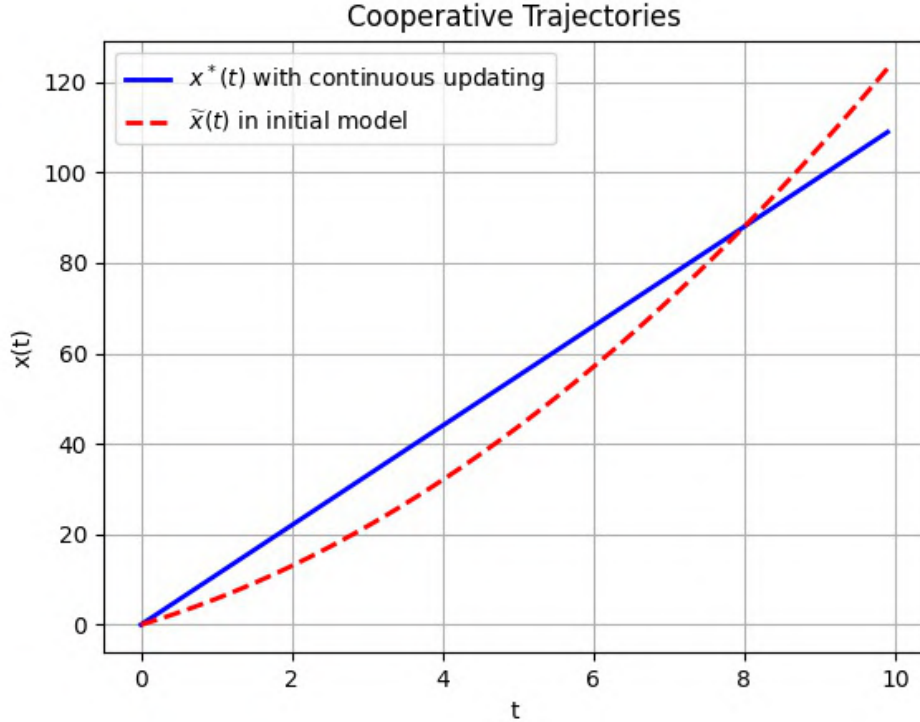


Рис. 4.22: Кооперативная траектория в исходной игре $\tilde{x}(t)$ (красная линия) и кооперативная траектория с непрерывным обновлением информации $x^*(t)$ (синяя линия).

Начиная с момента времени $t = 8$ (см. рис. 4.22), загрязнение атмосферы в игре с непрерывным обновлением информации меньше, чем в исходной игре; более того, рост загрязнения в исходной игре более резкий и, следовательно, негативнее для окружающей среды. При непрерывном обновлении информации это свойство в конечном счете приводит к снижению уровня загрязнения. Поэтому в долгосрочной перспективе использование нового метода для решения проблемы управления загрязнением лучше для окружающей среды.

Для данной подыгры $\Gamma(x^*(t), s, t + \bar{T})$ дифференциальной игры с непрерывным обновлением информации $\Gamma(x^*(t), t, t + \bar{T})$, характеристическая функция для гранд коалиции N задается выражением $\tilde{V}_t(N; x^*(t), s, t + \bar{T})$, которое может быть представлено в виде

$$\tilde{V}_t(N; x^*(t), s, t + \bar{T}) = \sum_{i=1}^n \int_s^{t+\bar{T}} \left((b_i - \frac{1}{2} \tilde{u}_i^*(t, \tau)) \tilde{u}_i^* - d_i x_t^*(\tau) \right) d\tau,$$

где $x_t^*(\tau)$ вычисляется по формуле (4.229) с $x = x^*(t)$, а $\tilde{u}_i^*(t, \tau)$ — по

формуле (4.228). Таким образом, получаем следующую функцию цены для гранд коалиции N :

$$\begin{aligned} \tilde{V}_t(N; x^*(t), s, t + \bar{T}) = & (t + \bar{T} - s) \left[\frac{1}{2} \tilde{b}_N - d_N x^*(t) + \frac{d_N b_N}{2} (t - \bar{T} - s) \right. \\ & \left. - \frac{1}{3} n d_N^2 ((t + \bar{T} - s)^2 - \frac{3}{2} \bar{T}^2) \right], \end{aligned} \quad (4.232)$$

где $d_N = \sum_{i=1}^n d_i$ и $b_N = \sum_{i=1}^n b_i$ найдены выше, а $\tilde{b}_N = \sum_{i=1}^n b_i^2$. Обращаем внимание, что $x^*(t)$ в (4.232) представляет собой кооперативный уровень загрязнения с непрерывным обновлением информации в момент времени t .

Для нашей задачи можно было применить и динамическое программирование на основе решения уравнения Гамильтона–Якоби–Беллмана. Несложно показать, что при выборе функции Беллмана $\tilde{V}_t = A(t, s)x_t(s) + B(t, s)$ получается тот же результат, что и из принципа максимума Понтрягина.

Шаг 2: Вычисление обобщенного равновесия по Нэшу в программных стратегиях с непрерывным обновлением информации

Гамильтониан для каждого игрока $i = 1, 2, \dots, n$ равен

$$H_i^t(\tau, x_t(\tau), u(t, \tau), \psi_i^t(\tau)) = (b_i - \frac{1}{2} u_i) u_i - \sum_{i=1}^n d_i x_t + \psi_i^t(\tau) (u_1 + u_2 + \dots + u_n). \quad (4.233)$$

Частные производные первого порядка по переменным u_i равны

$$\frac{\partial H_i^t}{\partial u_i}(\tau, x_t(\tau), u(t, \tau), \psi_i^t(\tau)) = b_i - u_i + \psi_i^t = 0, \quad (4.234)$$

а матрица Гессе $\frac{\partial^2 H_i^t}{\partial u_i^2}(\tau, x_t, u, \psi_i^t)$ является отрицательно определенной. Следовательно, гамильтониан H_i^t обладает свойством вогнутости по переменным u_i , а оптимальные управления можно записать в виде

$$\tilde{u}_i^{NE}(t, \tau) = b_i - d_i(t + \bar{T} - \tau), \quad i = 1, 2, \dots, n. \quad (4.235)$$

Рассмотрим подыгру, начинающуюся в момент времени $s \in [t, t + \bar{T}]$. Можно легко вывести соответствующую траекторию (для случая равновесия по Нэшу) подыгры $\Gamma(x^*(t), s, t + \bar{T})$ вдоль кооперативной траектории, другими словами,

$x_{t,s} = x_t^*(s)$ определяется выражением

$$\begin{aligned} x_t^{NE}(\tau) &= x_t^*(s) + b_N(\tau - s) + d_N \left(\frac{(t + \bar{T} - \tau)^2}{2} - \frac{(t + \bar{T} - s)^2}{2} \right) \\ &= x^*(t) + b_N(\tau - t) - nd_N(t + \bar{T})(s - t) + nd_N \frac{s^2 - t^2}{2} \\ &\quad + \frac{d_N}{2} ((t + \bar{T} - \tau)^2 - (t + \bar{T} - s)^2). \end{aligned} \quad (4.236)$$

Максимальный выигрыш для каждого игрока $i = 1, 2, \dots, n$ в подыгре, начинающейся с момента времени s из состояния $x^*(t)$, принимает вид:

$$\begin{aligned} \tilde{V}_i(\{i\}; x^*(t), s, t + \bar{T}) &= \int_s^{t+\bar{T}} \left((b_i - \frac{1}{2}\tilde{u}_i^{NE}(t, \tau))\tilde{u}_i^{NE} - d_i x_t^{NE}(\tau) \right) d\tau \\ &= (t + \bar{T} - s) \left[\frac{b_i^2}{2} - d_i x^*(t) + \frac{nd_i d_N}{2} (s - t)(t - s + 2\bar{T}) \right. \\ &\quad \left. - \frac{1}{6} d_i^2 (t + \bar{T} - s)^2 + \frac{1}{3} d_i d_N (t + \bar{T} - s)^2 - \frac{1}{2} d_i b_N (s + \bar{T} - t) \right]. \end{aligned} \quad (4.237)$$

Шаг 3: Вычисление характеристической функции для всех возможных оставшихся коалиций с непрерывным обновлением информации

Также можно вычислить управления и соответствующие функции цены для различных коалиций. Однако их форма будет зависеть от постановки соответствующих задач оптимального управления.

Построим характеристическую функцию на основе подхода δ -характеристической функции. Характеристическая функция коалиции S вычисляется в два этапа: первый уже выполнен (равновесные по Нэшу стратегии для всех игроков найдены на Шаге 2); на втором этапе предполагаем, что оставшиеся игроки $j \in N \setminus S$ реализуют свои оптимальные по Нэшу стратегии $\tilde{u}_j^{NE}(t, \tau)$ при условии, что игроки из коалиции S стремятся максимизировать их кооперативный выигрыш $\sum_{i \in S} K_i$. Рассмотрим случай коалиции S и проведем подробные расчеты. Соответствующий гамильтониан для коалиции S равен

$$H_S^t(\tau, x_t(\tau), u(t, \tau), \psi_S^t(\tau)) = \sum_{i \in S} \left((b_i - \frac{1}{2}u_i)u_i \right) - d_S x_t + \psi_S^t \left(\sum_{i \in S} u_i + \sum_{j \in N/S} \tilde{u}_j^{NE} \right),$$

где $d_S = \sum_{i \in S} d_i$. Заметим, что величины $u_j, j \in N/S$, заменены на \tilde{u}_j^{NE} , вычисленные ранее.

Оптимальные стратегии игроков из коалиции S равны $\tilde{u}_S^*(t, \tau)$, где

$$\tilde{u}_i^*(t, \tau) = b_i + \psi_S^t(\tau), \quad i \in S.$$

Дифференциальное уравнение для $\psi_S^t(\tau)$ имеет вид $\dot{\psi}_S^t(\tau) = d_S$ с дополнительным условием $\psi_S^t(t + \bar{T}) = 0$. В итоге получаем: $\psi_S^t(\tau) = -d_S(t + \bar{T} - \tau)$. При подстановке этого выражения для $\psi_S^t(\tau)$ в формулу \tilde{u}_i^* находим

$$\tilde{u}_i^*(t, \tau) = b_i - d_S(t + \bar{T} - \tau), \quad i \in S \quad (4.238)$$

где $d_S = \sum_{i \in S} d_i$. Очевидно, что игрок вне коалиции S реализует свою оптимальную стратегию \tilde{u}_S^* , в то время как остальные игроки придерживаются своего равновесия по Нэшу.

На следующем шаге проинтегрируем (4.224), начиная с точки $x_t^*(s)$, чтобы получить $x_t^S(\tau)$:

$$\begin{aligned} x_t^S(\tau) = & x_t^*(s) + b_N(\tau - s) + \frac{(t + \bar{T} - \tau)^2}{2}((k - 1)d_S + d_N) \\ & - \frac{(t + \bar{T} - s)^2}{2}((k - 1)d_S + d_N). \end{aligned} \quad (4.239)$$

$x_t^S(\tau)$ — траектория в подыгре $\Gamma(x^*(t), s, t + \bar{T})$ игры $\Gamma(x^*(t), t, t + \bar{T})$, начинающейся в момент времени s из позиции $x_t^*(s)$, когда игроки из коалиции S используют стратегии $\tilde{u}_S^*(t, \tau)$, а игроки из коалиции $N \setminus S$ используют $\tilde{u}^{NE}(t, \tau)$. При рассмотрении случая вдоль кооперативной траектории, можно подставить найденную $x_t^*(s)$ в (4.229), чтобы получить переменную состояния $x_t^S(\tau)$, зависящую от $x = x^*(t)$:

$$\begin{aligned} x_t^S(\tau) = & x^*(t) + b_N(\tau - t) - nd_N(t + \bar{T})(s - t) + nd_N \frac{s^2 - t^2}{2} + \\ & \frac{(t + \bar{T} - \tau)^2}{2}((k - 1)d_S + d_N) - \frac{(t + \bar{T} - s)^2}{2}((k - 1)d_S + d_N). \end{aligned} \quad (4.240)$$

Соответствующее значение характеристической функции $\tilde{V}_t(S; x^*(t), s, t + \bar{T})$

равно

$$\begin{aligned} \tilde{V}_t(S; x^*(t), s, t + \bar{T}) = & (t + \bar{T} - s) \left[\frac{1}{2} \tilde{b}_S - d_S x^*(t) + \frac{d_S d_N n}{2} (t + 2\bar{T} - s)(s - t) + \right. \\ & \left. (t + \bar{T} - s)^2 \left(\frac{k-2}{6} d_S^2 + \frac{d_S d_N}{3} \right) - \frac{d_S b_N}{2} (s + \bar{T} - t) \right]. \end{aligned} \quad (4.241)$$

Согласно Определению 4.3.8, характеристическая функция дифференциальной игры с непрерывным обновлением информации имеет следующий вид:

$$\begin{aligned} V(S; x^*(t), t, T) = & \int_t^T - \frac{d\tilde{V}_{\tau'}(S; x^*(\tau'), s, \tau' + \bar{T})}{ds} \Big|_{s=\tau'} d\tau' \\ = & (T - t) \left[\frac{1}{2} \tilde{b}_S - d_S x_0 + \bar{T}^2 \left(\frac{k-2}{2} d_S^2 + (1-n) d_S d_N \right) \right. \\ & \left. - \frac{d_S (b_N - n d_N \bar{T})}{2} (T + t - 2t_0) \right] \end{aligned} \quad (4.242)$$

Проверим условие супераддитивности (4.213) для построенной характеристической функции $V(S; x^*(t), t, T)$. Оказывается, для любых $S, P \subseteq N$, таких что $S \cap P = \emptyset$, $|S| = k \geq 1$ и $|P| = m \geq 1$, имеет место следующее соотношение:

$$\begin{aligned} V(S \cup P; x^*(t), t, T) - V(S; x^*(t), t, T) - V(P; x^*(t), t, T) \\ = \bar{T}^2 \left[(k + m - 2) d_S d_P + \frac{k}{2} d_P^2 + \frac{m}{2} d_S^2 \right] \geq 0. \end{aligned} \quad (4.243)$$

Таким образом, δ -характеристическая функция $V(S; x^*(t), t, T)$ является супераддитивной функцией без каких-либо дополнительных условий, предъявляемых к параметрам модели.

На следующем рисунке сравниваются характеристические функции для гранд коалиции N в исходной игре и дифференциальной игре с непрерывным обновлением информации.

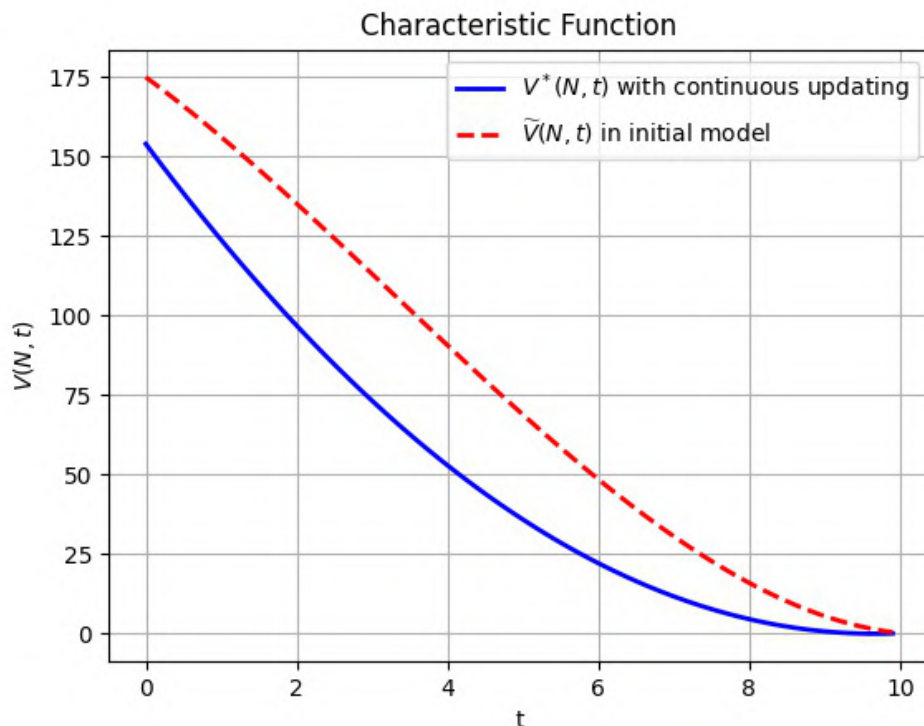


Рис. 4.23: Значение характеристической функции коалиции N в исходной игре (красная линия) и значение характеристической функции коалиции N с непрерывным обновлением информации $V^*(N, t, T)$ (синяя линия).

Почему значение характеристической функции в исходной модели больше, чем в модели с непрерывным обновлением информации? Причина состоит в том, что сложность информации в постановке с непрерывным обновлением информации может снизить эффективность коалиции. Следует отметить, что случай с непрерывным обновлением информации более реалистичен. Можно сделать вывод, что выигрыш коалиции уменьшается, поскольку со временем загрязнение накапливается в воздухе. Выигрыш игрока зависит от уровня загрязнения, и выигрыш уменьшается по мере увеличения загрязнения. Также обращаем внимание, что эффективность коалиции снижается в исходной игре более быстрыми темпами, чем в игре с непрерывным обновлением информации.

Шаг 4: Вычисление вектора Шепли на основе характеристической функции с непрерывным обновлением информации

Любой из известных принципов оптимальности может быть применен для поиска кооперативного решения. Прежде всего, величина $\sum d_j d_i^3$ отражает

³Отметим, что $d_k d_j = d_j d_k$.

распределение затрат между игроками. Теперь рассмотрим кооперативное решение дифференциальной игры с непрерывным обновлением информации. Следуя процедуре (4.215), строим вектор Шепли для любого $i \in N$ с непрерывным обновлением информации, используя характеристическую функцию с непрерывным обновлением информации подыгры. Получаем:

$$\begin{aligned}
 sh_i(x^*(t), t, T) &= \sum_{K \subseteq N, i \in K} \frac{(k-1)!(n-k)!}{n!} [V(K; x^*(t), t, T) - V(K \setminus i; x^*(t), t, T)] \\
 &= (T-t) \left[-d_i x_0 + \frac{1}{2} b_i^2 + \frac{nd_i d_N \bar{T} - d_i b_N}{2} (T+t-2t_0) \right. \\
 &\quad \left. + \bar{T}^2 \left(\frac{1-2n}{3} d_i d_N - \frac{4+n}{12} d_i^2 + \frac{1}{3} \left(\sum_{\substack{j, l \neq i \\ j \neq l \in N}} d_j d_l \right) + \frac{1}{4} \tilde{d}_N \right) \right].
 \end{aligned} \tag{4.244}$$

На рисунке 4.24 представлены векторы Шепли для подыгр с непрерывным обновлением информации и исходной игры вдоль оптимальной кооперативной траектории $x^*(t)$.

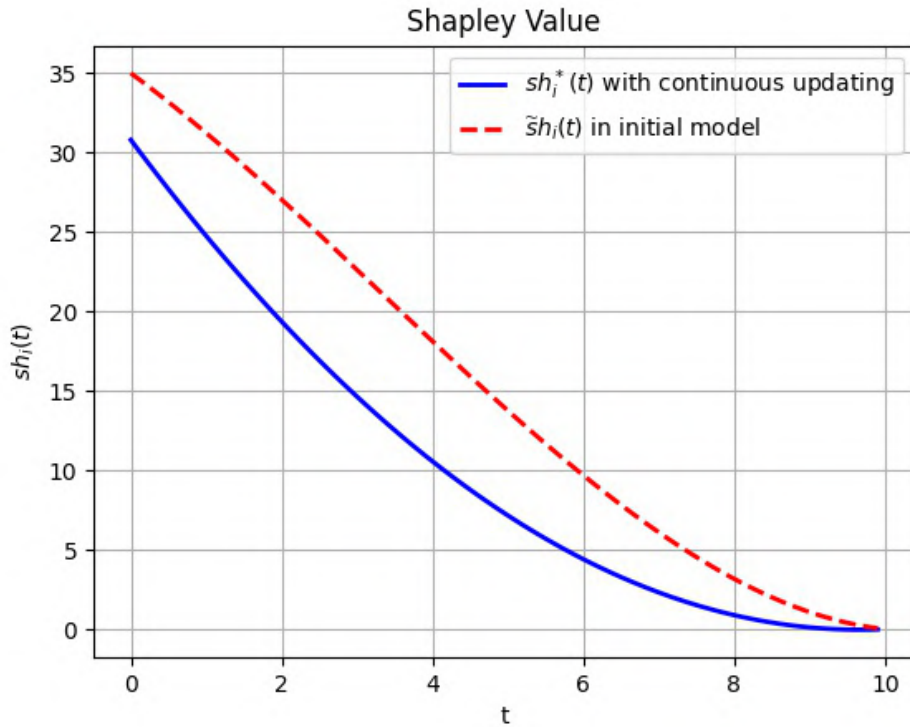


Рис. 4.24: Вектор Шепли игрока i в исходной игре (красная линия) и вектор Шепли с непрерывным обновлением информации $sh_i^*(t)$ (синяя линия).

Согласно рисунку 4.24, при рассмотрении проблемы в более реалистичном случае (игра с непрерывным обновлением информации) игрок может получить меньшую выплату от коалиции, чем в исходной игре. Дело в том, что на раннем этапе игры с непрерывным обновлением информации выбросы загрязняющих веществ в атмосферу выше, чем в исходной игре, а на последнем этапе — ниже, чем в исходной игре. Таким образом, начиная с момента времени $t = 0$, игроки получают больше в исходной игровой модели, но за тот же период все они в итоге получают 0. Таким образом, по мере усиления загрязнения производственные выгоды стран постепенно сокращаются.

4.3.4 Линейно-квадратичный случай кооперативной дифференциальной игры с непрерывным обновлением информации

Линейно-квадратичный случай этого класса игр особенно важен для практических задач, возникающих в технических системах при взаимодействии человека с машиной. Интересным результатом настоящего раздела является то, что программные стратегии используются для построения оптимальных стратегий, но впоследствии все же получаются позиционные стратегии. С помощью этих стратегий определены понятия вектора Шепли и равновесия по Нэшу в качестве принципов оптимальности соответственно для кооперативной и некооперативной постановок, а также представлены оптимальные стратегии для линейно-квадратичного случая. Дополнительно для построения характеристической функции с непрерывным обновлением информации представлена форма равновесия по Нэшу с непрерывным обновлением информации в программных стратегиях.

В этом разделе будем использовать только программные стратегии для построения кооперативных стратегий с непрерывным обновлением информации $u_{ol}^*(t, x)$ и программные равновесные по Нэшу стратегии $u_{ol}^{NE}(t, x)$ для построения характеристической функции с непрерывным обновлением информации. Поэтому введем упрощенные обозначения $u^*(t, x)$ и $u^{NE}(t, x)$.

4.3.4.1 Кооперативные стратегии с непрерывным обновлением информации

Чтобы построить кооперативную игру с непрерывным обновлением информации, необходимо определить кооперативные стратегии и соответствующую траекторию с непрерывным обновлением информации. Ниже представлены достаточные условия существования кооперативных стратегий для линейно-квадратичной дифференциальной игры с непрерывным обновлением информации.

Теорема 4.3.7 *Рассмотрим линейно-квадратичную дифференциальную игру N лиц с матрицами $Q_i \geq 0$, $R_{ij} \geq 0$ ($i, j \in N, i \neq j$). Предположим, что существует набор решений $\{Z^t, i \in N, t \geq t_0\}$ для матричных дифференциальных уравнений Риккати*

$$\begin{aligned} \dot{Z}^t(\tau) &= -\bar{T}A'Z^t(\tau) - \bar{T}Z^t(\tau)A + Z^t(\tau)SZ^t(\tau) - Q, \\ Z^t(1) &= 0, \end{aligned} \quad (4.245)$$

где $S = \bar{T}^2BR^{-1}B'$, $Q = \sum_{i \in N} Q_i$, $B = [B_1, \dots, B_n]$ и $R = \{R_{ij}\}_{i,j=1}^n$ — блочные матрицы. Тогда данная дифференциальная игра с непрерывным обновлением информации имеет кооперативное решение с непрерывным обновлением информации, задаваемое формулой

$$u^*(t, x) = -R^{-1}B'Z^t(0)\bar{T}x.$$

Доказательство Для доказательства теоремы введем замену переменных

$$\begin{aligned} s &= t + \bar{T}\tau, \\ y^t(\tau) &= x^t(t + \bar{T}\tau), \\ v_i^t(\tau, y) &= u_i(t + \bar{T}\tau, x), \quad i \in N. \end{aligned} \quad (4.246)$$

Подставляя (5.12) в уравнения движения (4.49) и функцию выигрыша (4.50), получаем:

$$\dot{y}^t(\tau) = \bar{T}Ay^t(\tau) + \sum_{i=1}^N \bar{T}B_i v_i^t(\tau, y) \quad (4.247)$$

и

$$K^t(y^t, \tau; v^t) = \sum_{i \in N} \int_0^1 \left((y^t(s))' Q_i y^t(s) + \sum_{j=1}^N (v_j^t(s, y))' R_{ij} v_j^t(s, y) \right) ds, \quad i \in N. \quad (4.248)$$

В силу Теоремы 5.1 из [224] и существования решения для системы дифференциальных уравнений (4.245), подыгра $\Gamma(x, t, \bar{T})$ имеет кооперативное решение

$$v^{t,*}(\tau, y_0) = -R^{-1} B' Z^t(\tau) \bar{T} \Phi^t(\tau) y_0,$$

где

$$\begin{aligned} \frac{d\Phi^t}{d\tau} &= (A - S Z^t(\tau)) \Phi^t(\tau), \\ \Phi^t(0) &= E. \end{aligned}$$

Возвращаясь к исходным переменным, получаем следующие стратегии:

$$u^t(s, x) = -R^{-1} B' Z^t \left(\frac{s-t}{\bar{T}} \right) \bar{T} \Phi^t \left(\frac{s-t}{\bar{T}} \right) x.$$

Тогда обобщенное кооперативное решение для игры с непрерывным обновлением информации имеет вид

$$\tilde{u}^*(t, s, x) = -R^{-1} B' Z^t \left(\frac{s-t}{\bar{T}} \right) \bar{T} \Phi^t \left(\frac{s-t}{\bar{T}} \right) x. \quad (4.249)$$

Наконец, применим процедуру (4.11) для определения равновесия по Нэшу с непрерывным обновлением информации с помощью обобщенного равновесия по Нэшу с непрерывным обновлением информации (4.249), $s = t$:

$$u^*(t, x) = -R^{-1} B' Z^t(0) \bar{T} x, \quad t \in [t_0, +\infty), \quad i \in N. \quad (4.250)$$

Теорема доказана. □

4.3.4.2 Равновесные по Нэшу стратегии с непрерывным обновлением информации для построения характеристической функции

Для построения характеристической функции с непрерывным обновлением информации в следующем разделе необходимо построить программные стратегии с непрерывным обновлением информации. Соответствующие условия оптимальности представлены ниже.

Теорема 4.3.8 *Рассмотрим линейно-квадратичную дифференциальную игру N лиц с матрицами $Q_i \geq 0$, $R_{ij} \geq 0$ ($i, j \in N, i \neq j$). Предположим, что существует набор решений $\{M_i^t, i \in N, t \geq t_0\}$ для матричных дифференциальных уравнений Риккати*

$$\begin{aligned} \frac{dM_i^t(\tau)}{d\tau} + \bar{T}M_i^t(\tau)A + \bar{T}A'M_i^t(\tau) + Q_i - \bar{T}^2 M_i^t(\tau) \sum_{j \in N} B_j R_{jj}^{-1} B_j' M_j^t(\tau) &= 0, \\ M_i^t(1) = 0, \quad i \in N. \end{aligned} \tag{4.251}$$

Тогда данная дифференциальная игра с непрерывным обновлением информации имеет равновесие по Нэшу в программных стратегиях с непрерывным обновлением информации, задаваемое формулой

$$u_i^{NE}(t, x) = -R_{ii}^{-1} B_i' M_i^t(0) \bar{T} x(t), \quad i \in N.$$

Доказательство Для доказательства теоремы введем замену переменных

$$\begin{aligned} s &= t + \bar{T}\tau, \\ y^t(\tau) &= x^t(t + \bar{T}\tau), \\ v_i^t(\tau, y) &= u_i(t + \bar{T}\tau, x), \quad i \in N. \end{aligned} \tag{4.252}$$

Подставляя (4.252) в уравнения движения (4.49) и функцию выигрыша (4.50), получаем:

$$\dot{y}^t(\tau) = \bar{T}A y^t(\tau) + \sum_{i=1}^N \bar{T}B_i v_i^t(\tau, y) \tag{4.253}$$

и

$$K_i^t(y^t, \tau; v^t) = \int_0^1 (y^t(s))' Q_i y^t(s) + \sum_{j=1}^N (v_j^t(s, y))' R_{ij} v_j^t(s, y) ds, \quad i \in N. \quad (4.254)$$

В силу Теоремы 6.12 из [80] и существования решения для системы дифференциальных уравнений (4.251), подыгра $\Gamma(x, t, \bar{T})$ имеет равновесие по Нэшу в программных стратегиях вида

$$v_i^{t, NE}(\tau, y_0) = -R_{ii}^{-1} B_i' M_i^t(\tau) \bar{T} \Phi^t(\tau) y_0,$$

где

$$\frac{d\Phi^t}{d\tau} = \left(A - \sum_{i \in N} B_i R_{ii}^{-1} B_i' \right) \Phi^t(\tau),$$

$$\Phi^t(0) = E.$$

Возвращаясь к исходным переменным, получаем следующие стратегии:

$$u_i^t(s, x) = -R_{ii}^{-1} B_i' M_i^t \left(\frac{s-t}{\bar{T}} \right) \bar{T} \Phi^t \left(\frac{s-t}{\bar{T}} \right) x.$$

Тогда обобщенное равновесие по Нэшу в программных стратегиях с непрерывным обновлением информации в игре с непрерывным обновлением информации имеет вид

$$\tilde{u}_i^{NE}(t, s, x) = -R_{ii}^{-1} B_i' M_i^t \left(\frac{s-t}{\bar{T}} \right) \bar{T} \Phi^t \left(\frac{s-t}{\bar{T}} \right) x. \quad (4.255)$$

Наконец, применим процедуру (4.11) для определения равновесия по Нэшу с непрерывным обновлением информации с помощью обобщенного равновесия по Нэшу с непрерывным обновлением информации (4.255), $s = t$:

$$u_i^{NE}(t, x) = -R_{ii}^{-1} B_i' M_i^t(0) \bar{T} x, \quad t \in [t_0, +\infty), \quad i \in N. \quad (4.256)$$

Теорема доказана. □

Замечание 3 *Обращаем внимание, что решение в программных стратегиях с непрерывным обновлением информации имеет форму обратной связи. Дру-*

гими словами, равновесие по Нэшу в программных стратегиях с непрерывным обновлением информации явно зависит от текущего состояния. Это происходит из-за способа построения решения на основе обобщенного равновесия по Нэшу с непрерывным обновлением информации.

4.3.4.3 Характеристическая функция для подыгры на временном интервале $[t, t + \bar{T}]$

Рассмотрим коалицию S в дифференциальной игре n лиц $\Gamma(x, t, \bar{T})$ (4.49), (4.50). Характеристическая функция определяется как суммарный выигрыш коалиции S в равновесии по Нэшу $u^{NE} = (u_1^{NE}, \dots, u_{n_S}^{NE})$ в игре $\Gamma^S(x, t, \bar{T})$ со следующим набором игроков: коалиция S , действующая как один игрок, и игроки из множества $N \setminus S$, т.е. в игре $n_S = |N \setminus S| + 1$ игроков.

Опишем процесс построения вспомогательной игры $\Gamma^S(x, t, \bar{T})$. Пусть для удобства первым игроком этой игры является игрок, связанный с коалицией S , а другие игроки — некоторым образом пронумерованные игроки из множества $N \setminus S$. Переименуем матрицы для игроков из $N \setminus S$ следующим образом: $A^S = A$, $B_i^S = B_{k_i}$, $Q_i^S = Q_{k_i}$, $R_{i,j} = R_{k_i, k_j}$, где $i, j = \overline{2, n_S}$ и i — новый индекс игрока k_i из игры $\Gamma(x, t, \bar{T})$ во вспомогательной игре $\Gamma^S(x, t, \bar{T})$. Некоторые матрицы для коалиционного игрока имеют блочную структуру: $B_1^S = [B_{m_1} \dots B_{m_c}]$, $R_{1,1}^S = \text{diag}(R_{m_1, m_1}, \dots, R_{m_c, m_c})$, $R_{i,1}^S = \text{diag}(R_{k_i, m_1}, \dots, R_{k_i, m_c})$; другие представляют собой сумму соответствующих матриц из $\Gamma(x, t, \bar{T})$: $Q_1^S = \sum_{m \in S} Q_m$, $R_{1,i}^S = \sum_{m \in S} R_{m, k_i}$, где $m_1, \dots, m_s \in S$ и $i = \overline{2, n_S}$.

Таким образом, уравнения движения в игре $\Gamma^S(x, t, \bar{T})$ имеют вид

$$\begin{aligned} \dot{x}^t(s) &= A^S x^t(s) + B_1^S u_1^t(s, x^t) + \dots + B_n^S u_n^t(s, x^t), \\ x^t(t) &= x. \end{aligned}$$

Определим функцию выигрыша игрока $i \in N^S$ в игре $\Gamma^S(x, t, \bar{T})$ следующим

образом:

$$K_i^{S,t}(x^t, t, \bar{T}; u^t) = \int_t^{t+\bar{T}} \left((x^t(s))' Q_i^S x^t(s) + \sum_{j=1}^n (u_j^t(s, x^t))' R_{ij}^S u_j^t(s, x^t) \right) ds,$$

где $x^t(s)$ и $u^t(s, x)$ обозначают траекторию и стратегии в игре $\Gamma^S(x, t, \bar{T})$.

Лемма 4.3.1 *Рассмотрим линейно-квадратичную дифференциальную игру N лиц $\Gamma(x, t, \bar{T})$ с матрицами $Q_i^S \geq 0$, $R_{ij}^S \geq 0$ ($i, j \in N, i \neq j$). Предположим, что существует набор решений $\{M_i^S, i \in N_S, t \geq t_0\}$ для матричных дифференциальных уравнений Риккати*

$$\begin{aligned} \frac{dM_i^S(\tau)}{d\tau} + \bar{T} M_i^S(\tau) A^S + \bar{T} (A^S)' M_i^S(\tau) + Q_i^S - \\ - \bar{T}^2 M_i^S(\tau) \sum_{j \in N_S} B_j^S (R_{jj}^S)^{-1} (B_j^S)' M_j^S(\tau) = 0, \end{aligned} \quad (4.257)$$

$$M_i^S(1) = 0, \quad i \in N_S.$$

Тогда характеристическая функция для игры $\Gamma(x, t, \bar{T})$ имеет вид

$$\begin{aligned} V^t(S, x, \xi, t + \bar{T}) = \int_{\xi}^{t+\bar{T}} \left((x^*(s, t, x))' Q_1^S x^*(s, t, x) + \right. \\ \left. + \sum_{j=1}^{n_S} (u_j^t(s, x))' R_{1j}^S u_j^t(s, x) \right) ds, \end{aligned} \quad (4.258)$$

где

$$u_i^t(s, x) = - (R_{ii}^S)^{-1} (B_i^S)' M_i^S \left(\frac{s-t}{\bar{T}} \right) \bar{T} \Phi^S \left(\frac{s-t}{\bar{T}} \right) x, \quad (4.259)$$

$\Phi^S(\tau)$ — решение системы

$$\begin{aligned} \frac{d\Phi^S}{d\tau} = \left(A^S - \sum_{i \in N_S} B_i^S (R_{ii}^S)^{-1} (B_i^S)' \right) \Phi^S(\tau), \\ \Phi^S(0) = E, \end{aligned}$$

а $x^*(s, t, x)$ — решение системы

$$\begin{aligned}\dot{x}^t(s) &= A^S x^t(s) + B_1^S u_1^t(s, x) + \dots + B_n^S u_n^t(s, x), \\ x^t(t) &= x.\end{aligned}\tag{4.260}$$

Доказательство Мы определили характеристическую функцию как суммарный выигрыш коалиции S в равновесии по Нэшу в игре $\Gamma^S(x, t, \bar{T})$.

Подобно доказательству Теоремы 4.3.8, используем замену переменных (4.252), Теорему 6.12 из [80] и существование решения для системы дифференциальных уравнений (4.257). В результате получаем, что программные равновесные по Нэшу стратегии в подыгре $\Gamma^S(x, t, \bar{T})$ имеют вид (4.259).

По построению вспомогательной игры $\Gamma^S(x, t, \bar{T})$, суммарный выигрыш коалиции S можно вычислить как $K_1^{S,t}(x^t, t, \bar{T}, u^t)$. Тогда функция (4.258) является характеристической, а динамика системы описывается дифференциальным уравнением (4.260). \square

4.3.4.4 Характеристическая функция для игр с непрерывным обновлением информации

Предположим, что функция $\tilde{V}^t(S; \tilde{x}_t^*(s), s, t + \bar{T})$, $S \subseteq N$, непрерывно дифференцируема по $s \in [t, \bar{T}]$ и интегрируема по $t \in [t_0, +\infty)$. Определим характеристическую функцию в игре с непрерывным обновлением информации $V(S; x^*(t), t)$ в виде (4.156).

Теорема 4.3.9 Рассмотрим коалицию S в линейно-квадратичной дифференциальной игре N лиц с матрицами $Q_i^S \geq 0$, $R_{ij}^S \geq 0$ ($i, j \in N, i \neq j$). Предположим, что существует набор решений $\{M_i^S, i \in N_S, t \geq t_0\}$ для матричных дифференциальных уравнений Риккати (4.257). Тогда характеристическая функция для игры с непрерывным обновлением имеет вид

$$V(S, x, t, T) = \int_t^T (x^*(s, t, x))' \left(Q_1^S - \bar{T}^2 \sum_{j=1}^{n_S} P_j' R_{1j}^S P_j \right) x^*(s, t, x) ds, \tag{4.261}$$

где

$$P_j = (R_{jj}^S)^{-1} (B_j^S)' M_j^S(0), \quad (4.262)$$

а $x^*(s, t, x)$ — решение системы

$$\begin{aligned} \dot{x}^t(s) &= \left(A^S - \sum_{i=1}^{n_S} B_i^S P_i \right) x^t(s), \\ x^t(t) &= x. \end{aligned} \quad (4.263)$$

Доказательство Согласно определению, характеристическая функция для игры с непрерывным обновлением информации имеет общий вид (4.156). Лемма 4.3.1 дает характеристическую функцию (4.258) для подыгры $\Gamma(x, t, \bar{T})$. Подставляя (4.258) в выражение (4.156), получаем:

$$\begin{aligned} V(S, x, t, T) &= \int_t^T -\frac{d}{ds} V^\tau(S; \tilde{x}_\tau^*(s), s, \tau + \bar{T})|_{s=\tau} d\tau = \\ &= \int_t^T \left((x^*(s, t, x))' Q_1^S x^*(s, t, x) + \right. \\ &\quad \left. + \sum_{j=1}^{n_S} (u_j^s(s, x^*(s, t, x)))' R_{1j}^S u_j^s(s, x^*(s, t, x)) \right) ds, \end{aligned} \quad (4.264)$$

С учетом $\Phi^S(0) = E$, имеем:

$$u_i^s(s, x) = - (R_{ii}^S)^{-1} (B_i^S)' M_i^S(0) \bar{T} x. \quad (4.265)$$

Подставляя (4.265) в выражение (4.264) и используя формулу (4.262), получаем:

$$\begin{aligned} V(S, x, t, T) &= \int_t^T \left((x^*(s, t, x))' Q_1^S x^*(s, t, x) - \right. \\ &\quad \left. - \bar{T}^2 \sum_{j=1}^{n_S} (x^*(s, t, x))' P_j' R_{1j}^S P_j x^*(s, t, x) \right) ds. \end{aligned} \quad (4.266)$$

По аналогии с (4.261), запишем выражение для (4.266). В силу (4.265) и (4.260), динамику системы можно представить в виде (4.263). \square

Ниже любое кооперативное решение с непрерывным обновлением информации будет определено с помощью характеристической функции с непрерывным обновлением информации.

4.3.4.5 Общественный фонд знаний в дифференциальной игре с непрерывным обновлением информации

4.3.4.6 Модель исходной игры

Рассмотрим модель, в которой два индивидуума инвестируют в общественный фонд знаний (см. также статью Докнера и соавторов [50]). Пусть $x(t)$ — запас знаний в момент времени t , а $u_i(t)$ — инвестиция игрока i в общественное знание в момент времени t . Предположим, что запас знаний накапливается согласно дифференциальному уравнению

$$\dot{x}(t) = -\beta x(t) + u_1(t, x_0) + u_2(t, x_0), \quad x(0) = x_0, \quad (4.267)$$

где β — скорость устаревания знаний. Предположим, что каждый игрок получает квадратичную полезность от потребления запасов знаний, а инвестиционные затраты являются квадратичной возрастающей функцией инвестиционных усилий. То есть функция затрат обоих игроков имеет вид

$$K_i(x_0, t_0, T; u) = \int_0^T (-q_i x^2(t) + r_i u_i^2(t, x_0)) dt, \quad i = 1, 2.$$

4.3.4.7 Модель игры с непрерывным обновлением информации

Теперь рассмотрим случай с непрерывным обновлением информации. Предположим, что в каждый момент времени $t \in [t_0, +\infty)$ два индивидуума используют информацию об уравнениях движения и функциях выигрыша на временном интервале $[t, t + \bar{T}]$. По мере продвижения текущего времени t , временной интервал, на котором доступна информация, также смещается. Уравнения движения для модели игры с непрерывным обновлением информации имеют

вид

$$\dot{x}^t(s) = -\beta x^t(s) + u_1^t(s, x) + u_2^t(s, x), \quad x^t(t) = x, \quad t \in [t_0, +\infty).$$

4.3.4.7.1 Некооперативный случай

Определим функцию выигрыша игрока $i \in N$ в игре с непрерывным обновлением информации следующим образом:

$$K_i^t(x^t, t, \bar{T}; u^t) = \int_t^{t+\bar{T}} \left(- (x^t(s))^2 q_i + (u_i^t(s, x))^2 r_i \right) ds, \quad i = 1, 2.$$

В качестве примера, рассмотрим симметричный случай $r_1 = r_2 = r$, $q_1 = q_2 = q$. Согласно Теореме 4.3.8, определяющей явный вид равновесия по Нэшу в программных стратегиях с непрерывным обновлением информации, на первом шаге необходимо решить дифференциальное уравнение

$$\begin{cases} \dot{k}(\tau) = 2\beta\bar{T}k(\tau) + \frac{2\bar{T}k^2(\tau)}{r} + q, \\ k(1) = 0. \end{cases} \quad (4.268)$$

Решением уравнения (4.268) является функция

$$k(\tau) = \frac{r(\beta - v)}{2\bar{T}} \left(\frac{2v}{v - \beta + (v + \beta)e^{2v\bar{T}(1-\tau)}} - 1 \right), \quad (4.269)$$

где $v = \sqrt{\beta^2 - \frac{2q}{r}}$. Согласно Теореме 4.3.8, равновесие по Нэшу в программных стратегиях с непрерывным обновлением информации имеет вид

$$\tilde{u}_i^{NE}(t, x) = -\frac{k(0)x\bar{T}}{r}. \quad (4.270)$$

При подстановке выражения (4.269) в формулу (4.270) получаем:

$$\tilde{u}_i^{NE}(t, x) = \frac{\beta - v}{2} \left(\frac{2v}{v - \beta + (v + \beta)e^{2v\bar{T}(1-\tau)}} - 1 \right) x. \quad (4.271)$$

Наконец, подставляя выражение (4.271) в уравнение (4.267), устанавливаем, что

функция $\tilde{x}^{NE}(t)$ — решение дифференциального уравнения

$$\dot{\tilde{x}}^{NE}(t) = -\beta\tilde{x}^{NE}(t) + \tilde{u}_1^{NE}(t, x) + \tilde{u}_2^{NE}(t, x), \quad \tilde{x}^{NE}(0) = x_0. \quad (4.272)$$

4.3.4.7.2 Кооперативный случай

Определим функцию выигрыша игрока $i \in N$ в игре с непрерывным обновлением информации следующим образом:

$$K^t(x^t, t, \bar{T}; u^t) = \sum_{i=1}^2 \int_t^{t+\bar{T}} \left(- (x^t(s))^2 q_i + (u_i^t(s, x))^2 r_i \right) ds.$$

Вновь рассмотрим симметричный случай $r_1 = r_2 = r$, $q_1 = q_2 = q$. Согласно Теореме 2, определяющей явный вид кооперативных стратегий с непрерывным обновлением информации, на первом шаге необходимо решить дифференциальное уравнение

$$\begin{cases} \dot{k}(\tau) = 2\beta\bar{T}k(\tau) + \frac{2\bar{T}k^2(\tau)}{r} + 2q, \\ k(1) = 0. \end{cases} \quad (4.273)$$

Решением уравнения (4.273) является функция

$$k(\tau) = \frac{r(\beta - v_1)}{2\bar{T}} \left(\frac{2v_1}{v_1 - \beta + (v_1 + \beta)e^{2v_1\bar{T}(1-\tau)}} - 1 \right), \quad (4.274)$$

где $v_1 = \sqrt{\beta^2 - \frac{4q}{r}}$. Согласно Теореме 4.3.7, кооперативные стратегии с непрерывным обновлением информации имеют вид

$$\tilde{u}_i^*(t, x) = -\frac{k(0)x\bar{T}}{r}. \quad (4.275)$$

При подстановке выражения (4.274) в формулу (4.275) получаем:

$$\tilde{u}_i^*(t, x) = \frac{\beta - v_1}{2} \left(\frac{2v_1}{v_1 - \beta + (v_1 + \beta)e^{2v_1\bar{T}(1-\tau)}} - 1 \right) x. \quad (4.276)$$

Далее, подставляя выражение (4.276) в уравнение (4.267), устанавливаем, что функция $\tilde{x}^*(t)$ — решение дифференциального уравнения

$$\dot{\tilde{x}}^*(t) = -\beta\tilde{x}^*(t) + \tilde{u}_1^*(t, x) + \tilde{u}_2^*(t, x), \quad \tilde{x}^*(0) = x_0. \quad (4.277)$$

С учетом (4.276) и (4.277), имеем характеристическую функцию вида

$$\tilde{V}(S, x, t, T) = \sum_{i \in S} \int_t^T \left(-(\tilde{x}^*(s))^2 q_i + (\tilde{u}_i^*(s, x))^2 r_i \right) ds. \quad (4.278)$$

Наконец, при подстановке (4.278) в формулу (4.164) получаем вектор Шепли с непрерывным обновлением информации $\tilde{S}h_i(\tilde{x}^*(t), t, T)$.

4.3.4.8 Модель игры на бесконечном временном интервале

Рассмотрим классический подход для равновесия по Нэшу в игре на бесконечном временном интервале $[0, +\infty)$. Уравнения движения имеют вид

$$\dot{x}(t) = -\beta x(t) + u_1(t, x) + u_2(t, x), \quad x(0) = x_0. \quad (4.279)$$

Определим функцию выигрыша игрока $i \in N$ следующим образом:

$$K_i(x_0; u) = \lim_{T \rightarrow \infty} \int_0^T \left(-q_i x^2(t) + r_i u_i^2(t, x) \right) dt, \quad i = 1, 2.$$

4.3.4.8.1 Некооперативный случай

Согласно [224], в симметричном случае ($r_1 = r_2 = r$, $q_1 = q_2 = q$) равновесные по Нэшу программные стратегии имеют вид

$$u_i^*(t, x_0) = -\frac{kx_0}{r} e^{-(\beta + \frac{2k}{r})t}, \quad (4.280)$$

где k — решение квадратного уравнения

$$\frac{2k^2}{r} + 2\beta k + q = 0.$$

Подставляя выражение (4.280) в уравнения (4.279), устанавливаем, что $x^{NE}(t)$ — решение дифференциального уравнения

$$\dot{x}^{NE}(t) = -\beta x^{NE}(t) - \frac{2kx_0}{r} e^{-(\beta + \frac{2k}{r})t}, \quad x^{NE}(0) = x_0. \quad (4.281)$$

4.3.4.8.2 Кооперативный случай

В этом случае определим функцию выигрыша следующим образом:

$$K(x_0; u) = \sum_{i=1}^2 \lim_{T \rightarrow \infty} \int_0^T (-q_i x^2(t) + r_i u_i^2(t, x)) dt.$$

Согласно [224], в симметричном случае ($r_1 = r_2 = r$, $q_1 = q_2 = q$) равновесные по Нэшу программные стратегии имеют вид

$$u_i^*(t, x_0) = -\frac{kx_0}{r} e^{-(\beta + \frac{2k}{r})t}, \quad (4.282)$$

где k — решение квадратного уравнения

$$\frac{2k^2}{r} + 2\beta k + 2q = 0.$$

Подставляя выражение (4.282) в уравнения (4.279), устанавливаем, что $x^*(t)$ — решение дифференциального уравнения

$$\dot{x}^*(t) = -\beta x^*(t) - \frac{2kx_0}{r} e^{-(\beta + \frac{2k}{r})t}, \quad x^*(0) = x_0. \quad (4.283)$$

С учетом (4.282) и (4.283), имеем характеристическую функцию вида

$$V(S, x, t, T) = \sum_{i \in S} \int_t^T \left(-(x^*(s))^2 q_i + (u_i^*(s, x))^2 r_i \right) ds. \quad (4.284)$$

Наконец, при подстановке (4.284) в формулу (4.164) получаем вектор Шепли на бесконечном временном интервале $Sh_i(x^*(t), t, T)$.

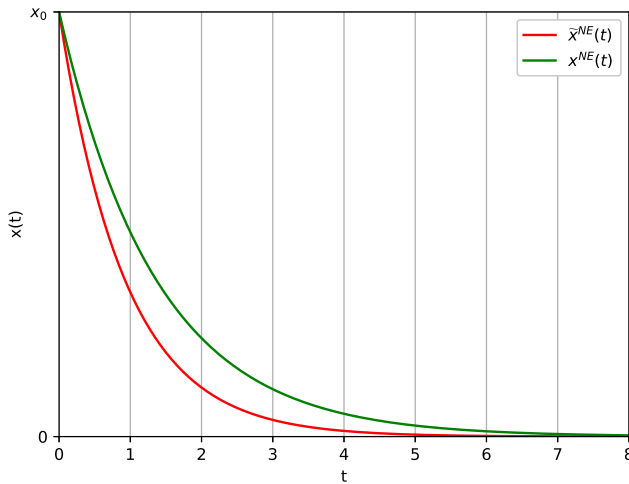


Рис. 4.25: $\tilde{x}^{NE}(t)$ (4.356) — красная линия снизу, $x^{NE}(t)$ (4.328) — зеленая линия сверху.

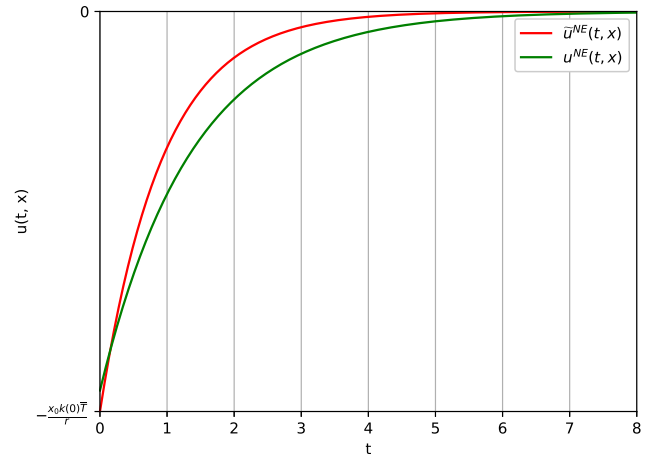


Рис. 4.26: $\tilde{u}^{NE}(t)$ (4.271) — красная линия сверху, $u^{NE}(t)$ (4.280) — зеленая линия снизу.

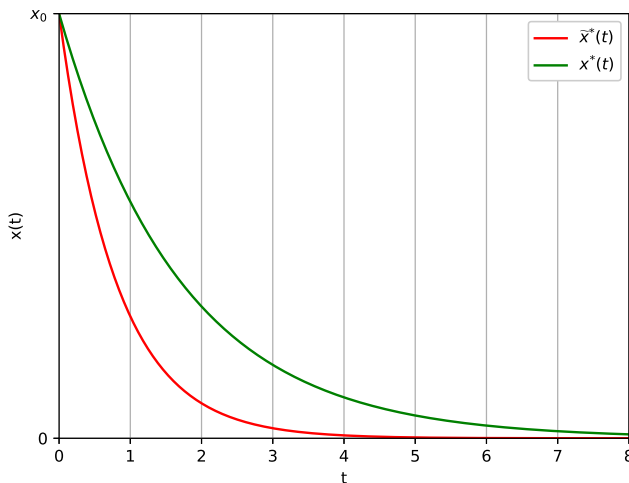


Рис. 4.27: $\tilde{x}^*(t)$ (4.277) — красная линия снизу, $x^*(t)$ (4.283) — зеленая линия сверху.

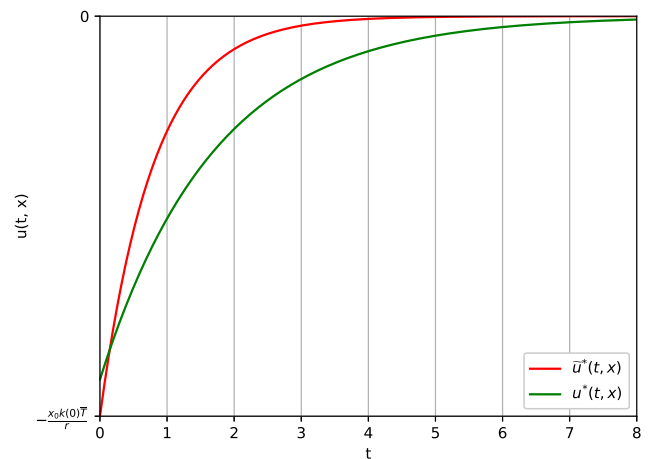


Рис. 4.28: $\tilde{u}^*(t)$ (4.276) — красная линия сверху, $u^*(t)$ (4.282) — зеленая линия снизу.

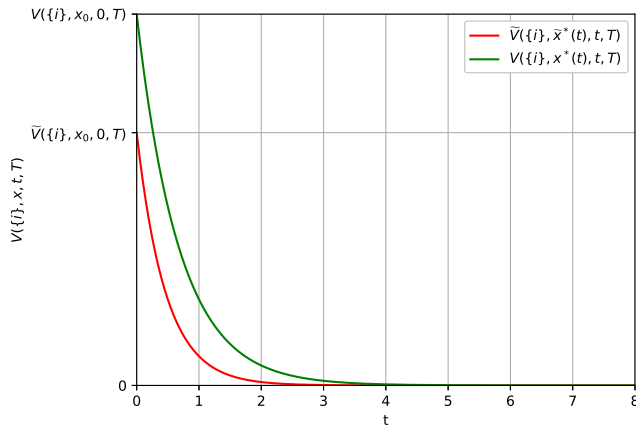


Рис. 4.29: Функция выигрыша с непрерывным обновлением $\tilde{V}(\{i\}, \tilde{x}^*(t), t, T)$ — красная линия снизу, функция выигрыша на бесконечном временном интервале $V(\{i\}, x^*(t), t, T)$ — зеленая линия сверху.

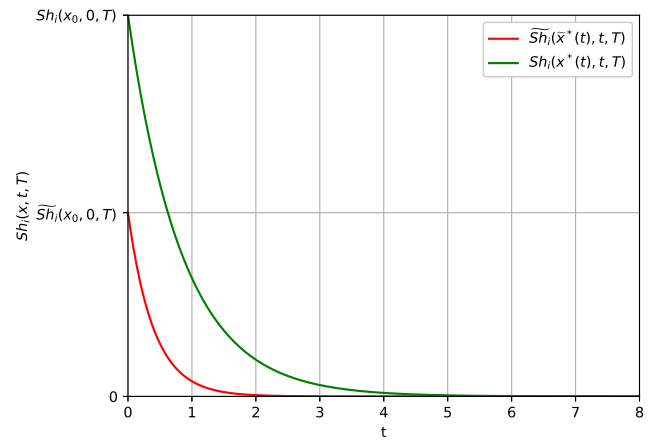


Рис. 4.30: Вектор Шепли с непрерывным обновлением информации $\tilde{Sh}_i(\tilde{x}^*(t), t, T)$ — красная линия снизу, вектор Шепли на бесконечном временном интервале $Sh_i(x^*(t), t, T)$ — зеленая линия сверху.

4.3.4.9 Численное моделирование

Рассмотрим результаты численного моделирования для представленной выше игровой модели на временном интервале $[0, 8]$, т.е. $t_0 = 0$ и $T = 8$. В начальный момент времени $t_0 = 0$ запас знаний равен 100, т.е. $x_0 = 100$. Остальные параметры модели таковы: $\beta = 0.9$, $r = 6$, $q = -1$, и $\bar{T} = 3$. На рисунке 4.25 представлено сравнение равновесия по Нэшу с непрерывным обновлением информации (красные линии) и равновесия по Нэшу в игре на бесконечном интервале времени (зеленые линии). На рисунке 4.26 показаны аналогичные результаты для соответствующих стратегий. На рисунках 4.27–4.28 представлено сравнение кооперативных решений, а на рисунке 4.29 — сравнение функций выигрыша в некооперативном случае. Наконец, на рисунке 4.30 представлено сравнение векторов Шепли для этих двух случаев.

4.4 Класс дифференциальных игр с нетрансферабельной полезностью и непрерывным обновлением информации

Настоящий раздел посвящен классу кооперативных дифференциальных игр с нетрансферабельной полезностью и непрерывным обновлением информации. Описан процесс построения оптимальной по Парето стратегии с непрерывным обновлением информации и траектории Парето. Еще одним важным результатом является свойство временной состоятельности, введенное для класса игр с непрерывным обновлением информации. В качестве примера рассмотрена игровая модель добычи ресурса. Построены оптимальные по Парето стратегии и соответствующая траектория, а также представлено множество оптимальных по Парето стратегий, удовлетворяющих свойству временной состоятельности. В этом разделе используются позиционные оптимальные по Парето стратегии, поэтому приняты упрощенные обозначения.

4.4.1 Парето-оптимальные стратегии с непрерывным обновлением информации

В рамках непрерывно обновляемой информации важно моделировать поведение игроков. Для этого используем концепцию Парето-оптимальности. Для класса дифференциальных игр с непрерывным обновлением информации хотелось бы ввести эту концепцию в следующем виде:

- для любого фиксированного $t \in [t_0, +\infty)$, $u^P(t, x) = (u_1^P(t, x), \dots, u_n^P(t, x))$ совпадает с Парето-оптимальной ситуацией в игре (4.5), (4.6), определенной на временном интервале $[t, t + \bar{T}]$ в момент времени t .

Прямое применение классических подходов для определения Парето-оптимальных стратегий с непрерывным обновлением информации невозможно. Для построения таких стратегий рассмотрим концепцию обобщенных Парето-оптимальных

стратегий как принцип оптимальности:

$$\tilde{u}^P(t, x; s, x_t) = (\tilde{u}_i^P(t, x; s, x_t), i = 1 \dots n), \quad t \in [t_0, +\infty], \quad s \in [t, t + \bar{T}], \quad (4.285)$$

который будет использован в дальнейшем для построения стратегий $u^P(t, x)$.

Определение 4.4.1 Ситуация $\tilde{u}^P(t, x; s, x_t) = (\tilde{u}_1^P(t, x; s, x_t), \dots, \tilde{u}_n^P(t, x; s, x_t))$ называется обобщенной Парето-оптимальной ситуацией в игре с непрерывным обновлением информации, если для любого фиксированного $t \in [t_0, +\infty)$ ситуация $\tilde{u}^P(t, x; s, x_t)$ является Парето-оптимальной в игре $\Gamma(x, t, \bar{T})$.

Определение 4.4.2 Ситуация $u^P(t, x)$ называется Парето-оптимальной ситуацией с непрерывным обновлением информации, если

$$u^P(t, x) = \tilde{u}^P(t, x; s, x_t)|_{s=t}, \quad t \in [t_0, +\infty), \quad (4.286)$$

где $\tilde{u}^P(t, x; s, x_t)$ — обобщенная Парето-оптимальная ситуация согласно Определению (4.4.1).

Траектория $x^*(t)$, соответствующая Парето-оптимальной ситуации с непрерывным обновлением информации $u^P(t, x)$ может быть получена из системы (4.8).

4.4.2 Уравнение Гамильтона–Якоби–Беллмана с непрерывным обновлением информации

Чтобы построить Парето-оптимальные стратегии и соответствующие траектории с непрерывным обновлением информации, необходимо для каждого вектора весовых коэффициентов $\alpha : \alpha_i \in (0, 1), \sum_{i=1}^n \alpha_i = 1$, рассмотреть следующую задачу оптимизации [234]:

$$\sum_{i=1}^n \alpha_i K_i^t(x, t; u^t) \longrightarrow \max_{u_1^t, \dots, u_n^t} \text{ при условии (4.5),} \quad (4.287)$$

где $t \in [t_0, +\infty]$ — текущий момент времени. (Парето-оптимальную ситуацию будем обозначать через $u^\alpha(t, x)$, а обобщенную Парето-оптимальную ситуацию

— через $\tilde{u}^\alpha(t, x; s, x_t)$.) Чтобы решить (4.287) для фиксированного вектора весовых коэффициентов α или найти ситуацию $u^\alpha(t, x)$, необходимо определить обобщенную Парето-оптимальную ситуацию $\tilde{u}^\alpha(t, x; s, x_t)$. Для этого воспользуемся модернизированной версией динамического программирования. Комбинируя всевозможные $u^\alpha(t, x)$ для различных векторов α , получаем набор Парето-оптимальных ситуаций с непрерывным обновлением информации. В классической постановке весовые коэффициенты α определяют соглашение между игроками, но они могут быть пересмотрены. В данном разделе считаем, что весовые коэффициенты α зафиксированы в начале игры, $\alpha(t) = \alpha$.

Обозначим через $W^\alpha(t; s, x)$ функцию Беллмана в подыгре, начинающейся в момент времени s из состояния x игры, начинающейся в текущий момент времени t :

$$W^\alpha(t; s, x) = \max_{u_1^t, \dots, u_n^t} \sum_{i=1}^n \alpha_i K_i^t(x, s, u^t) \text{ при условии (5.4)}. \quad (4.288)$$

Уравнение Гамильтона–Якоби–Беллмана (ГЯБ) имеет следующий вид.

Теорема 4.4.1 *Ситуация $\tilde{u}^\alpha(t, x; s, x_t)$ является обобщенной Парето-оптимальной ситуацией в дифференциальной игре с непрерывным обновлением информации, если существуют непрерывно дифференцируемые по s и x функции $W^\alpha(t; s, x) : [t_0, +\infty) \times [t, t + \bar{T}] \times R \rightarrow R$, которые удовлетворяют следующей системе уравнений в частных производных:*

$$\begin{aligned} -W_s^\alpha(t; s, x) &= \max_{\phi_i} \left\{ \sum_{i=1}^n \alpha_i g^i(s, x, \tilde{u}_{-i}^\alpha) + W_x^{\alpha_i}(t; s, x) f(s, x, \tilde{u}_{-i}^\alpha) \right\} \\ &= \alpha_i g^i(s, x, \tilde{u}^\alpha) + W_x^\alpha(t; s, x) f(s, x, \tilde{u}^\alpha), \\ W^\alpha(t; t + \bar{T}, x) &= 0, \quad i \in N, \end{aligned} \quad (4.289)$$

где $\tilde{u}_{-i}^\alpha(\phi_i) = (\tilde{u}_1^{\alpha_1}, \dots, \phi_i, \dots, \tilde{u}_n^{\alpha_n})$.

Доказательство Согласно определению обобщенной Парето-оптимальной ситуации, ситуация $\tilde{u}^\alpha(t, x; s, x_t)$ должна быть Парето-оптимальной для любого фиксированного t .

Зафиксировав t в формулировке Теоремы 4.4.1 и, в частности, в (4.289), получим классические достаточные условия Парето-оптимальной ситуации в дифференциальной игре с заданной продолжительностью $[t, t + \bar{T}]$ (см. [80]). Следовательно, для любого фиксированного t выполнены условия определения обобщенной Парето-оптимальной ситуации. Теорема доказана. \square

Мы рассматриваем только класс обобщенных Парето-оптимальных ситуаций, в котором для Парето-оптимальной ситуации с непрерывным обновлением информации решение системы (4.8) удовлетворяет условиям существования, единственности и продолжимости А.Ф. Филиппова [135]. Если можно найти обобщенную Парето-оптимальную ситуацию $\tilde{u}^\alpha(t, x; s, x_t)$ с помощью уравнений (4.289), то в результате применения процедуры (4.286) получаем желаемую ситуацию $u^\alpha(t, x)$.

4.4.3 Временная состоятельность в игре с непрерывным обновлением информации

В модели кооперации с нетрансферабельными выигрышами игроки договариваются о некотором соглашении (принципе оптимальности) относительно того, как будет развиваться кооперативная игра. В частности, выбранный принцип оптимальности должен удовлетворять групповой оптимальности (i) и индивидуальной рациональности (ii) вдоль выбранной траектории $x^*(t)$ (в нашем случае, Парето-оптимальной траектории). Свойство временной состоятельности решения в игре начинающейся в текущий момент времени требует, чтобы решение игры в подыгре, начинающейся в любой более поздний момент времени и в любом возможном состоянии вдоль траектории, вызванной оптимальным решением в текущей игре, оставалось оптимальным. Для этого необходимы как групповая оптимальность, так и индивидуальная рациональность. При групповой оптимальности игроки ищут набор кооперативных стратегий или управлений, который дает Парето-оптимальное решение. Решение должно удовлетворять индивидуальной рациональности: все игроки получают более высокий выигрыш в случае кооперации, а не индивидуального поведения.

Согласно процедуре (4.286), Парето-оптимальные стратегии и траектории

с непрерывным обновлением информации удовлетворяют свойству групповой оптимальности. Но свойство индивидуальной рациональности не всегда выполняется [234]. Для класса игр с непрерывным обновлением информации свойство индивидуальной рациональности принимает вид:

$$K_i^{t,NE}(x^*(t), t; \tilde{u}^{NE}(t, x; s, x_t)) \leq K_i^{t,\alpha}(x^*(t), t; \tilde{u}^\alpha(t, x; s, x_t)), \forall i, \forall t, \text{ где}$$

$K_i^{t,NE}(x^*(t), t; \tilde{u}^\alpha(t, x; s, x_t))$ — индивидуальный выигрыш игрока i в равновесии по Нэшу в игре, заданной на временном интервале $[t, t + \bar{T}]$ и начинающейся на Парето-оптимальной траектории $x^*(t)$, $K_i^{t,\alpha}(x^*(t), t; \tilde{u}^\alpha(t, x; s, x_t))$ — индивидуальный выигрыш при кооперации (4.286) в игре, начинающейся на Парето-оптимальной траектории $x^*(t)$. Временная состоятельность может быть формально определена следующим образом.

Определение 4.4.3 *Парето-оптимальная ситуация $u^\alpha(t, x)$ называется состоятельной во времени, если соответствующая обобщенная Парето-оптимальная ситуация $\tilde{u}^\alpha(t, x; s, x_t)$ такова, что выполняются следующие условия:*

- (i) *Групповая оптимальность: $(K_i^{t,\alpha}(x^*(t), t; \tilde{u}^\alpha(t, x; s, x_t)), i \in N)$, Парето-оптимальна;*
- (ii) *Индивидуальная рациональность:*

$$K_i^{t,NE}(x^*(t), t; \tilde{u}^{NE}(t, x; s, x_t)) \leq K_i^{t,\alpha}(x^*(t), t; \tilde{u}^\alpha(t, x; s, x_t)), \forall i \text{ и } \forall t.$$

Предположим, что существует множество A весовых коэффициентов α такое, что выполнены условия (i) и (ii). Множество Парето-оптимальных стратегий $u^\alpha(t, x)$, $\alpha \in A$, будем называть состоятельным во времени кооперативным решением с непрерывным обновлением информации.

4.4.3.1 Дифференциальная игра добычи невозобновляемого ресурса

В качестве наглядного примера рассмотрим дифференциальную игру с непрерывным обновлением информации для добычи невозобновляемого ресурса (см. [50]).

4.4.3.1.1 Исходная игра

Пусть вектор состояния $x(t)$ описывает запас ресурса в момент времени t . Обозначим через $u_i(t, x)$ скорость извлечения ресурса игроком i в момент времени t при условии, что запас ресурса равен x . Предположим, что $u_i(t, x) \geq 0$ и $u_i(t, x) = 0$ при $x(t) = 0$.

Динамика запаса задается дифференциальным уравнением

$$\dot{x}(t) = - \sum_{i=1}^n b_i u_i(t, x), \quad x(t_0) = x_0, \quad (4.290)$$

где $b_i > 0$ при всех $i = 1, \dots, n$ и $x_0 > 0$. Функция выигрыша игрока i имеет вид

$$K_i(x_0, T - t_0) = \int_{t_0}^T \ln u_i(t, x) dt, \quad i \in N. \quad (4.291)$$

4.4.3.2 Парето-оптимальные стратегии с непрерывным обновлением информации

Согласно Разделу 4.4.1, для определения Парето-оптимальных стратегий в игре с непрерывным обновлением информации рассмотрим семейство вспомогательных подыгр $\Gamma(x, t, t + \bar{T})$ с продолжительностью \bar{T} , начинающихся в момент времени t из состояния x . Чтобы определить Парето-оптимальные стратегии $\tilde{u}^\alpha(t, x; s, x_t)$ во вспомогательной подыгре $\Gamma(x, t, t + \bar{T})$, используем динамическое программирование. Для текущего момента времени t , обозначим через $W^\alpha(t, s, x)$ функцию Беллмана в подыгре, начинающейся в момент времени s :

$$W^\alpha(t; s, x) = \max_{u_1^t, \dots, u_n^t} \sum_{i=1}^n \alpha_i K_i(x, s; u_t) \quad (4.292)$$

$$\text{при условии : } \dot{x}_t(s) = - \sum_{i=1}^n b_i u_i^t(s, x_t), \quad x_t(t) = x.$$

Уравнение Гамильтона–Якоби–Беллмана (ГЯБ) имеет следующий вид:

$$-\frac{\partial W^\alpha(t; s, x)}{\partial s} = \max_{u_1^t, \dots, u_n^t} \left\{ \sum_{i=1}^n \alpha_i \ln u_i^t(s, x) - \frac{\partial W^\alpha(t; s, x)}{\partial x} \sum_{i=1}^n b_i u_i^t(s, x) \right\},$$

$$\lim_{s \rightarrow t+\bar{T}} W^\alpha(t; s, x) = 0. \quad (4.293)$$

Найдем решение уравнения (4.293) в виде $W^\alpha(t; s, x) = A(t, s) \ln x + B(t, s)$. Частные производные этой функции равны

$$W_s^\alpha(t; s, x) = \dot{A}(t, s) \ln x + \dot{B}(t, s),$$

$$W_x^\alpha(t; s, x) = \frac{A(t, s)}{x}, \quad (4.294)$$

где $\dot{A}(t, s)$ и $\dot{B}(t, s)$ — производные функции по переменной s . Максимизируя выражение в правой части (4.293) и подставляя результат в (4.294), получаем:

$$\tilde{u}_i^\alpha(t, x; s, x_t) = \frac{x \alpha_i}{b_i A(t, s)}, \quad \sum_{i=1}^n \alpha_i = 1. \quad (4.295)$$

Подстановка формул (4.294) и (4.295) в уравнение (4.293) дает систему дифференциальных уравнений

$$\dot{A}(t, s) \ln x + \dot{B}(t, s) = -\ln x + \ln A(t, s) + \ln b^N - \sum_{i=1}^n \alpha_i \ln \alpha_i + n, \quad (4.296)$$

$$\lim_{s \rightarrow t+\bar{T}} A(t, s) \ln x + B(t, s) = 0,$$

где $b^N = \prod_{i=1}^n b_i$. Тогда

$$\dot{A}(t, s) = -1, \quad \lim_{s \rightarrow t+\bar{T}} A(t, s) = 0, \quad (4.297)$$

$$\dot{B}(t, s) = \ln A(t, s) + \ln b^N - \sum_{i=1}^n \alpha_i \ln \alpha_i + n, \quad \lim_{s \rightarrow t+\bar{T}} B(t, s) = 0.$$

Решение системы (4.296) имеет вид

$$B(s, t) = -(t + \bar{T} - s) \left(- \sum_{i=1}^n \alpha_i \ln \alpha_i + \ln b^N + n + \ln n(t + \bar{T} - s) \right), \quad (4.298)$$

$$A(s, t) = n(t + \bar{T} - s), \quad s \in [t, t + \bar{T}), \quad t \in [t_0, +\infty). \quad (4.299)$$

Таким образом, приходим к следующим Парето-оптимальным стратегиям во вспомогательной подыгре $\Gamma(x, t, t + \bar{T})$:

$$\tilde{u}_i^\alpha(t, x; s, x_t) = \frac{x \alpha_i}{b_i n(t + \bar{T} - s)}, \quad s \in [t, t + \bar{T}), \quad t \in [t_0, +\infty), \quad \sum_{i=1}^n \alpha_i = 1. \quad (4.300)$$

Кроме того,

$$\begin{aligned} W^\alpha(t; s, x) &= \\ &= (t + \bar{T} - s) \left(\ln b^N + \sum_{i=1}^n \alpha_i \ln \alpha_i - n - n \ln \frac{x}{n(t + \bar{T} - s)} \right), \quad s \in [t, t + \bar{T}). \end{aligned} \quad (4.301)$$

Следуя процедуре (4.286), строим Парето-оптимальные стратегии с непрерывным обновлением информации:

$$u^\alpha(t, x) = \tilde{u}^\alpha(t, x; s, x_t)|_{s=t} = \left(\frac{x \alpha_i}{b_i \bar{T}}, \quad i = 1 \dots n \right), \quad s \in [t, t + \bar{T}], \quad \sum_{i=1}^n \alpha_i = 1. \quad (4.302)$$

Подставляя выражение (4.302) в дифференциальное уравнение (4.290), получаем Парето-оптимальную траекторию $x^*(t)$ с непрерывным обновлением информации вида

$$x^*(t) = x_0 \exp \frac{-t}{\bar{T}}. \quad (4.303)$$

Очевидно, Парето-оптимальная ситуация (4.302) зависит от весовых коэффициентов α , но Парето-оптимальная траектория (4.303) не зависит от него, поскольку сумма весовых коэффициентов α равна 1.

4.4.3.3 Временная состоятельность Парето-оптимального решения

В этом разделе построим состоятельную во времени Парето-оптимальную ситуацию $u^\alpha(t, x)$, используя условия из Определения 4.4.3. Для этого необходимо вычислить $K_i^{t,\alpha}(x^*(t), t; \tilde{u}^\alpha(t, x; s, x_t))$ и $K_i^{t,NE}(x^*(t), t; \tilde{u}^{NE}(t, x; s, x_t))$. Сначала нужно найти $\tilde{x}_t^\alpha(s)$ и $\tilde{x}_t^{NE}(s)$, которые являются соответственно мнимой (ожидаемой) Парето-оптимальной траекторией и равновесной по Нэшу траекторией для игры, определенной на временном интервале $[t, t + \bar{T}]$, начинающейся на Парето-оптимальной траектории $x^*(t)$. Учитывая, что $x_0 = x^*(t)$, $x(t) = x_t(s)$, подставляем Парето-оптимальную ситуацию (4.300) в (4.290) и получаем:

$$\tilde{x}_t^\alpha(s) = x^*(t) \cdot \frac{t + \bar{T} - s}{\bar{T}} \cdot \exp^{\frac{1}{n}}, \quad s \in [t, t + \bar{T}]. \quad (4.304)$$

Тогда можно вычислить Парето-оптимальную ситуацию $\tilde{u}^\alpha(t, x; s, \tilde{x}_t^\alpha(s))$ вдоль траектории (4.304) на временном интервале $[t, t + \bar{T}]$:

$$\tilde{u}^\alpha(t, x; s, \tilde{x}_t^\alpha(s)) = \frac{x^*(t) \cdot \alpha_i \exp^{\frac{1}{n}}}{b_i n \bar{T}}, \quad s \in [t, t + \bar{T}], \quad \sum_{i=1}^n \alpha_i = 1. \quad (4.305)$$

Теперь необходимо получить выражения для Парето-оптимального выигрыша игрока i в момент времени t с непрерывным обновлением информации, подставив (4.305) в функцию выигрыша с непрерывным обновлением информации (4.291) (с пределами интегрирования t и $t + \bar{T}$):

$$K_i^{t,\alpha}(x^*(t), t; \tilde{u}^\alpha(t, x; s, x_t)) = \int_t^{t+\bar{T}} \ln \frac{x^*(t) \cdot \alpha_i \exp^{\frac{1}{n}}}{b_i n \bar{T}} ds. \quad (4.306)$$

Используя определение из статьи [78], можем найти равновесные по Нэшу стратегии с непрерывным обновлением информации для этой конкретной игровой модели. Функция выигрыша игрока i в момент времени s в равновесии по Нэшу в игре, заданной на временном интервале $[t, t + \bar{T}]$, имеет вид

$$\begin{aligned} K_i^{t,NE}(x^*(t), t; \tilde{u}^{NE}(t, x; s, x_t)) &= \int_t^{t+\bar{T}} \ln \tilde{u}^\alpha(t, x; s, \tilde{x}_t^{NE}(s)) ds = \\ &= \int_t^{t+\bar{T}} \ln \frac{x^*(t) \cdot (t + \bar{T} - s)}{b_i (t + \bar{T} - s)} ds = \int_t^{t+\bar{T}} \ln \frac{x^*(t)}{b_i} ds. \end{aligned} \quad (4.307)$$

Определим набор весовых коэффициентов α , который удовлетворяет (4.308) и, следовательно, гарантирует временную состоятельность Парето-оптимальных стратегий в этой игре с непрерывным обновлением информации:

$$K_i^{t,\alpha}(x^*(t), t; \tilde{u}^\alpha(t, x; s, x_t)) \geq K_i^{t,NE}(x^*(t), t; \tilde{u}^{NE}(t, x; s, x_t)), \quad \sum_{i=1}^n \alpha_i = 1. \quad (4.308)$$

Решением является:

$$\begin{aligned} \ln \frac{x^*(t) \cdot \alpha_i \exp^{\frac{1}{n}}}{b_i n \bar{T}} &\geq \ln \frac{x^*(t)}{b_i}, \\ \ln \alpha_i &\geq \ln \frac{x^*(t)}{b_i} - \ln \frac{x^*(t) \cdot \exp^{\frac{1}{n}}}{b_i n \bar{T}}, \\ \alpha_i &\geq \frac{n \bar{T}}{\exp^{\frac{1}{n}}}, \quad \sum_{i=1}^n \alpha_i = 1. \end{aligned} \quad (4.309)$$

Это решение показывает, при каких условиях Парето-оптимальная ситуация с непрерывным обновлением информации может удовлетворять свойству временной состоятельности.

Рассмотрим числовой пример с $n = 2$, $x_0 = 5$, $b_i = 2$, $\bar{T} = 0.5$, $\alpha_i \in \{0.1, 0.2, \dots, 1\}$. На рисунке 4.31 представлено сравнение траектории в исходной игре и игре с непрерывным обновлением информации. На рисунке 4.32 показано различие между Парето-оптимальными стратегиями в исходной игре и игре с непрерывным обновлением информации для весовых коэффициентов $\alpha_1 = 0,664$ и $\alpha_2 = 0,336$. На рисунке 4.33 представлена Парето-оптимальная стратегия игрока i для различных значений α . На рисунке 4.34 изображены функции выигрыша (4.306) для разных весовых коэффициентов α и равновесный по Нэшу выигрыш (4.307). Очевидно, что красная линия (4.307) делит синие линии (4.306) на две части. Набор синих линий над красной линией — это то, что нужно найти (4.309). Синие линии (4.306) под красной линией (4.307) не удовлетворяют свойству временной состоятельности.

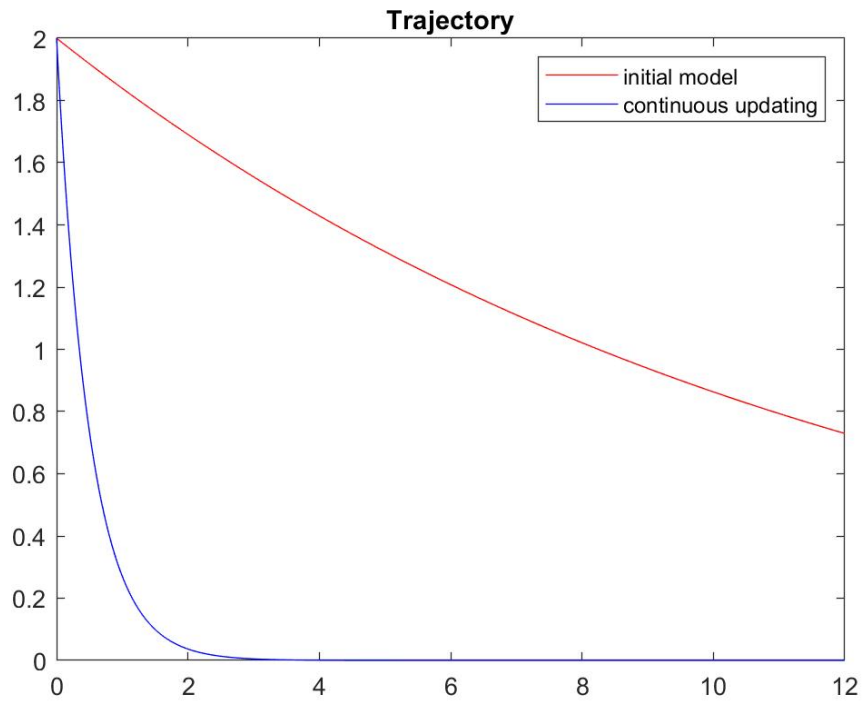


Рис. 4.31: Парето-оптимальная траектория с непрерывным обновлением информации (синяя линия) и в исходной игре (красная линия).

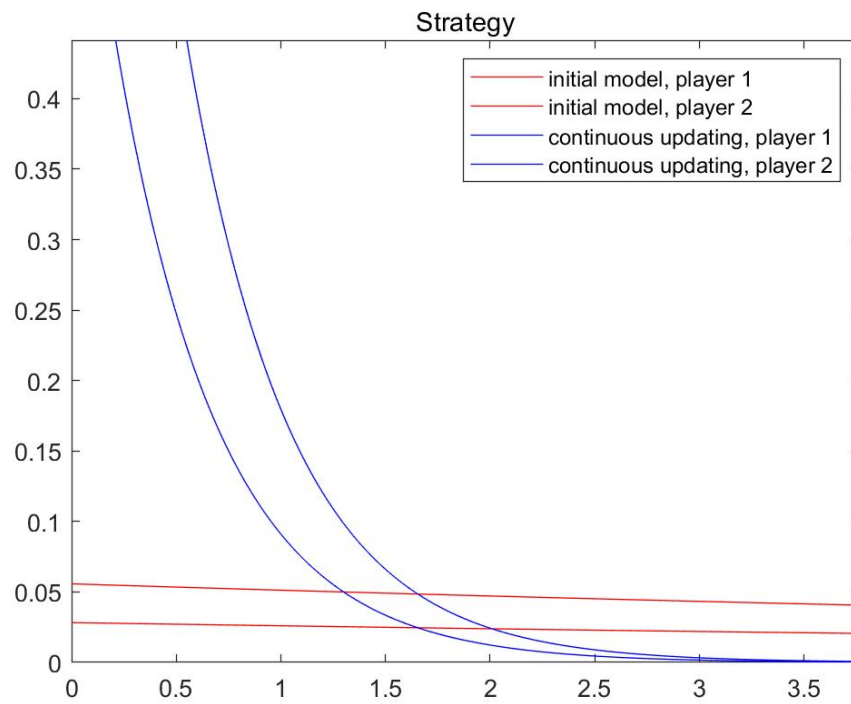


Рис. 4.32: Парето-оптимальные стратегии игроков 1 и 2 в исходной игре и игре с непрерывным обновлением информации при $\alpha_1 = 0.664$ и $\alpha_2 = 0.336$.

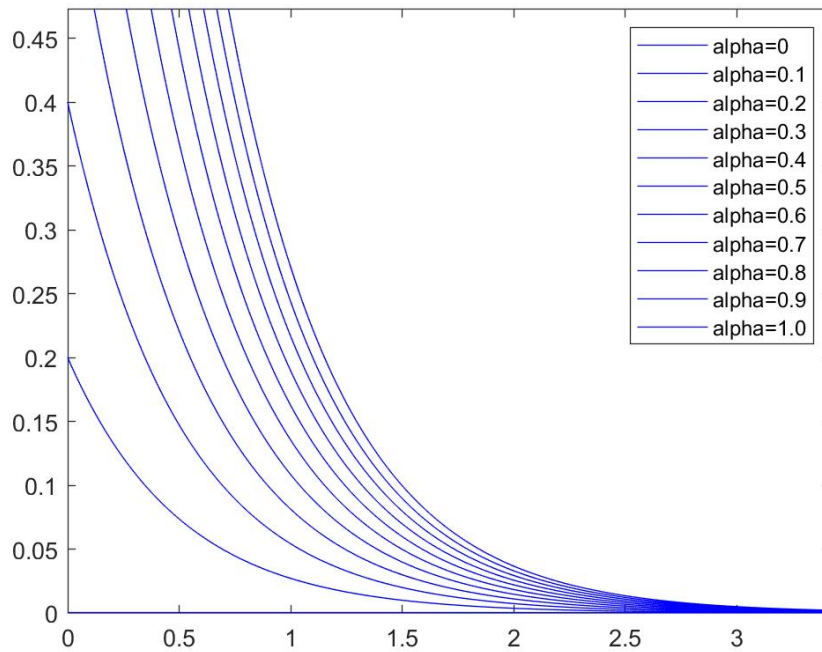


Рис. 4.33: Парето-оптимальная стратегия игрока i с непрерывным обновлением информации при $\alpha_i = (0.1, 0.2, \dots, 1)$.

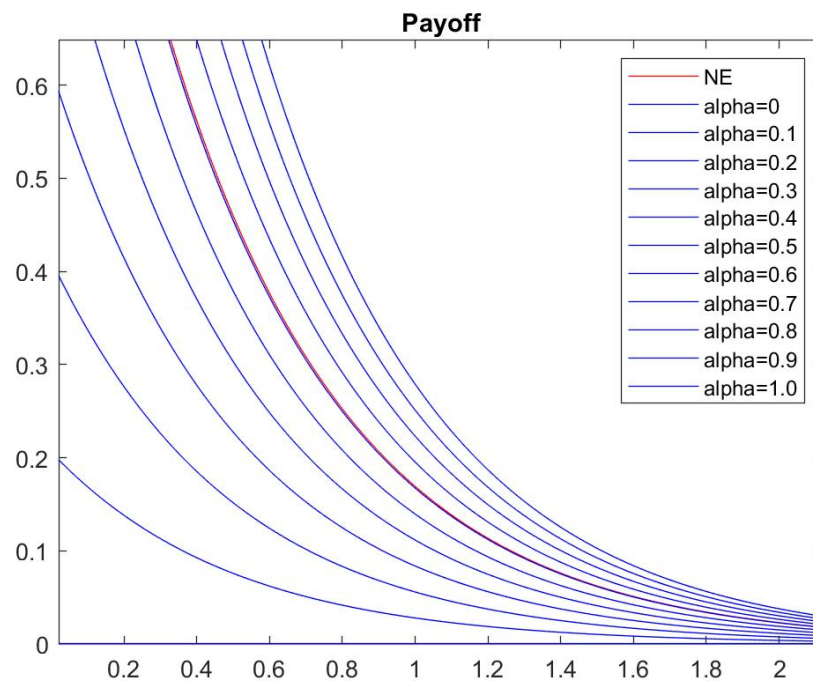


Рис. 4.34: Функция выигрыша (4.306) игрока i в Парето-оптимальной ситуации (синие линии) и функция выигрыша (4.307) игрока i в равновесии по Нэшу (красные линии) с непрерывным обновлением информации.

4.5 Дифференциальная игра добычи ресурса с непрерывным и динамическим обновлением информации

В этом разделе описывается модель добычи ресурса для случая динамического и непрерывного обновления информации. Доказано, что оптимальное управление (кооперативные стратегии) и позиционные равновесные по Нэшу стратегии равномерно сходятся к соответствующим стратегиям в игре с непрерывным обновлением информации при стремлении к бесконечности количества моментов обновления информации. Аналогичные результаты представлены для оптимальной траектории (кооперативной траектории), равновесной траектории и соответствующих выигрышей.

В этом разделе будем использовать только позиционные стратегии для построения кооперативных стратегий с непрерывным обновлением информации $u_{fb}^*(t, x)$ и позиционные равновесные по Нэшу стратегии $u_{fb}^{NE}(t, x)$ для построения характеристической функции с непрерывным обновлением информации. Поэтому введем упрощенные обозначения $u^*(t, x)$ и $u^{NE}(t, x)$.

4.5.1 Исходная игра добычи ресурса

Рассмотрим классическую теоретико-игровую модель добычи невозобновляемого ресурса с симметричными игроками [168]. Здесь количество ресурса напрямую зависит от скоростей добычи, которые выбирают игроки (компании). В игре участвует n симметричных игроков (множество игроков обозначим через N) с функциями полезности, зависящими от скоростей добычи $h_i(t, x, u_i) = \log u_i$. Игра начинается в момент времени t_0 и заканчивается в момент времени T , т.е. игра определена на интервале $[t_0, T]$. Количество ресурса в начале игры равно x_0 .

Обозначим через $x(t) \in R^1$ количество ресурса, доступного игрокам в момент времени t , а через $u_i(t, x)$ — стратегию игрока $i \in N$, которая представляет собой скорость добычи ресурса в зависимости от t и x . Будем искать стратегии в классе позиционных стратегий, в предположении, что $\forall t$, $u_i(t, x) \geq 0$ и $u_i(t, x) = 0$ при $x(t) = 0$. Количество ресурса $x(t)$ как функция

времени t зависит от скоростей добычи (стратегий игроков $u_i(t, x)$) следующим образом:

$$\begin{aligned}\dot{x} &= - \sum_{i=1}^n u_i(t, x), \\ x(t_0) &= x_0.\end{aligned}\tag{4.310}$$

Функция выигрыша игрока $i \in N$ имеет вид

$$K_i(x_0, T - t_0) = \int_{t_0}^T \log(u_i(\tau, x)) d\tau, \quad i \in N.\tag{4.311}$$

Предположим, что для любого набора n стратегий $u_1(\cdot), \dots, u_n(\cdot)$ выполнены условия существования, единственности и продолжимости решений (4.310) [160]. Учитывая симметрию игроков, положим $u(t, x) = u_i(t, x)$ для каждого $i \in N$.

В следующем разделе рассчитываются оптимальные стратегии (управления) для двух основных классов дифференциальных игр: кооперативных и некооперативных дифференциальных игр.

4.5.1.1 Кооперативная дифференциальная игра

Рассмотрим кооперативную постановку для игры добычи ресурса [168]. Здесь игроки объединяются в гранд коалицию $S = N$ и, действуя как один игрок, максимизируют суммарный выигрыш. Соответствующая задача оптимального управления формулируется следующим образом:

$$\sum_{i=1}^n K_i(x_0, T - t_0) = n \int_{t_0}^T \log(u(\tau, x)) d\tau \rightarrow \max_u \tag{4.312}$$

при условии

$$\begin{aligned}\dot{x} &= -nu(t, x), \\ x(t_0) &= x_0 > 0, \\ u(t, x) &\geq 0.\end{aligned}\tag{4.313}$$

Для решения задачи оптимизации (4.312), (4.313) используем метод динамического программирования, предложенный в этой диссертации. Определим функцию Беллмана как максимальное значение функционала (4.311) в подыгре $\Gamma(x, T - t)$, начинающейся в момент времени t из позиции x :

$$V(t, x) = \max_u \left\{ \sum_{i=1}^n K_i(x, T - t) \right\} = \max_u \left\{ n \int_t^T \log u(\tau, x) d\tau \right\}. \quad (4.314)$$

Следующий факт был установлен в статье [2]: если существует непрерывно дифференцируемая функция $V(t, x)$, удовлетворяющая уравнению Гамильтона–Якоби–Беллмана

$$\begin{aligned} -V_t(t, x) &= \max_u \{ n \log u - nuV_x(t, x) \}, \\ V(T, x) &= 0, \end{aligned} \quad (4.315)$$

то управление $u^*(t, x)$, полученное при максимизации правой части (4.315), оптимально в задаче управления (4.312), (4.313).

Из необходимых условий оптимальности первого порядка для (4.315) находим

$$u^* = \frac{1}{V_x(t, x)}.$$

Подставляя это управление в уравнение (4.315), получаем:

$$\begin{aligned} V_t(t, x) &= n \log V_x(t, x) + n, \\ V(T, x) &= 0. \end{aligned}$$

Определим функцию Беллмана в виде

$$V(t, x) = A(t) \log x + B(t).$$

Подстановка этого выражения в уравнение (4.315) дает

$$\begin{aligned} \dot{A}(t) \log x + \dot{B}(t) &= n \log A(t) - n \log x + n, \\ A(T) &= B(T) = 0. \end{aligned} \quad (4.316)$$

Решением системы (4.316) являются функции

$$\begin{aligned} A(t) &= n(T - t), \\ B(t) &= -n(T - t) \log n(T - t). \end{aligned} \quad (4.317)$$

Итак, функция Беллмана принимает вид

$$V(t, x) = n(T - t) \log \frac{x}{n(T - t)}, t \in [t_0, T], \quad (4.318)$$

а оптимальное управление — вид

$$u^*(t, x) = \frac{1}{V_x(t, x)} = \frac{x}{n(T - t)}, t \in [t_0, T]. \quad (4.319)$$

Подставляя оптимальное управление в уравнения движения (4.313), получаем дифференциальное уравнение для траектории, соответствующей оптимальному управлению:

$$\begin{aligned} \dot{x}(t) &= -\frac{x(t)}{T - t}, \\ x(t_0) &= x_0. \end{aligned} \quad (4.320)$$

Решением задачи Коши (4.320) является функция

$$x^*(t) = x_0 \frac{T - t}{T - t_0}, t \in [t_0, T]. \quad (4.321)$$

Траектория $x^*(t)$ и стратегия (управление) $u^*(t, x)$ называются кооперативными.

Кооперативные стратегии вдоль кооперативной траектории имеют вид

$$u^*(t, x^*) = \frac{x_0}{n(T - t)}, t \in [t_0, T]. \quad (4.322)$$

На рисунке 4.35 зеленой линией показана кооперативная траектория $x^*(t)$ (4.321), а на рисунке 4.36 — соответствующие оптимальные стратегии (4.322) в игре добычи невозобновляемого ресурса $\Gamma(x_0, T - t_0)$.

4.5.1.2 Некооперативная дифференциальная игра

Рассмотрим случай, когда каждый игрок принимает индивидуальное решение о скорости добычи ресурса $u_i(t, x)$. В качестве принципа оптимальности будем использовать равновесие по Нэшу в позиционных стратегиях. Введем следующее обозначение для выигрыша игрока i в равновесии по Нэшу в позиционных стратегиях $u^{NE}(t, x) = (u_1^{NE}(t, x), \dots, u_n^{NE}(t, x))$ в подыгре $\Gamma(x, T - t)$, начинающейся в момент времени t из позиции x :

$$V_i(t, x) = \int_t^T \log(u_i^{NE}(t, x)) d\tau \quad (4.323)$$

при условии

$$\begin{aligned} \dot{x}(t) &= - \sum_{i=1}^n u_i^{NE}(t, x), \\ x(t) &= x. \end{aligned} \quad (4.324)$$

Чтобы найти равновесные по Нэшу стратегии $u_1^{NE}(t, x), \dots, u_n^{NE}(t, x)$, также применим принцип динамического программирования, описанный в [80]. Соответствующая система уравнений Гамильтона–Якоби–Беллмана имеет вид

$$\begin{aligned} -V_t^i(t, x) &= \max_{u_i} \left\{ \log u_i - V_x^i(t, x) \left(u_i + \sum_{k=1, k \neq i}^n u_k^{NE} \right) \right\}, \\ V_i(T, x) &= 0, \quad i \in N. \end{aligned} \quad (4.325)$$

Согласно [80], если существуют непрерывно дифференцируемые функции $V_i(t, x)$, удовлетворяющие (4.325), то $u_i^{NE}(t, x)$ является равновесием по Нэшу в позиционных стратегиях. Будем искать функцию Беллмана в виде:

$$V_i(t, x) = A_i(t) \log x + B_i(t), \quad i \in N.$$

Решая (4.325), получаем следующую функцию Беллмана:

$$V_i(t, x) = (T - t) \left(\log \frac{x}{T - t} - n + 1 \right), \quad t \in [t_0, T], \quad i \in N. \quad (4.326)$$

Таким образом, равновесием по Нэшу в позиционных стратегиях является ситуация

$$u_i^{NE}(t, x) = \frac{x}{T - t}, \quad t \in [t_0, T], \quad i \in N, \quad (4.327)$$

а соответствующая траектория имеет вид

$$x^{NE}(t) = x_0 \left(\frac{T - t}{T - t_0} \right)^n, \quad t \in [t_0, T]. \quad (4.328)$$

Равновесные по Нэшу стратегии $(u_1^{NE}(t, x), \dots, u_n^{NE}(t, x))$ вдоль траектории $x^{NE}(t)$ равны

$$u_i^{NE}(t, x^{NE}) = x_0 \frac{(T - t)^{n-1}}{(T - t_0)^n}, \quad t \in [t_0, T], \quad i \in N. \quad (4.329)$$

На рисунке 4.37 зеленой линией показаны позиционные равновесные по Нэшу стратегии $u^{NE}(t, x)$ (4.329), а на рисунке 4.38 — соответствующая равновесная по Нэшу траектория $x^{NE}(t)$ (4.328).

4.5.2 Игра с динамическим обновлением информации

При динамическом обновлении информации, временной интервал $[t_0, T]$ разбивается на l отрезков длины $\Delta t = \frac{T - t_0}{l}$, а поведение игроков на каждом временном отрезке $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, l$, моделируется с помощью усеченной подыгры.

4.5.2.1 Кооперативная игра с динамическим обновлением информации

Во-первых, рассмотрим кооперативную постановку для игры добычи ограниченного ресурса с динамическим обновлением информации. Поскольку структура усеченной подыгры на каждом временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ соответствует исходной игре, определенной на временном интервале $[t_0, T]$, то аналогично определяется и решение для каждой подыгры с номером j . Основное различие заключается в параметрах модели, а именно: $t_0 + j\Delta t$ —

начальный момент времени подыгры, $t_0 + j\Delta t + \bar{T}$ — конечный момент времени подыгры, а x_0^j — начальное количество ресурса в усеченной подыгре. Таким образом, кооперативные стратегии $u_j^*(x, t)$ для каждой подыгры $\bar{\Gamma}_j(x_0^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ имеют вид

$$u_j^*(t, x) = \frac{x}{n(t_0 + j\Delta t + \bar{T} - t)}, \quad t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}], \quad (4.330)$$

а соответствующая кооперативная траектория $x_j^*(t)$ равна

$$x_j^*(t) = x_0^j \frac{t_0 + j\Delta t + \bar{T} - t}{\bar{T}}, \quad t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]. \quad (4.331)$$

Заметим, что x_0^j зависит от конечной точки оптимальной траектории в предыдущей усеченной игре $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$:

$$x_0^j = x_{j-1}(t_0 + j\Delta t), \quad \text{где } x_0^0 = x_0.$$

Теорема 4.5.1 *Результирующая кооперативная траектория $\hat{x}^*(t)$ в игре с динамическим обновлением информации имеет следующий вид:*

$$\hat{x}^*(t) = x_0 \left(1 - \frac{\Delta t}{\bar{T}}\right)^j \frac{t_0 + j\Delta t + \bar{T} - t}{\bar{T}}, \quad t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}], \quad j = 0, \dots, l. \quad (4.332)$$

Доказательство Чтобы вывести явное выражение для результирующей траектории с использованием формулы (4.331), нужно определить параметр x_0^j для любой усеченной подыгры $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$.

Рассмотрим участки траектории с номерами $j - 1$ и j :

$$\begin{aligned} x_{j-1}^*(t) &= x_0^{j-1} \frac{t_0 + (j-1)\Delta t + \bar{T} - t}{\bar{T}}, \\ x_j^*(t) &= x_0^j \frac{t_0 + j\Delta t + \bar{T} - t}{\bar{T}}, \end{aligned}$$

где $x_0^j = x_{j-1}^*(t_0 + j\Delta t)$. Тогда:

$$\begin{aligned} x_0^j &= x_0^{j-1} \left(1 - \frac{\Delta t}{\bar{T}}\right), \\ x_0^j &= x_0^{j-2} \left(1 - \frac{\Delta t}{\bar{T}}\right)^2, \\ &\dots, \\ x_0^j &= x_0^0 \left(1 - \frac{\Delta t}{\bar{T}}\right)^j. \end{aligned}$$

С учетом $x_0^0 = x_0$, окончательно получаем (4.333):

$$x_0^j = x_0 \left(1 - \frac{\Delta t}{\bar{T}}\right)^j. \quad (4.333)$$

Подставляя выражение (4.333) в формулу (4.331), приходим к (4.332):

$$\hat{x}^*(t) = x_0 \left(1 - \frac{\Delta t}{\bar{T}}\right)^j \frac{t_0 + j\Delta t + \bar{T} - t}{\bar{T}}, \quad t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}], \quad j = 0, \dots, l.$$

□

Следовательно, кооперативные стратегии вдоль траектории (4.332) имеют вид

$$u_j^*(t, x_j^*) = \frac{x_0}{n\bar{T}} \left(1 - \frac{\Delta t}{\bar{T}}\right)^j, \quad t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]. \quad (4.334)$$

Обозначим через $\hat{u}^*(t, x)$ результирующие стратегии $\hat{u}(t, x)$ (4.110), а через $\hat{x}^*(t)$ — соответствующую результирующую кооперативную траекторию.

На рисунке 4.35 представлено сравнение результирующей кооперативной траектории $\hat{x}^*(t)$ (4.332) в игре с динамическим обновлением информации и кооперативной траектории $\hat{x}^*(t)$ (4.321) в исходной игре с заданной продолжительностью. На рисунке 4.36 представлено сравнение результирующих кооперативных стратегий $\hat{u}^*(t)$ (4.334) в игре с динамическим обновлением информации и кооперативных стратегий $u^*(t)$ (4.322) в исходной игре с заданной продолжительностью.

□

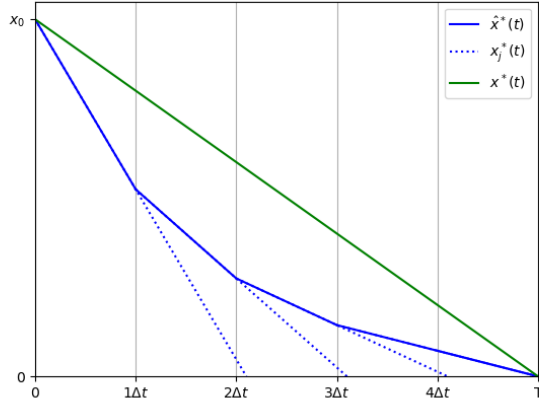


Рис. 4.35: Результирующая кооперативная траектория $\hat{x}^*(t)$ (4.332) (синяя линия) и кооперативная траектория в исходной игре $x^*(t)$ (4.321) (зеленая линия).

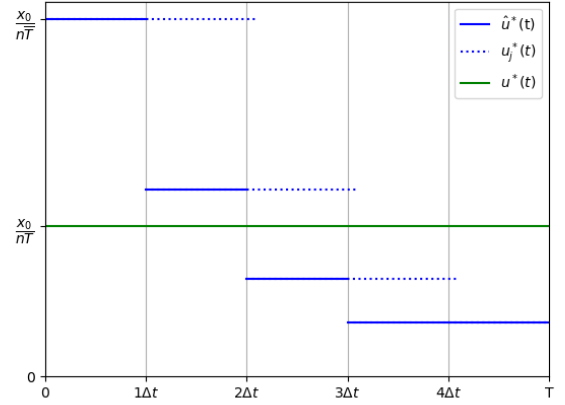


Рис. 4.36: Результирующие кооперативные стратегии $\hat{u}^*(t, \hat{x}^*)$ (4.334) (синяя линия) и кооперативные стратегии в исходной игре $u^*(t, x^*)$ (4.322) (зеленая линия).

4.5.2.2 Некооперативная игра с динамическим обновлением информации

Теперь рассмотрим модель некооперативной игры с динамическим обновлением информации. Здесь предполагается, что игроки действуют индивидуально. Как и в исходной игре, будем использовать равновесие по Нэшу в позиционных стратегиях в качестве принципа оптимальности. Выполняя вычисления, аналогичные приведенным в Разделе 4.5.1.2, получаем:

$$u_j^{NE}(t, x) = \frac{x}{\bar{T} + j\Delta t + t_0 - t}, \quad t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}], j = 0, \dots, l, \quad (4.335)$$

$$x_j^{NE}(t) = x_0^j \left(\frac{\bar{T} + j\Delta t + t_0 - t}{\bar{T}} \right)^n, \quad t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}], j = 0, \dots, l. \quad (4.336)$$

Теорема 4.5.2 Результирующая равновесная по Нэшу траектория $\hat{x}^{NE}(t)$ в

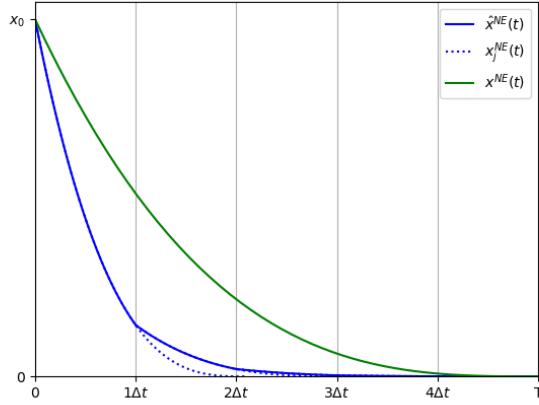


Рис. 4.37: Результирующая равновесная по Нэшу траектория $\hat{x}^{NE}(t)$ (4.337) (синяя линия) и равновесная по Нэшу траектория в исходной игре $x^{NE}(t)$ (4.321) (зеленая линия).

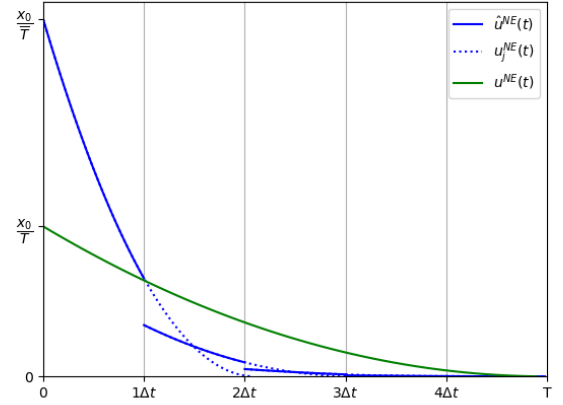


Рис. 4.38: Результирующие равновесные по Нэшу стратегии $\hat{u}^{NE}(t, \hat{x}^{NE})$ (4.338) (синяя линия) и равновесные по Нэшу стратегии в исходной игре $u^{NE}(t, x^{NE})$ (4.322) (зеленая линия).

игре с динамическим обновлением информации имеет следующий вид:

$$\hat{x}^{NE}(t) = x_0 \left(1 - \frac{\Delta t}{\bar{T}}\right)^{jn} \left(\frac{t_0 + j\Delta t + \bar{T} - t}{\bar{T}}\right)^n, \quad t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]. \quad (4.337)$$

Доказательство Аналогично доказательству Теоремы 4.5.1. □

Позиционные равновесные по Нэшу стратегии вдоль равновесной траектории $\hat{x}^{NE}(t)$ на временном интервале $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, l$, имеют вид

$$u_j^{NE}(t, x_j^{NE}(t)) = x_0 \left(1 - \frac{\Delta t}{\bar{T}}\right)^{jn} \frac{(\bar{T} + j\Delta t + t_0 - t)^{n-1}}{\bar{T}^n}. \quad (4.338)$$

4.5.3 Игра с непрерывным обновлением информации

Рассмотрим случай, когда длина Δt временного интервала между моментами обновления информации пренебрежимо мала. Другими словами, информация обновляется непрерывно во времени ($\Delta t \rightarrow 0$ or $l \rightarrow \infty$).

В соответствии с процедурой, описанной в статье [78], на первом этапе рассмотрим следующую игру $\Gamma(x, t, t + \bar{T})$ с начальным t и конечным $t + \bar{T}$ моментами времени:

$$K_i^t(x, t, t + \bar{T}; u_1^t, \dots, u_n^t) = \int_t^{t+\bar{T}} \log(u_i^t(\tau, x_t)) d\tau, \quad (4.339)$$

при условии

$$\begin{aligned} \dot{x}_t &= - \sum_{i=1}^n u_i^t(\tau, x_t), \\ x_t(t) &= x, \\ x_t(t) &\geq 0, \end{aligned} \quad (4.340)$$

где x — некоторая начальная позиция. Предположим, что параметр t является фиксированной константой. Будем использовать игровую модель (4.339), (4.340), чтобы построить стратегии для игры с непрерывным обновлением информации.

Следуя процедуре из [78], построим уравнения Гамильтона–Якоби–Беллмана, чтобы найти кооперативные (некооперативные) стратегии, а именно оптимальные стратегии (равновесие по Нэшу в позиционных стратегиях). Для этого необходимо определить функцию Беллмана для данной игры:

$$\begin{aligned} V(t; \tau, x) &= \max_{u^t} \left\{ \sum_{i=1}^n K_i^t(x, t + \bar{T} - \tau; u_1^t, \dots, u_n^t) \right\} = \\ &= \max_{u^t} \left\{ n \int_{\tau}^{t+\bar{T}} \log(u^t(\tau, x_t)) d\tau \right\}, \quad \tau \in [t, t + \bar{T}] \end{aligned} \quad (4.341)$$

$$\begin{aligned} \left(V_i(t; \tau, x) &= K_i^t(x, t + \bar{T} - \tau; u_1^{t,NE}, \dots, u_n^{t,NE}) = \right. \\ &= \left. \int_{\tau}^{t+\bar{T}} \log(u_i^{t,NE}(\tau, x_t)) d\tau, \quad \tau \in [t, t + \bar{T}], \quad i \in N \right), \end{aligned} \quad (4.342)$$

где функция Беллмана $V(t; \tau, x)$ — максимальное значение суммарного выигрыша игроков в подыгре $\Gamma(x, \tau, t + \bar{T})$, начинающейся в момент времени $\tau \in [t, t + \bar{T}]$ из позиции x .

(Функция Беллмана $V_i(t; \tau, x)$ — выигрыш игрока $i \in N$ в равновесии по Нэшу в позиционных стратегиях $u^{t, NE}(\tau, x) = (u_1^{t, NE}(\tau, x), \dots, u_n^{t, NE}(\tau, x))$ в подыгре $\Gamma(x, \tau, t + \bar{T})$, начинающейся в момент времени t из позиции x .)

Используя подход, описанный в разделах 4.5.1.1 и 4.5.1.2, можно вывести соответствующее уравнение Гамильтона–Якоби–Беллмана для определения кооперативных стратегий (позиционных равновесных по Нэшу стратегий):

$$\begin{aligned} -V_\tau(t; \tau, x) &= \max_{u^t} \{n \log u^t - nu^t V_x(t; \tau, x)\}, \\ V(t; t + \bar{T}, x) &= 0 \end{aligned} \quad (4.343)$$

$$\left(\begin{aligned} -V_\tau^i(t; \tau, x) &= \max_{u^t} \{n \log u^t - nu^t V_x^i(t; \tau, x)\}, \\ V(t; t + \bar{T}, x) &= 0 \end{aligned} \right). \quad (4.344)$$

Метод решения (4.343) и (4.344) аналогичен тому, что использован для классических задач управления в разделах 4.5.1.1 и 4.5.1.2. В иллюстративных целях, представим решение только для кооперативной постановки (4.343), т.е. определим кооперативные стратегии $u^{t,*}(\tau, x)$.

Из необходимых условий оптимальности первого порядка для (4.343) находим

$$u^{t,*} = \frac{1}{V_x(t; \tau, x)}.$$

Подставляя это управление в уравнение (4.343), получаем:

$$\begin{aligned} V_t(t; \tau, x) &= n \log V_x(t, \tau, x) + n, \\ V(t; t + \bar{T}, x) &= 0. \end{aligned}$$

Определим функцию Беллмана в виде

$$V(t; \tau, x) = A(t, \tau) \log x + B(t, \tau).$$

Подстановка этого выражения в уравнение (4.343) дает

$$\begin{aligned} \dot{A}_\tau(t, \tau) \log x + \dot{B}_\tau(t, \tau) &= n \log A(t, \tau) - n \log x + n, \\ A(t, t + \bar{T}) &= B(t, t + \bar{T}) = 0. \end{aligned} \quad (4.345)$$

Решением системы (4.345) являются функции

$$\begin{aligned} A(t, \tau) &= n(t + \bar{T} - \tau), \\ B(t, \tau) &= -n(t + \bar{T} - \tau) \log n(t + \bar{T} - \tau), \quad \tau \in [t, t + \bar{T}]. \end{aligned} \quad (4.346)$$

Итак, функция Беллмана принимает вид

$$V(t; \tau, x) = n(t + \bar{T} - \tau) \log \frac{x}{n(t + \bar{T} - \tau)}, \quad \tau \in [t, t + \bar{T}], \quad (4.347)$$

а оптимальное управление — вид

$$u^{t,*}(\tau, x) = \frac{1}{V_x} = \frac{x}{n(t + \bar{T} - \tau)}, \quad \tau \in [t, t + \bar{T}]. \quad (4.348)$$

Управление или стратегия $u^{t,*}(\tau, x)$ (4.348) оптимальна в игре $\Gamma(x, t, t + \bar{T})$, определенной на временном интервале $[t, t + \bar{T}]$. Следовательно, при фиксированном t стратегия $u^{t,*}(\tau, x)$ (4.348) определена на том же временном интервале, т.е. $[t, t + \bar{T}]$. Однако, изменяя параметр t , мы меняем начальный момент игры $\Gamma(x, t, t + \bar{T})$ и, следовательно, автоматически меняем временной интервал, на котором рассчитано оптимальное управление (равновесие по Нэшу в позиционных стратегиях). Важно отметить, что оптимальное управление (4.348) явно зависит от параметра t , а также от параметров \bar{T} и τ .

Используя $u^{t,*}(\tau, x)$ (4.348), построим стратегию, которую будем применять для доказательства сходимости. Предположим, что в момент времени t в качестве результирующей стратегии в игре с непрерывным обновлением информации используется $\tilde{u}^*(t, x) = u^{t,*}(\tau, x)$, где $\tau = t$. Это означает, что в любой момент времени t игроки ориентируются (определяют свои оптимальные

стратегии) по информации на временном интервале $[t, t + \bar{T}]$.

Аналогичная процедура может быть проделана для равновесных стратегий, которые мы обозначим через $\tilde{u}^{NE}(t, x)$:

$$\tilde{u}^*(t, x) = \frac{x}{n\bar{T}} \quad \left(\tilde{u}^{NE}(t, x) = \frac{x}{\bar{T}} \right). \quad (4.349)$$

Далее, необходимо показать, что построенные таким образом стратегии являются стратегиями в игре с непрерывным обновлением информации, а именно в случае, когда временной интервал между моментами обновления информации стремится к нулю: $\Delta t \rightarrow 0$ ($l \rightarrow \infty$).

Теорема 4.5.3 *При $\Delta t \rightarrow 0$ ($l \rightarrow \infty$) результирующие стратегии $\hat{u}^*(t, x) = \hat{u}_i^*(t, x)$ (4.334) ($\hat{u}^{NE}(t, x) = \hat{u}_i^{NE}(t, x)$ (4.335)) в игре с динамическим обновлением информации равномерно сходятся к стратегиям $\tilde{u}^*(t, x)$ ($\tilde{u}^{NE}(t, x)$), которые являются результирующими в игре с непрерывным обновлением информации (4.349):*

$$\hat{u}_i^*(t, x) \xrightarrow{[t_0, T]} \tilde{u}^*(t, x) \quad \left(\hat{u}_i^{NE}(t, x) \xrightarrow{[t_0, T]} \tilde{u}^{NE}(t, x) \right). \quad (4.350)$$

Доказательство Из решения исходной игры, представленного в разделе 4.5.1 следует, что стратегия игрока i в некооперативной игре отличается от соответствующей стратегии только коэффициентом n . Поэтому при доказательстве мы рассмотрим только кооперативный случай.

Используем критерий равномерной сходимости последовательности функций $\hat{u}_i^*(t, x)$ к функции $\tilde{u}^*(t, x)$ при $l \rightarrow \infty$, где l — число моментов обновления информации:

$$\lim_{\Delta t \rightarrow 0} \sup_{t \in [t_0, T]} |\hat{u}_i^*(t, x) - \tilde{u}^*(t, x)| = \lim_{l \rightarrow \infty} \sup_{t \in [t_0, T]} |\hat{u}_i^*(t, x) - \tilde{u}^*(t, x)| = 0. \quad (4.351)$$

Предположим, что наибольшее значение разностной функции $\hat{u}_i^*(t, x) - \tilde{u}^*(t, x)$ (4.351) принадлежит интервалу $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$. Тогда необходимо

доказать следующее равенство:

$$\lim_{\Delta t \rightarrow 0} \sup_{t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]} |u_j^*(t, x) - \tilde{u}^*(t, x)| = 0. \quad (4.352)$$

Возможны три случая достижения наибольшего значения разностной функции $\hat{u}_j^*(t, x) - \tilde{u}^*(t, x)$; см. ниже.

- 1 Во внутренней точке интервала $(t_0 + j\Delta t, t_0 + (j+1)\Delta t)$; в этом случае максимум разностной функции (4.352) может быть найден с помощью необходимого условия оптимальности первого порядка:

$$\begin{aligned} \frac{d}{dt}(u_j^*(t, x) - \tilde{u}^*(t, x)) &= \frac{d}{dt} \left(\frac{x}{n(\bar{T} + j\Delta t + t_0 - t)} - \frac{x}{n\bar{T}} \right) = \\ &= \frac{x}{n(\bar{T} + j\Delta t + t_0 - t)^2}. \end{aligned}$$

Производная не равна нулю на всем интервале $(t_0 + j\Delta t, t_0 + (j+1)\Delta t)$, что говорит о строгой монотонности разностной функции на этом интервале.

- 2 В левой части интервала $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, т.е. в момент времени $t = t_0 + j\Delta t$:

$$\begin{aligned} &\lim_{\Delta t \rightarrow 0} |u_j^*(t_0 + j\Delta t, x) - \tilde{u}^*(t_0 + j\Delta t, x)| = \\ &= \lim_{\Delta t \rightarrow 0} \left| \frac{x}{n(\bar{T} + j\Delta t + t_0 - (t_0 + j\Delta t))} - \frac{x}{n\bar{T}} \right| = \lim_{\Delta t \rightarrow 0} \left| \frac{x}{n\bar{T}} - \frac{x}{n\bar{T}} \right| = 0. \end{aligned}$$

- 3 В правой части интервала $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, т.е. в момент времени $t = t_0 + (j+1)\Delta t$:

$$\begin{aligned} &\lim_{\Delta t \rightarrow 0} |u_j^*(t_0 + (j+1)\Delta t, x) - \tilde{u}^*(t_0 + (j+1)\Delta t, x)| = \\ &= \lim_{\Delta t \rightarrow 0} \left| \frac{x}{n(\bar{T} + j\Delta t + t_0 - (t_0 + (j+1)\Delta t))} - \frac{x}{n\bar{T}} \right| = \\ &\quad \lim_{\Delta t \rightarrow 0} \left| \frac{x}{n(\bar{T} - \Delta t)} - \frac{x}{n\bar{T}} \right| = 0. \end{aligned}$$

Значение разностной функции (4.352) внутри интервала не превышает

значений на его концах, а слева и справа значение разностной функции стремится к нулю и, как следствие, значение на интервале также стремится к нулю.

В соответствии с критерием (4.351) последовательность результирующих управлений $\hat{u}^*(t, x) = \hat{u}_l^*(t, x)$ с динамическим обновлением информации равномерно сходится к управлению $\tilde{u}^*(t, x)$ с непрерывным обновлением информации при $l \rightarrow \infty$. \square

Теорема 4.5.4 При $\Delta t \rightarrow 0$ ($l \rightarrow \infty$) кооперативная траектория в игре с динамическим обновлением информации $\hat{x}^*(t) = \hat{x}_l^*(t)$ равномерно сходится к траектории в игре с непрерывным обновлением информации $\tilde{x}^*(t)$.

Доказательство Заметим, что

$$\begin{aligned}\hat{u}^*(t, x) &= \hat{u}_l^*(t, x) = x\hat{u}^*(t, 1), \\ \tilde{u}^*(t, x) &= x\tilde{u}^*(t, 1).\end{aligned}$$

Из уравнений движения для случая динамического обновления информации получаем:

$$\begin{aligned}\frac{d\hat{x}_l^*}{dt}(t) &= -nu^*(t, x) = -n\hat{x}_l^*\hat{u}^*(t, 1), \\ \log \hat{x}_l^*(t) &= -n \int_{t_0}^t \hat{u}^*(\tau, 1) d\tau.\end{aligned}$$

При непрерывном обновлении информации имеем:

$$\begin{aligned}\frac{d\tilde{x}^*}{dt}(t) &= -nu^*(t, x) = -n\tilde{x}^*(t)\tilde{u}^*(t, 1), \\ \log \tilde{x}^*(t) &= -n \int_{t_0}^t \tilde{u}^*(\tau, 1) d\tau.\end{aligned}$$

Из свойств равномерной сходимости стратегий следует:

$$\begin{aligned} \hat{u}_i^*(t, x) \underset{[t_0, T]}{\rightrightarrows} \tilde{u}^*(t, x) &\implies n \int_{t_0}^t \hat{u}_i^*(\tau, x) d\tau \underset{[t_0, T]}{\rightrightarrows} n \int_{t_0}^t \tilde{u}^*(\tau, x) d\tau \implies \\ &\implies \exp \left\{ -n \int_{t_0}^t \hat{u}_i^*(\tau, x) d\tau \right\} \underset{[t_0, T]}{\rightrightarrows} \exp \left\{ -n \int_{t_0}^t \tilde{u}^*(\tau, x) d\tau \right\}. \end{aligned} \quad (4.353)$$

Поскольку функция e^{-s} ограничена на интервале $s \in [0, \infty)$, получаем:

$$\hat{x}_i^*(t, x) \underset{[t_0, T]}{\rightrightarrows} \tilde{x}^*(t, x). \quad (4.354)$$

□

Построим результирующую траекторию с непрерывным обновлением информации, подставив $\tilde{u}^*(t, x)$ ($\tilde{u}^{NE}(t, x)$) в уравнения движения (4.310):

$$\begin{aligned} \dot{x} &= -\frac{x}{T}, & \left(\dot{x} &= -n\frac{x}{T} \right) \\ x(t_0) &= x_0 & \left(x(t_0) &= x_0 \right). \end{aligned} \quad (4.355)$$

Тогда соответствующие решения — это траектории в игре с непрерывным обновлением информации:

$$\tilde{x}^*(t) = x_0 e^{-\frac{t-t_0}{T}} \quad \left(\tilde{x}^{NE}(t) = x_0 e^{-n\frac{t-t_0}{T}} \right). \quad (4.356)$$

Оптимальные стратегии (позиционные равновесные по Нэшу стратегии) вдоль результирующей траектории в игре с непрерывным обновлением информации имеют следующий вид:

$$\tilde{u}^*(t, \tilde{x}^*(t)) = \frac{x_0}{nT} e^{-\frac{t-t_0}{T}} \quad \left(\tilde{u}^{NE}(t, \tilde{x}^{NE}(t)) = \frac{x_0}{T} e^{-n\frac{t-t_0}{T}} \right). \quad (4.357)$$

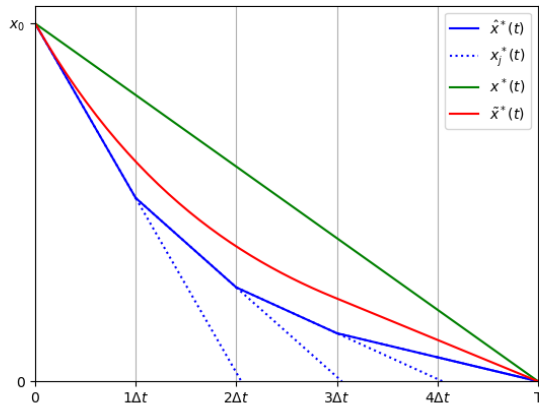


Рис. 4.39: Результирующая кооперативная траектория $\tilde{x}^*(t)$ (4.356) с непрерывным обновлением информации (красная линия), результирующая кооперативная траектория $\hat{x}^*(t)$ (4.332) с динамическим обновлением информации (синяя линия) и кооперативная траектория в исходной игре $x^*(t)$ (4.321) (зеленая линия).

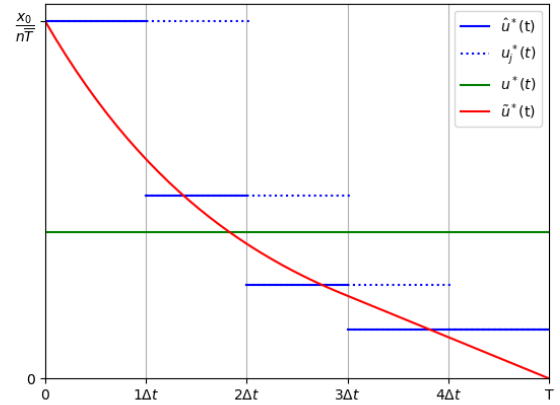


Рис. 4.40: Результирующие кооперативные стратегии $\tilde{u}^*(t)$ (4.349) с непрерывным обновлением информации (красная линия), результирующие кооперативные стратегии $\hat{u}^*(t)$ (4.334) с динамическим обновлением информации (синяя линия) и кооперативные стратегии в исходной игре $u^*(t)$ (4.322) (зеленая линия).

4.5.4 Численное моделирование

Рассмотрим результаты численного моделирования для игры с тремя симметричными игроками ($n = 3$) на временном интервале $[0, 100]$, т.е. $t_0 = 0$ и $T = 100$. В начальный момент времени $t_0 = 0$ запас ресурса равен 2000, т.е. $x_0 = 2000$. Предположим, что при динамическом обновлении информации (синие сплошные и пунктирные линии, см. рис. 4.39–4.42) интервалы между моментами обновления информации равны $\Delta t = 20$, и следовательно $l = 5$. На рисунках 4.39 и 4.41 представлено сравнение результирующих стратегий в исходной игре заданной продолжительности (красные линии), игре с динамическим обновлением информации (синяя линия) и игре с непрерывным обновлением информации (красные линии) для кооперативного и некооперативного случаев. На рисунках 4.40 и 4.42 представлены аналогичные результаты для соответствующих стратегий.

Для иллюстрации Теорем 4.1 и 4.2 о сходимости результирующих стратегий

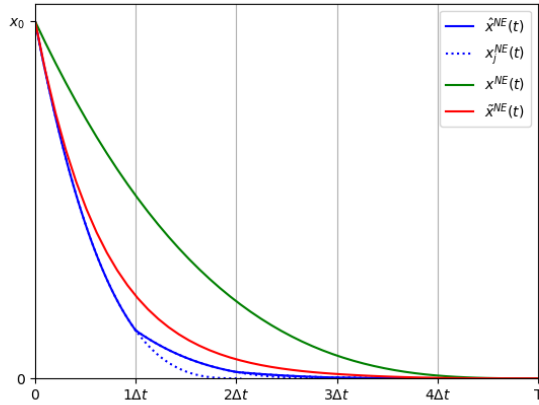


Рис. 4.41: Результирующая равновесная по Нэшу траектория $\tilde{x}^{NE}(t)$ (4.356) с непрерывным обновлением информации (красная линия), результирующая равновесная траектория $\hat{x}^{NE}(t)$ (4.337) с динамическим обновлением информации (синяя линия) и равновесная по Нэшу траектория в исходной игре $x^{NE}(t)$ (4.328) (зеленая линия).

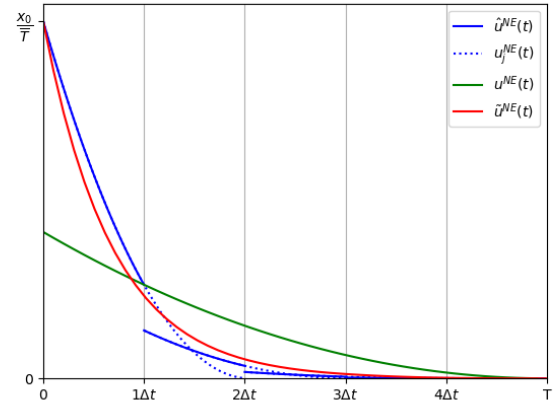


Рис. 4.42: Результирующие равновесные по Нэшу стратегии $\tilde{u}^{NE}(t)$ (4.349) с непрерывным обновлением информации (красная линия), результирующие равновесные по Нэшу стратегии $\hat{u}^{NE}(t)$ (4.338) с динамическим обновлением информации (синяя линия) и равновесные по Нэшу стратегии в исходной игре $u^{NE}(t)$ (4.329) (зеленая линия).

и соответствующей траектории, рассмотрим результаты моделирования для случая частого обновления информации, а именно $l = 50$. На рисунках 4.43-4.46 показаны те же решения, что и на рисунках 4.39-4.42, но для случая $\Delta t = 2$. Таким образом, сходимость стратегий и траектории подтверждается численными экспериментами, представленными ниже.

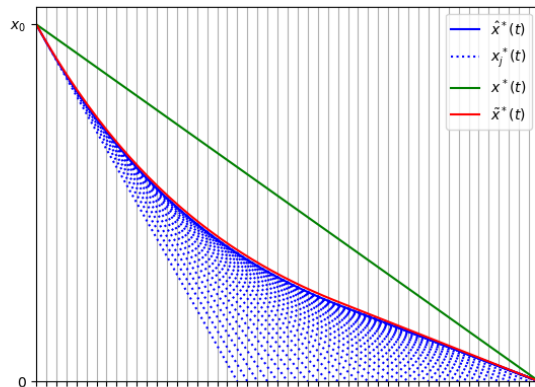


Рис. 4.43: $\tilde{x}^*(t)$ (4.356) (красная линия), $\hat{x}^*(t)$ (4.332) (синяя линия) и $x^*(t)$ (4.321) (зеленая линия).

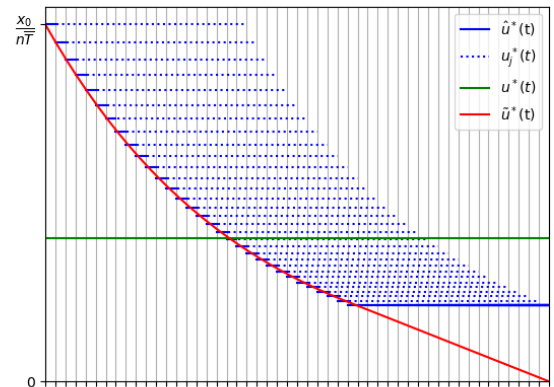
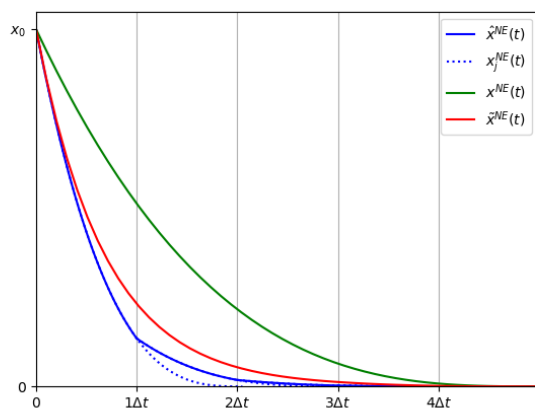
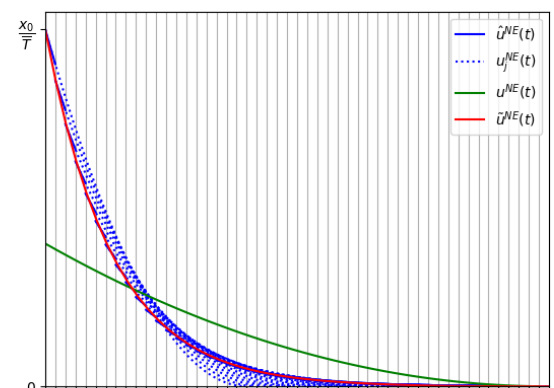


Рис. 4.44: $\tilde{u}^*(t)$ (4.349) (красная линия), $\hat{u}^*(t)$ (4.334) (синяя линия) и $u^*(t)$ (4.322) (зеленая линия).



labelformat=empty

Рис. 4.45: $\tilde{x}^{NE}(t)$ (4.356) (красная линия), $\hat{x}^{NE}(t)$ (4.337) (синяя линия) и $x^{NE}(t)$ (4.328) (зеленая линия).



labelformat=empty

Рис. 4.46: $\tilde{u}^{NE}(t)$ (4.349) (красная линия), $\hat{u}^{NE}(t)$ (4.338) (синяя линия) и $u^{NE}(t)$ (4.329) (зеленая линия).

Глава 5

Обратная задача теории управления с непрерывным обновлением информации

В теории оптимального управления значительное внимание уделяется моделированию поведения людей и процесса принятия решений людьми. Большинство подходов основаны на фиксированном (или бесконечном) временном горизонте в предположении, что вся информация доступна в начале временного интервала. Тем не менее, есть основания считать, что в каждый момент времени человек использует информацию, определенную на непрерывно движущемся информационном горизонте, и соответствующим образом адаптируется к ней. Поэтому в Главе 5 предлагается подход к оптимальному управлению с обратной связью, основанный на парадигме непрерывного обновления информации. Параметры модели, определяющие индивидуальное поведение человека, состоят из параметров функции затрат и длины информационного горизонта. Значения данных параметров могут быть найдены с помощью соответствующего подхода обратного оптимального управления. Применимость указанного подхода продемонстрирована на примере потенциального приложения — задачи идентификации человеческого поведения в системе помощи водителю.

5.1 Оптимальное управление с непрерывным обновлением информации

В этом разделе представлены результаты относительно оптимального управления с непрерывным обновлением информации. Прежде чем описать наш подход с непрерывным обновлением информации, напомним классическую постановку задачи оптимального управления.

5.1.1 Исходная задача оптимального управления

Рассмотрим следующую задачу оптимального управления, заданную на временном интервале $[t_0, T]$:

$$J(x_0, t_0; u) = \int_{t_0}^T g[t, x(t), u(t, x)] dt \rightarrow \min_u \quad (5.1)$$

при условии

$$\begin{aligned} \dot{x}(t) &= f(t, x, u), \\ x(t_0) &= x_0, \\ x &\in \text{comp}\mathbb{R}^n, u = u(t, x) \in U \subset \text{comp}\mathbb{R}^m, t \in [t_0, T], \end{aligned} \quad (5.2)$$

где $\text{comp}\mathbb{R}^m$ — компактное множество в m -мерном пространстве вещественных чисел; функции $g[t, x(t), u(t, x(t))]$ и $f(t, x, u)$ являются интегрируемыми; $x(t) \in \mathbb{R}^n$ обозначает решение задачи Коши (5.2) при фиксированном $u(t, x) \in \mathbb{R}^m$. Управление $u(t, x) \in \mathbb{R}^m$ называется допустимым, если задача (5.2) имеет единственное непрерывное решение.

5.1.2 Задача оптимального управления с непрерывным обновлением информации

Используя исходную задачу оптимального управления, заданную на временном отрезке $[t_0, T]$, построим соответствующую задачу оптимального управления с непрерывным обновлением информации.

Рассмотрим следующую задачу оптимального управления, заданную на временном интервале $[t, t + \bar{T}]$, $0 < \bar{T} < +\infty$:

$$J(x, t; u^t) = \int_t^{t+\bar{T}} g[s, x_t(s), u^t(s, x_t)] ds \rightarrow \min_{u^t} \quad (5.3)$$

при условии

$$\begin{aligned} \dot{x}_t(s) &= f(s, x_t, u^t), \\ x_t(t) &= x, \\ x_t &\in \text{comp} \mathbb{R}^n, u^t = u^t(s, x_t) \in U \subset \text{comp} \mathbb{R}^m, s \in [t, t + \bar{T}], \end{aligned} \quad (5.4)$$

где $u^t(s, x_t) \in \mathbb{R}^m$ и $x_t(s) \in \mathbb{R}^n$ — соответственно оптимальное управление и оптимальная траектория системы на временном интервале $[t, t + \bar{T}]$.

Для задачи оптимального управления с непрерывным обновлением информации характерны следующие особенности: *текущее время* $t \in [t_0, +\infty)$ непрерывно изменяется, и в результате человек непрерывно получает новую информацию об уравнениях движения и целевой функции на временном интервале $[t, t + \bar{T}]$.

Управление $u(t, x)$ в задаче оптимального управления с непрерывным обновлением информации имеет вид

$$u(t, x) = u^t(s, x)|_{s=t}, \quad t \in [t_0, +\infty), \quad (5.5)$$

где $u^t(s, x)$, $s \in [t, t + \bar{T}]$, — управление в задаче, определенной на временном интервале $[t, t + \bar{T}]$, а $u^t(s, x)|_{s=t}$ — часть этого управления в начальный момент времени $s = t$. Основная идея (5.5) заключается в следующем: по мере продвижения текущего времени t , информация обновляется; следовательно, для моделирования поведения человека необходимо учитывать управление $u^t(s, x)$ только в точках $s = t$.

Траектория $x(t)$ в задаче оптимального управления с непрерывным обновлением информации определяется в соответствии с динамикой системы (5.2) при управлении $u = u(t, x)$ (5.5). Предполагаем, что управление с непрерывным обновлением информации, полученное с помощью (5.5), является допустимым.

Задача управления с непрерывным обновлением информации существенно отличается от классической задачи оптимального управления, определенной на временном отрезке. В исходной задаче оптимального управления, лицо, принимающее решение, руководствуется значением целевой функции, которое будет в конечном итоге получено на временном интервале $[t_0, T]$. В задаче управления с непрерывным обновлением информации в момент времени t система ориентируется на ожидаемое значение целевой функции (5.3), которое вычисляется с использованием информации на временном интервале $[t, t + \bar{T}]$ (информации, доступной системе в момент времени t).

5.1.3 Оптимальное управление с непрерывным обновлением информации

Для модели с непрерывно обновляемой информацией будем использовать понятие оптимального управления с обратной связью $u^*(t, x)$. Кроме того, потребуем, чтобы для любого фиксированного текущего момента времени $t \in [t_0, +\infty)$, величина $u^*(t, x)$ совпадала с оптимальным управлением в момент времени t в задаче (5.3), (5.4), определенной на временном интервале $[t, t + \bar{T}]$.

Однако прямое применение классических подходов к оптимальному управлению с обратной связью невозможно. Рассмотрим другой текущий момент времени $t + \epsilon$, где $\epsilon \ll \bar{T}$. Согласно вышеупомянутому требованию, величина $u^*(t + \epsilon, x)$ в момент времени $t + \epsilon$ должна совпадать с оптимальным управлением в задаче, заданной на временном интервале $[t + \epsilon, t + \epsilon + \bar{T}]$.

Следовательно, $u^*(t, x)$, $t \in [t_0, +\infty)$, следует определять как бесконечную комбинацию оптимальных управлений на временных интервалах $[t, t + \bar{T}]$ для каждого текущего момента времени $t \in [t_0, +\infty)$.

Для построения таких управлений рассмотрим понятие обобщенного оптимального управления с обратной связью

$$\tilde{u}^*(t, x; s, x), \quad t \in [t_0, +\infty), \quad s \in [t, t + \bar{T}], \quad (5.6)$$

которое будет использовано в дальнейшем для построения управления $u^*(t, x)$.

Определение 5.1.0.1 Управление $\tilde{u}^*(t, x; s, x)$ называется обобщенным оптимальным управлением в задаче с непрерывным обновлением информации, если оно оптимально для любого фиксированного $t \in [t_0, +\infty)$ в задаче с целевой функцией (5.3) и динамикой системы (5.4).

С помощью обобщенного оптимального управления можно определить концепцию решения для задачи оптимального управления с непрерывным обновлением информации.

Определение 5.1.0.2 Управление $u^*(t, x)$ называется оптимальным управлением с непрерывным обновлением информации, если

$$u^*(t, x) = \tilde{u}^*(t, x; s, x)|_{s=t}, \quad t \in [t_0, +\infty), \quad (5.7)$$

где $\tilde{u}^*(t, x; s, x)$ — обобщенное оптимальное управление в смысле Определения 5.1.0.1.

5.2 Линейно-квадратичное оптимальное управление с непрерывным обновлением информации

В этом разделе представлено решение задачи оптимального управления с непрерывным обновлением информации согласно Определению 5.1.0.2 для линейно-квадратичного случая.

5.2.1 Постановка задачи для линейно-квадратичного случая с непрерывным обновлением информации

В линейно-квадратичном случае, общая постановка задачи оптимального управления с непрерывным обновлением информации принимает следующий вид.

Рассмотрим задачу оптимального управления, заданную на временном

интервале $[t, t + \bar{T}]$, где $0 < \bar{T} < +\infty$:

$$J(x, t; u^t) = \int_t^{t+\bar{T}} x_t^T(s) Q x_t(s) + u^t(s, x_t)^T R u^t(s, x_t) ds \rightarrow \min_{u^t} \quad (5.8)$$

при условии

$$\begin{aligned} \dot{x}_t(s) &= Ax_t(s) + Bu^t(s, x_t), \\ x_t(t) &= x, \\ x_t &\in \mathbb{R}^n, u^t = u^t(s, x_t) \in U \subset \text{comp} \mathbb{R}^m, s \in [t, t + \bar{T}]. \end{aligned} \quad (5.9)$$

Определения обобщенного оптимального управления и оптимального управления с непрерывным обновлением информации аналогичны Определениям 5.1.0.1 и 5.1.0.2.

5.2.2 Оптимальное управление с непрерывным обновлением информации

Ниже представлен явный вид оптимального управления с непрерывным обновлением информации на основе системы дифференциальных уравнений Риккати.

Теорема 5.2.1 *При любом начальном состоянии x_0 , задача линейно-квадратичного оптимального управления с непрерывным обновлением информации (5.8), (5.9) имеет решение тогда и только тогда, когда дифференциальное уравнение Риккати (5.10)*

$$\begin{aligned} \dot{K}(\tau) &= -\bar{T}A^T K(\tau) - \bar{T}K(\tau)A + \bar{T}K(\tau)SK(\tau) - \bar{T}Q, \\ K(1) &= 0, \end{aligned} \quad (5.10)$$

где $S = BR^{-1}B^T$, матрицы Q и R симметричны, а матрица R положительно определена, имеет симметричное решение $K(\cdot)$ на временном интервале $[0, 1]$.

Если данная задача с непрерывным обновлением информации имеет решение, то оно единственное и имеет вид обратной связи

$$u^*(t, x) = -R^{-1}B^TK(0)x. \quad (5.11)$$

Доказательство Введем следующую замену переменных:

$$\begin{aligned} s &= t + \bar{T}\tau, \\ y_t(\tau) &= x(t + \bar{T}\tau), \\ v_t(\tau, y_t) &= u(t + \bar{T}\tau, x). \end{aligned} \quad (5.12)$$

Поставляя формулу (5.12) в уравнения движения (5.9) и целевую функцию (5.8), получаем:

$$\dot{y}_t(\tau) = \bar{T}Ay_t(\tau) + \bar{T}Bv_t(\tau, y_t) \quad (5.13)$$

и

$$J(y_t, t; v_t) = \int_0^1 \bar{T}y_t^T(\tau)Q_iy_t(\tau) + \bar{T}(v_t(\tau, y_t))^TRv_t(\tau, y_t)d\tau. \quad (5.14)$$

Известно (например, см. [224, Theorem 5.1]), что критерием существования оптимального управления с обратной связью является наличие симметричного решения для системы дифференциальных уравнений (5.10). Оптимальное управление имеет вид

$$v_t^*(\tau, y_t) = -R^{-1}B^TK(\tau)y_t. \quad (5.15)$$

Из (5.12) находим

$$\tau = \frac{s-t}{\bar{T}}.$$

Возвращаясь к исходным переменным, получаем следующее управление:

$$u^t(s, x) = -R^{-1}B^TK\left(\frac{s-t}{\bar{T}}\right)x.$$

По построению оно является оптимальным в задаче, определенной на временном интервале $[t, t + \bar{T}]$. Уравнения движения (5.13), целевая функция (5.14) и решение (5.15) имеют одинаковый вид для всех значений t . Тогда обобщенное оптимальное управление в задаче управления с непрерывным обновлением

информации рассчитывается по формуле

$$\tilde{u}^*(t, s, x) = -R^{-1}B^TK \left(\frac{s-t}{\bar{T}} \right) x. \quad (5.16)$$

Используя (5.7), при $s = t$ будем считать, что оптимальное управление с непрерывным обновлением информации равно обобщенному оптимальному управлению (5.16):

$$u^*(t, x) = -R^{-1}B^TK(0)x, \quad t \in [t_0, +\infty). \quad (5.17)$$

Таким образом, управление (5.15) в подзадаче существует при любом начальном состоянии y_t ($t \geq t_0$). Следовательно, управление (5.17) в задаче оптимального управления с непрерывным обновлением информации существует при любых x_0 . \square

Используя формулу (5.11) для оптимального управления с непрерывным обновлением информации $u^*(t, x)$, можно предложить следующую постановку обратной задачи оптимального управления с непрерывным обновлением информации (см. ниже).

5.3 Обратная задача оптимального управления с непрерывным обновлением информации

Предположим, что функция $g[t, x(t), u(t, x(t))]$ в (5.1) параметризована. Следовательно, в обратной задаче оптимального управления эта параметризация неизвестна и должна оцениваться по наблюдениям управления и вектора состояния. Кроме того, неизвестно значение информационного горизонта \bar{T} . Обозначим через θ набор неизвестных параметров, включая параметризацию целевой функции и информационный горизонт. Обратная задача управления с непрерывным обновлением информации включает в себя оценку набора неизвестных параметров θ по наблюдениям управления $\hat{u}(t)$ и соответствующей траектории $\hat{x}(t)$ на временном интервале $[t_0, T]$. Необходимо получить оценки оптимального управления $u_\theta^*(t, x_\theta^*(t))$ и соответствующей траектории $x_\theta^*(t)$,

при которых разность между модельными и наблюдаемыми траекториями минимальна.

Поэтому обратная задача оптимального управления с непрерывным обновлением информации решается ниже с помощью подхода, аналогичного стандартному двухуровневому подходу к задачам оптимального управления, представленному [235]. В указанной статье этот подход успешно применен к экспериментальным данным. Параметры модели θ определяются исходя из минимизации квадрата ошибки между наблюдаемыми и модельными траекториями:

$$Div = \int_{t_0}^T \|x_\theta(t) - \hat{x}(t)\|^2 + \|u_\theta(t, x_\theta^*(t)) - \hat{u}(t)\|^2 dt \rightarrow \min_{\theta}, \quad (5.18)$$

где x_θ и u_θ — оптимальные траектория и управление, соответствующие минимуму целевой функции (5.1) с параметрами θ . Обращаем внимание, что в рамках подхода обратного оптимального управления с непрерывным обновлением информации, величина $u_\theta(t, x_\theta(t))$ определяется как в (5.7), а траектория $x_\theta(t)$ представляет собой решение соответствующего уравнения динамики. Поэтому ниже рассмотрим линейно-квадратичную задачу оптимального управления с непрерывным обновлением информации, используя управление (5.11) для решения (5.18).

5.4 Численное моделирование

В этом разделе продемонстрируем предлагаемый нами подход к оптимальному управлению с непрерывным обновлением информации; в частности, как с помощью соответствующего решения обратной задачи можно определить параметры для описания поведения водителя, управляющего поперечным движением транспортного средства.

5.4.1 Модель рулевого управления одноколейным ТС

Рассмотрим модель рулевого управления одноколейным транспортным средством (ТС) с шестью параметрами состояния; подробное описание можно найти в [236]. Модель представляет собой экспериментально подтвержденную линейную аппроксимацию рулевого управления легковым автомобилем среднего размера.

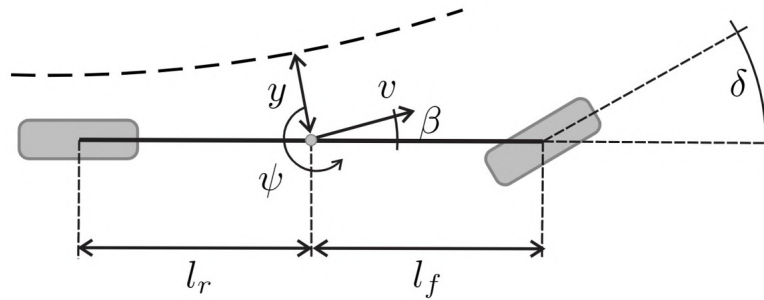


Рис. 5.1: Геометрические соотношения модели одноколейного транспортного средства.

Вектор состояния имеет вид

$$x(t) = \left[\beta(t) \quad \psi(t) \quad \dot{\psi}(t) \quad y(t) \quad \delta(t) \quad \dot{\delta}(t) \right]^T,$$

где $\beta(t)$ — угол бокового скольжения, $\psi(t)$ — угол рыскания, $\dot{\psi}(t)$ — угловая скорость рыскания, $y(t)$ — поперечное расстояние от средней полосы до ТС, $\delta(t)$ — угол поворота рулевого колеса, а $\dot{\delta}(t)$ — угловая скорость рулевого колеса. Значения всех параметров состояния указаны в единицах СИ. Управление $u(t, x) \in \mathbb{R}$ в данной системе — крутящий момент, приложенный водителем.

Динамика системы описывается дифференциальным уравнением

$$\dot{x}(t) = A x(t) + B u(t) \tag{5.19}$$

с заданным начальным состоянием $x(t_0) = x_0$ и

$$A = \begin{bmatrix} \frac{-C_f - C_r}{Mv} & \frac{C_r l_r - Mv^2 - C_f l_f}{Mv^2} & 0 & 0 & \frac{C_f}{Mv i_s} & 0 \\ \frac{C_r l_r - C_f l_f}{J_z} & \frac{-C_f l_f^2 - C_r l_r^2}{J_z v} & 0 & 0 & \frac{C_f l_f}{J_z i_s} & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ v & 0 & v & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \frac{-C_s}{J_s} & \frac{-D_s}{J_s} \end{bmatrix}, \quad (5.20)$$

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \frac{1}{J_s} \end{bmatrix}^T.$$

Геометрические соотношения модели схематически изображены на рисунке 5.1. Параметры системы представлены в Таблице 5.1.

Таблица 5.1: Параметры линейной модели рулевого управления [236]

Обозначение	Значение	Ед. измерения	Описание
v	20	$\frac{\text{м}}{\text{с}}$	скорость ТС,
l_f	1.5	м	расстояние от центра масс, до передней колесной базы,
l_r	1	м	расстояние от центра масс, до задней колесной базы,
C_f	137.5	$\frac{\text{кН}}{\text{рад}}$	проскальзывание передних колес,
C_r	137.5	$\frac{\text{кН}}{\text{рад}}$	проскальзывание задних колес,
M	1500	кг	масса ТС,
J_z	1800	кг м ²	инерция рысканья ТС,
D_s	1	$\frac{\text{Нм с}}{\text{рад}}$	постоянная демпфирования рулевого колеса,
C_s	1	$\frac{\text{Нм}}{\text{рад}}$	постоянная втягивания рулевого колеса,
J_s	0.1	$\frac{\text{Нм с}^2}{\text{рад}}$	инерция рулевого колеса,
i_s	16	—	постоянная передачи рулевого управления

Следуя, например, [237], будем считать, что целевая функция водителя имеет квадратичную форму. Поэтому подынтегральное выражение в (5.1) и

(5.8) запишем в виде

$$g[t, x(t), u(t, x), \theta] = x^T(t)Qx(t) + u(t, x)^T Ru(t, x). \quad (5.21)$$

Необходимо провести идентификацию набора неизвестных параметров θ по сгенерированным данным с помощью подхода (5.18). Набор θ состоит из элементов матриц Q и R , а также временного горизонта \bar{T} :

$$\theta = \{Q, R, \bar{T}\}. \quad (5.22)$$

5.4.2 Оптимальное управление с непрерывным обновлением информации

Чтобы продемонстрировать различные решения, возникающие при оптимальном управлении с непрерывным обновлением информации, сначала определим достоверные параметры целевой функции человека-водителя в виде:

$$Q = \text{diag}(1, 10, 1, 40, 10, 1), \quad R = [1]. \quad (5.23)$$

Пусть начальное состояние в момент времени $t_0 = 0$ равно

$$x_0 = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0.1 \end{bmatrix}.$$

Применим оптимальное управление с непрерывным обновлением информации (Теорему 5.2.1), чтобы сгенерировать наборы данных, имитирующие поведение человека при рулевом управлении. Для этого с помощью решателя `ode45` среды MATLAB были найдены численные решения уравнения Риккати (5.10). Полученные наборы данных D1–D4 состоят из оптимальных траекторий системы и управления на временном интервале $t \in [0, 10]$ и соответствуют следующим значениям информационного горизонта:

$$\bar{T} = 1, \quad \bar{T} = 1.5, \quad \bar{T} = 2, \quad \bar{T} = 2.5. \quad (5.24)$$

Все наборы данных сгенерированы на основе одних и тех же параметров целевой функции (5.23) и системной динамики (5.20). В качестве примера

на рисунках 5.2, 5.3 и 5.4 показаны соответственно оптимальные управления, поперечное расстояние от средней полосы до ТС и угол поворота рулевого колеса.

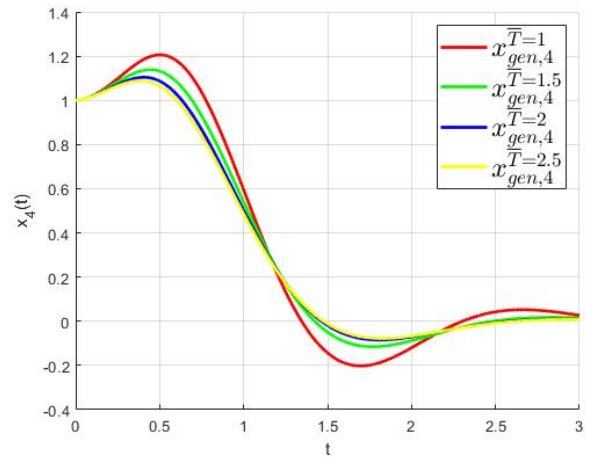
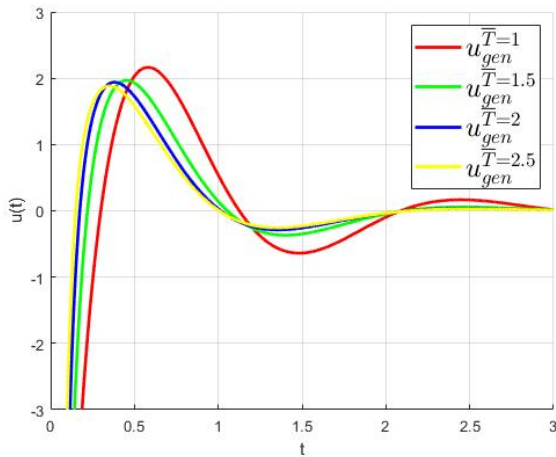


Рис. 5.2: Оптимальные управления с непрерывным обновлением информации при различных значениях информационного горизонта.

Рис. 5.3: $y(t)$ в задаче оптимального управления с непрерывным обновлением информации при различных значениях информационного горизонта.

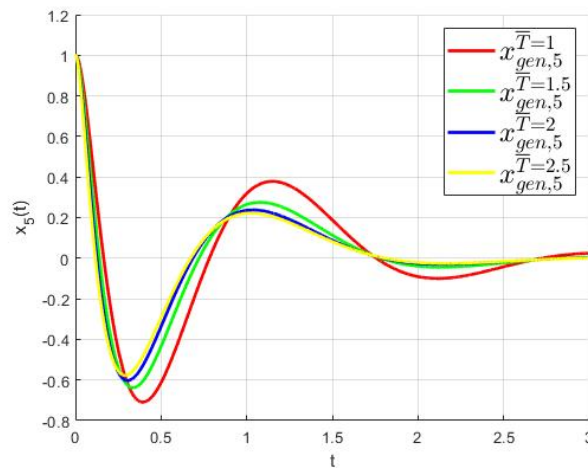


Рис. 5.4: $\delta(t)$ в задаче оптимального управления с непрерывным обновлением информации при различных значениях информационного горизонта.

5.4.3 Обратная задача оптимального управления

Теперь решим обратную задачу оптимального управления (5.18), применяя подход с непрерывным обновлением информации, чтобы оценить θ ; см. (5.22). Используем набор данных D2, т.е. оптимальные управления с непрерывным обновлением информации с $\bar{T} = 1.5$ в качестве наблюдаемых траекторий $\hat{x}(t)$ и $\hat{u}(t)$ на временном интервале $[0, 10]$. Чтобы показать важность идентификации \bar{T} помимо матриц целевой функции, представим оптимальные матрицы $[Q, R]$ для различных фиксированных значений \bar{T} , определенных в (5.24). Кроме того, представим соответствующие значения целевой функции Div (5.18). Во избежание неоднозначности параметров, все параметры целевой затрат нормированы относительно матрицы $R = [1.00]$.

$$\begin{aligned}
 Div^{\bar{T}=1} &= 7.7944 \cdot 10^{-9}, R = 1.00 \\
 Q^{\bar{T}=1} &= \text{diag}(17.11, 38.56, 727.49, 91.64, 65.98, 2.50), \\
 Div^{\bar{T}=1.5} &= 5.4986 \cdot 10^{-13}, R = 1.00 \\
 Q^{\bar{T}=1.5} &= \text{diag}(11.95, 9.91, 1.00, 40.00, 10.01, 1.00), \\
 Div^{\bar{T}=2} &= 0.017448, R = 1.00 \\
 Q^{\bar{T}=2} &= \text{diag}(0.00, 0.00, 0.00, 18.54, 0.00, 0.34), \\
 Div^{\bar{T}=2.5} &= 0.11415, R = 1.00 \\
 Q^{\bar{T}=2.5} &= \text{diag}(0.00, 0.00, 0.00, 9.79, 0.00, 0.02).
 \end{aligned} \tag{5.25}$$

Подход обратного оптимального управления описан в разделе 5.3, где также проведена идентификация информационного горизонта \bar{T} . Решение (5.18) найдено методом последовательного квадратичного программирования (SQP) с помощью решателя `fmincon` среды MATLAB. Значения всех параметров ограничены положительной полуосью. Результаты идентификации представлены ниже.

$$\begin{aligned}
 \bar{T}_{opt} &= 1.4815 \\
 Div^{\bar{T}_{opt}} &= 1.3065 \cdot 10^{-10}, R = 1.00, \\
 Q^{\bar{T}_{opt}} &= \text{diag}(177.12, 8.82, 9.84, 41.26, 10.99, 1.03).
 \end{aligned} \tag{5.26}$$

На основе идентифицированных параметров $Q^{\bar{T}_{opt}}$ и \bar{T}_{opt} сгенерированы траектории $u_{id}^{\bar{T}_{opt}}$ и $x_{id}^{\bar{T}_{opt}}$, соответствующие оптимальному управлению с непрерывным обновлением информации. На рисунке 5.5 показано исходное оптимальное управление (из набора данных D2, черная линия) и оптимальное управление с непрерывным обновлением информации на основе идентифицированных параметров (желтая линия). На рисунках 5.6 и 5.7 показаны достоверные и идентифицированные траектории модели: поперечное расстояние $y(t)$ от средней полосы до ТС и угол поворота рулевого колеса $\delta(t)$.

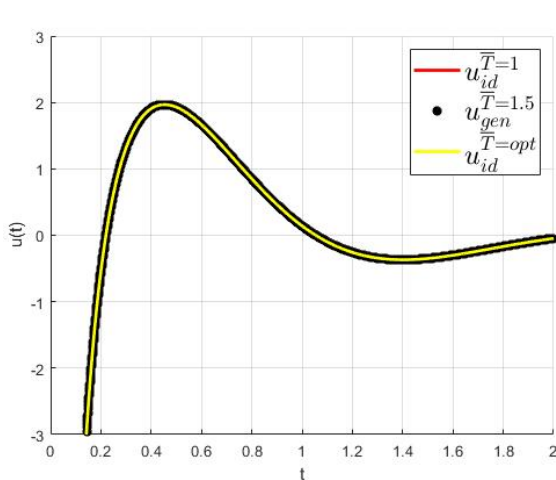


Рис. 5.5: Достоверное управление (синяя линия) и идентифицированное оптимальное управление с непрерывным обновлением информации (желтая линия), сгенерированное при $\bar{T}_{opt} = 1.4815$ и $Q^{\bar{T}_{opt}}$.

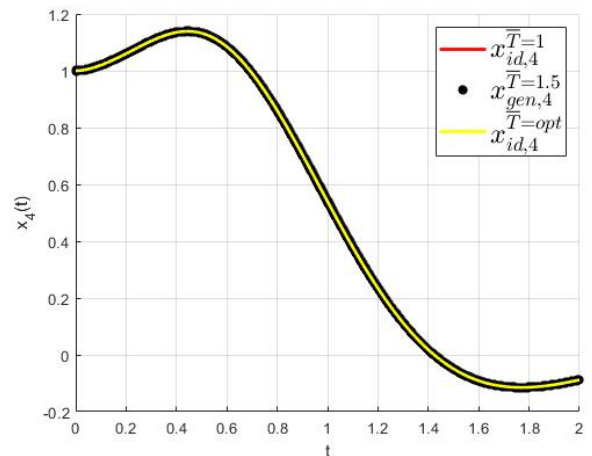


Рис. 5.6: Достоверная функция $y(t)$ (синяя линия) и идентифицированное оптимальное управление с непрерывным обновлением информации (желтая линия), сгенерированное при $\bar{T}_{opt} = 1.4815$ и $Q^{\bar{T}_{opt}}$.

5.4.4 Обсуждение

Очевидно (см. рисунок 5.2), что оптимальные траектории управления различаются в зависимости от предполагаемой величины информационного горизонта \bar{T} . Кроме того, как показано на рисунках 5.3 и 5.4, больше информации приводит к более эффективной стабилизации состояний.

Решения в (5.25) указывают на следующий факт. В предположении, что человек действует в соответствии с движущимся информационным горизон-

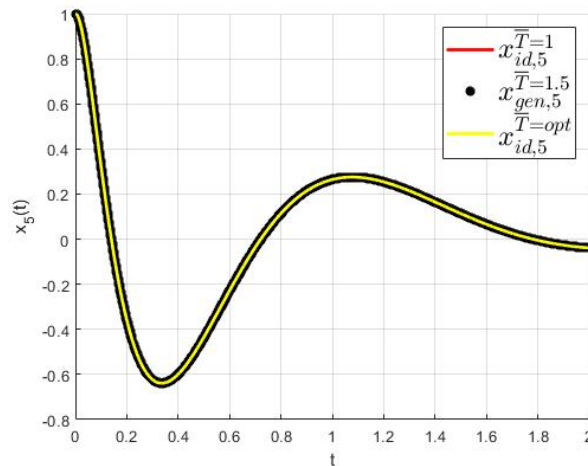


Рис. 5.7: Достоверная функция $\delta(t)$ (черная линия) и идентифицированное оптимальное управление с непрерывным обновлением информации (желтая линия), сгенерированное при $\bar{T}_{opt} = 1.4815$ и $Q^{\bar{T}_{opt}}$.

том, допущение фиксированного, — возможно, неправильного информационного горизонта — может привести к субоптимальным результатам. Таким образом, важна идентификация для оптимального значения информационного горизонта \bar{T} по наблюдениям, которая проведена с помощью реализованного нами подхода обратного оптимального управления. Идентифицированные параметры целевой функции $Q^{\bar{T}_{opt}}$ с оптимальным значением \bar{T} наиболее близки к достоверным (см. (5.23)), но отличаются, в частности, первым диагональным элементом матрицы Q (отклонение угла скольжения). В этом маневре значение угла скольжения не оказывает значительного влияния на другие параметры состояния. Поэтому корректно определить соответствующие параметры непросто. Тем не менее, идентифицированные параметры адекватно отражают достоверные траектории.

5.5 Заключение

В этом разделе представлены подход оптимального управления и соответствующая обратная задача с непрерывным обновлением информации: лицо,

принимающее решение, изменяет свое поведение на основе новой доступной информации, возникающей благодаря движущемуся временному горизонту. Результаты численного моделирования показывают практическую применимость подхода, а также подчеркивают важность идентификации информационного горизонта \bar{T}_{opt} . Данный подход предоставляет возможности для более глубокого моделирования технических систем, включающих человека. Дальнейшие исследования будут посвящены тестированию обратного оптимального управления с непрерывным обновлением информации по реальным данным измерений поведения различных водителей.

Заключение

Дадим краткий обзор результатов, полученных в данной работе. Во введении представлен обзор литературы по теме исследования, а также рассмотрены актуальность исследования, его теоретическая и практическая ценность, и научная новизна.

В Главе 1 представлено ПРД-ядро — новое кооперативное решение для дифференциальных игр. Это решение строится с использованием специальных условий для процедур распределения дележа (ПРД) и по построению обладает сильной динамической устойчивостью. Кроме того, в Главе 1 предложен численный алгоритм для анализа непустоты IDP-ядра.

В Главе 2 представлен класс динамических игр с обновлением информации. Рассмотрены кооперативные и некооперативные случаи и приведены условия для связанных принципов оптимальности (равновесие по Нэшу и кооперативные стратегии). Эти условия основаны на уравнениях Беллмана для динамической постановки. Кроме того, введена процедура обновления случайного горизонта для случая динамического обновления информации. Определена характеристическая функция для случая динамического обновления информации и доказана связь между кооперативным решением для каждой усеченной подыгры (на усеченном временном интервале) и кооперативным решением для игры, определенной на всем временном интервале. Для численного моделирования использована динамическая игровая модель олигополии с рекламой.

В Главе 3 представлен класс кооперативных дифференциальных игр с динамическим обновлением информации. Рассмотрены модели динамического обновления информации, а именно: классическая модель (в которой игроки имеют определенную информацию о функциях выигрыша и уравнениях дви-

жения) и модель обновления со случайным информационным горизонтом и неопределенной (случайной) будущей динамикой. Для построения кооперативных стратегий и характеристической функции с динамическим обновлением информации в качестве условий оптимальности использованы уравнения Гамильтона–Якоби–Беллмана. Для кооперативного решения с динамическим обновлением информации доказано свойство Δt -динамической устойчивости. Введено понятие характеристической функции с динамическим обновлением информации и доказана связь между кооперативным решением для каждой усеченной подыгры и кооперативным решением для игры, определенной на всем временном интервале.

Глава 4 посвящена классу кооперативных дифференциальных игр с непрерывным обновлением информации. Здесь процедура обновления выполняется непрерывно во времени. Обновление информации основано на классическом подходе, когда у игроков есть определенная информация о функции выигрыша и уравнениях движения. Принципы оптимальности для кооперативных и равновесных по Нэшу стратегий, как программных, так и позиционных, представлены в виде уравнений Гамильтона–Якоби–Беллмана, принципа максимума Понтрягина с непрерывным обновлением информации, а также условий для частного случая линейно-квадратичных автономной и неавтономной моделей с непрерывным обновлением информации. Кроме того, представлены результаты относительно сходимости решений для игр с динамическим и непрерывным обновлением информации для линейно-квадратичных автономных и неавтономных моделей и одной специальной классической дифференциальной игры добычи ресурса. Доказана сходимость кооперативных и равновесных по Нэшу стратегий и связанных траекторий с динамическим обновлением информации к соответствующим стратегиям и траекториям с непрерывным обновлением информации. Доказана сильная динамическая устойчивость кооперативного решения с непрерывным обновлением информации.

В Главе 5 подход с непрерывным обновлением информации применен для решения технических задач. Сформулирована и решена новая обратная задача оптимального управления с непрерывным обновлением информации. Эта задача и соответствующее решение важны для определения человеческих параметров в технических системах человеко-машинного типа (системы помо-

щи водителю, адаптивные экзоскелеты и т.д.). Другие результаты связаны с применением подхода с непрерывным обновлением информации к прикладной обратной задаче оптимального управления для систем помощи водителю.

Таким образом, основные цели и задачи диссертационной работы были выполнены полностью. Соискатель выражает благодарность и большую признательность д.физ.-мат.н., заведующему кафедрой Математического моделирования энергетических систем СПбГУ, проф. В.В. Захарову за поддержку, помощь, обсуждение результатов и передачу богатого исследовательского опыта.

Дальнейшие исследования могут проводиться в направлении изучения новых видов обновления информации. Байесовское обновление информации представляется многообещающим. Байесовский подход может применяться в различных технических системах человеко-машинного типа. Кроме того, важно установить условия существования и единственности кооперативных стратегий и равновесных по Нэшу стратегий с непрерывным обновлением информации.

Список литературы

1. Isaacs R. Differential Games. –New York : John Wiley and Sons, 1965.
2. Bellman R. Dynamic Programming. –Princeton : Princeton University Press, 1957.
3. Bercovitz L. D. A variational approach to differential games // Adv. in game theory, Ann. of Math. Studies. –1964. –P. 127–175.
4. Leitmann G. Cooperative and Non-Cooperative Many Players Differential Games. –N.Y.:Springer-Verlag, 1974.
5. Fleming W. H. Rishel. R. W. Deterministic and Stochastic Optimal Control. –New York: Springer-Verlag, 1975.
6. A. Friedman. Differential Games. –Wiley, N.Y., 1971.
7. L.S. Pontryagin. On the theory of differential games // Uspekhi matematicheskikh nauk. –1966. –Vol. 21, no. 4 (130). –P. 219–274.
8. L.S. Pontryagin V. G. Boltyansky R. V. Gamkrelidze E. F. Mishchenko. Mathematical theory of optimal processes. –M .: Science, 1976.
9. Petrosyan L. A., Tomsy G. V. Dynamic games and their applications. –Leningrad : Ed. Leningrad State University, 1982.
10. A. Petrosyan L. Differential games of pursuit. –L .: Publishing house Leningrad. University, 1977.
11. N.N. Krasovskiy. Game tasks about the meeting of movements. –M .: Nauka, 1970.

12. Krasovsky N. N. A. N. Kotelnikov. On the differential game for interception // Proceedings of the Mathematical Institute. V.A. Steklova. –2010. –no. T. 268. –P. 168–214.
13. Pshenichny B.N. Ostapenko V.V. Differential games. –Kiev: Nauk. Dumka, 1992.
14. N.N. Krasovskiy A.I. Subbotin. Positional differential games. –M.: Nauka, 1974.
15. A.I. Subbotin A.G. Chentsov. Optimization of guarantees in control problems. –M.: Nauka, 1981.
16. Subbotin A.I. Generalized solutions of differential equations of the 1st order. Prospects for dynamic optimization. –M.: Izhevsk, 2003.
17. Nash J. F. Equilibrium points in n-person games. // Proc. Natl. Acad. Sci. United States. –1950. –no. 36. –P. 48–49.
18. E. M. Vaysbord V. I. Zhukovsky. Introduction to differential games of several persons and their applications. –M.: Sov. radio, 1980.
19. Grigorenko N. L. Differential pursuit games by several objects. –M.: Publishing house of Moscow State University, 1983.
20. N.L.Grigorenko. To the theory of differential games of three persons // Proceedings of the IMM UB RAS. –2006. –Vol. 12, no. 1. –P. 78–86.
21. V.I Zhukovsky A. A. Chikriy. Linear-quadratic differential games. –Kiev: Naukova Dumka, 1994.
22. V. I. Zhukovsky M. E. Salakudze. Some game control problems and their applications. –Tbilisi, 1998.
23. V. I. Zhukovsky N. T. Tynyansky. Equilibrium controls of multicriteria dynamical systems. –M.: Publishing house Moscow State University, 1994.
24. Kleimenov A. F. Tasks of conflict management // App. Mathematics and Mechanics. –1975. –Vol. 39, no. 2. –P. 225–234.

25. Kleimenov A. F. On the cooperative theory of noncooperative positional differential games // Dokl. Academy of Sciences of the USSR. –1990. –Vol. T.39, no. 1. –P. 32–35.
26. Kleimenov A. F. Non-antagonistic positional differential games. –Ekaterinburg : Science, 1993.
27. Kleimenov A. F. Cooperative solutions in the position differential game of many individuals with continuous payment functions // Applied Mathematics and Mechanics. –1990. –T. 54, № 3. –C. 389–394.
28. Kleimenov A. F. On solutions in a nonantagonistic positional differential game // App. Mathematics and Mechanics. –1997. –Vol. 61, no. 5. –P. 739–746.
29. Kononenko A. F. The structure of an optimal strategy in dynamic controlled systems // Zh. Vychisl. mat. and mat. fiz. –1980. –Vol. 20, no. 5. –P. 1105–1116.
30. V. N. Lagunov V.V.Sushkin. Multistep positional games of N persons. –Tver, 1993.
31. Malafeev O.A. On the existence of a equilibrium situation in dynamic games with dependent movements // Zh. Vychisl. mat. and mat. fiz. –1974. –Vol. 14, no. 1. –P. 88–98.
32. Malafeev O.A. On the existence of a equilibrium situation in differential coalition-free games of two persons with independent movements // Bulletin of Leningrad State University. –1980. –no. 7.
33. Malafeev O.A. Stationary strategies in differential games // Zh. Vychisl. mat. and mat. fiz. –1977. –Vol. 17, no. 1. –P. 42–51.
34. Malafeev O.A. Controlled conflict systems. –SPb .: Publishing house of SPbSU, 2000.
35. F.F Nikitin S.V Chistyakov. On differential antagonistic games with unlimited duration // Bulletin of St. Petersburg State University. –1980. –Vol. 1, no. 3. –P. 37–43.

36. Petrov N.N. Proof of the existence of the value of the pursuit game with bounded time // *Differential Uravn.* –1970. –Vol. 6, no. 5. –P. 784–797.
37. Petrov N.N. On the existence of the meaning of the game of pursuit // *DAN SSSR.* –1970. –Vol. 190, no. 6. –P. 1289–1291.
38. G. Chentsov A. On the game problem of convergence at a given moment of time // *Mat. Sat.* –1976. –Vol. 99, no. 3. –P. 394–420.
39. G. Chentsov A. On an example of an irregular differential game // *PMM.* –1976. –Vol. 40, no. 6. –P. 1113–1116.
40. Chistyakov S. V. Dynamic aspect of the solution of classical cooperative games // *Dokl. RAS.* –1993. –Vol. 330, no. 6. –P. 707–709.
41. Chistyakov S. V. On non-cooperative differential games // *Reports of the USSR Academy of Sciences.* –1981. –T. 259, № 5. –C. 1052–1055.
42. Chistyakov S. V. On the construction of strongly dynamically stable solutions cooperative differential games // *Vestn. SPbSU.* –1992. –no. 1. –P. 50–54.
43. Chistyakov S. V. Program iterations and universal delta - optimal strategy in positional differential game // *Dokl. AN SSSR.* –1991. –Vol. 319, no. 6. –P. 1333–1336.
44. S. V. Chistyakov FF Nikitin. Existence and uniqueness theorem for the solution of the generalized Isaacs-Bellman equation // *Differential Equations.* –2007. –Vol. 43, no. 6. –P. 757–766.
45. Subbotina N. N. The method of characteristics for Hamilton-Jacobi equations and applications to dynamical optimization // *Journal of math. sciences.* –2007. –Vol. 135, no. 3. –P. 2955–3091.
46. Zhukovskii V.I. *Lyapunov Functions in Differential Games.* –London and New York: Taylor Francis, 2003.
47. E. J. Dockner Van Long N. International pollution control: Cooperative versus noncooperative strategies // *Journal of Environmental Economics and Management.* –1993. –Vol. 25. –P. 13–29.

48. H. S. Bierman L. Fernandez. Game Theory with Economic Applications. –Massachusetts, USA : Addison: Wesley Publishing Company, Inc., 1993.
49. Gibbons R. Game Theory for Applied Economists. –Princeton, New Jersey : Princ. Univ. Press, 1992.
50. Dockner E. J., Jorgensen S., et al. Differential Games in Economics and Management Science. –Cambridge : Cambridge University Press, 2000.
51. L. A. Petrosyan N. A. Zenkevich E. V. Shevkoplyas. Game Theory. –SPb: BHV-Petersburg, 2012.
52. V.V Zakharov L.A. Petrosyan. Introduction to mathematical ecology. –L .: Publishing house of Leningrad State University, 1986.
53. V.V Zakharov L.A. Petrosyan. Mathematical models in ecology. –SPb .: Publishing house of St. Petersburg. University, 1996.
54. V. A. Gorelik A. F. Kononenko. Game-theoretic models of decision making in ecological and economic systems. –M., 1982.
55. V. V. Mazalov A. N. Rettieva. Equilibrium according to Nash in the problems of environmental protection // Mat. modeling. –2006. –Vol. 18, no. 5. –P. 73–90.
56. Rettieva A. N. Cooperative regulating condition in the problem of separation of biological resources // UBS. –2009. –Vol. 26.1. –P. 366–384.
57. Breton M., Zaccour G., Zahaf M. A differential game of joint implementation of environmental projects // Automatica. –2005. –Vol. 41, no. 10. –P. 1737–1749.
58. P. Chander H. Tulkens. The Core of an Economy with Multilateral Environmental Externalities // International Journal of Game Theory. –1997. –Vol. 23. –P. 379–401.
59. Chander P. The Gamma-Core and Coalition Formation // international journal of game theory. –2007. –no. 35. –P. 539–556.

60. P. Dasgupta R. Gilbert J. Stiglitz. Strategic considerations in invention and innovation: the case of natural resources // *Econometrica*. –1983. –Vol. 51. –P. 1439–1448.
61. E. J. Dockner A. Gaunersdorfer. On the Profitability of Horizontal Mergers in Industries with Dynamic Competition // *Japan and the World Economy*. –2001. –no. 4. –P. 195–216.
62. J. A. Filar P. S. Gaertner. A regional allocation of world CO₂ emission reductions // *Mathematics and Computers in Simulation*. –1997. –Vol. 43. –P. 269–275.
63. B. C. Giri S. K. Goyal. Recent trends in modeling of deteriorating inventory // *European Journal of Operational Research*. –2001. –Vol. 134, no. 1. –P. 1–16.
64. A. Haurie G. Zaccour. Differential game models of global environmental management // *Annals of Dynamic Games, Boston*. –1994. –P. 124–132.
65. al A. Haurie.et. *Games and dynamic games*. –World Scientific Books, 2012.
66. S. Jørgensen G. Martin-Herran G. Zaccour. Agreeability and Time Consistency in Linear-State Differential Games // *Journal of Optimization Theory and Applications*. –2003. –Vol. 119, no. 1. –P. 49–63.
67. S. Jørgensen G. Zaccour. *Developments in Differential Game Theory and Numerical Methods: Economic and Management Applications // Computational Management Science*. –2007. –Vol. 4, no. 2. –P. 159–182.
68. S. Jørgensen G. Sorger. Feedback Nash equilibria in a problem of optimal fishery management // *Journal of Optimization Theory and Applications*. –1990. –Vol. 64. –P. 293–310.
69. Jorgensen S., Yeung D. W. K. Inter- and intragenerational renewable resource extraction // *Annals of Operations Research*. –1999. –Vol. 88. –P. 275–289.
70. S. Jørgensen G. Zaccour. Time-Consistent Side Payment in a Dynamic Game of Downstream Pollution // *Journal of Economic Dynamics and Control*. –2001. –Vol. 25. –P. 1973–1987.

71. Kaitala. V. Equilibria in a stochastic resource management game under imperfect information // *European Journal of Operational Research*. –1993. –Vol. 71. –P. 439–453.
72. V. Kaitala M. Pohjola. Sustainable international agreements on greenhouse warming // *A game theory study*, in Carraro and Filar (eds.), *Control and Game Theoretic Models of the Environment*, *Annals of the International Society of Dynamic Games*. –1995. –Vol. II.
73. J. Krawczyk G. Zaccour. Management of pollution from decentralized agents by the local government // *International Journal of Environment and Pollution*. –1999. –Vol. 12, no. 2/3. –P. 343–357.
74. Sorger G. Competitive dynamic advertising: a modification of the case game // *Journal of Economic Dynamics and Control*. –1989. –Vol. 13. –P. 55–80.
75. Zakharov V. One approach to allocating damage to environment. –*System Modelling and Optimization*, New York.: Springer-Verl., 1994.
76. Petrosian O. L. Looking Forward Approach in Cooperative Differential Games. // *International Game Theory Review*. –2016. –no. 2.
77. Petrosian O., Barabanov A. Looking Forward Approach in Cooperative Differential Games with Uncertain-Stochastic Dynamics // *Journal of Optimization Theory and Applications*. –2017. –Vol. 172, no. 1. –P. 328–347.
78. Petrosian O. L., Tur A. V. Hamilton-Jacobi-Bellman Equations for Non-cooperative Differential Games with Continuous Updating. // *Communications in Computer and Information Science*. –2019. –Vol. 1090. –P. 178–191.
79. Kuchkarov I. I., Petrosian O. L. On class of linear quadratic non-cooperative differential games with continuous updating. // *Lecture Notes in Computer Science*. –2019. –Vol. 11548. –P. 635–650.
80. Basar T., Olsder G. J. *Dynamic noncooperative game theory*. –London : Academic Press, 1995.

81. Shevkoplyas E. V. Optimal Solutions in Differential Games with Random Duration // Journal of Mathematical Sciences. –2014. –Vol. 199, no. 6. –P. 715–722.
82. Petrosyan L. A., Murzov N. V. Game-theoretic Problems in Mechanics // Lithuanian Mathematical Collection. –1966. –№ 3. –C. 423–433.
83. Pontryagin L. S. On Theory of Differential Games // Successes of Mathematical Sciences. –1966. –№ 26, 4 (130). –C. 219–274.
84. Goodwin G. C., Seron M. M., Dona J. A. Constrained Control and Estimation: An Optimisation Approach. –New York : Springer, 2005.
85. Kwon W. H., Han S. H. Receding Horizon Control: Model Predictive Control for State Models. –New York : Springer, 2005.
86. Rawlings J. B., Mayne D. Q. Model Predictive Control: Theory and Design. –Madison : Nob Hill Publishing, 2009.
87. Wang L. Model Predictive Control System Design and Implementation Using MATLAB. –New York : Springer, 2005.
88. The explicit linear quadratic regulator for constrained systems / Bemporad A., Morari M., Dua V., and Pistikopoulos E. // Automatica. –2002. –Vol. 38, no. 1. –P. 3–20.
89. Hempel A., Goulart P., Lygeros J. Inverse Parametric Optimization With an Application to Hybrid System Control // IEEE Transactions on Automatic Control. –2015. –Vol. 60, no. 4. –P. 1064–1069.
90. Kwon W., Bruckstein A., Kailath T. Stabilizing state-feedback design via the moving horizon method // 21st IEEE Conference on Decision and Control. –1982.
91. Kwon W., Pearson A. A modified quadratic cost problem and feedback stabilization of a linear system // IEEE Transactions on Automatic Control. –1977. –Vol. 22(5). –P. 838–842.
92. Mayne D., Michalska H. Receding horizon control of nonlinear systems // IEEE Transactions on Automatic Control. –1990. –Vol. 35(7). –P. 814–824.

93. Shaw L. Nonlinear control of linear multivariable systems via state-dependent feedback gains // IEEE Transactions on Automatic Control. –1979. –Feb. –Vol. 24, no. 1. –P. 108–112. –Conference Name: IEEE Transactions on Automatic Control.
94. Haurie A. Feedback equilibria in differential games with structural and modal uncertainties // Advances in Large Scale Systems. –1984.
95. Joao B. R. do Val, Basar T. Receding horizon control of jump linear systems and a macroeconomic policy problem // Journal of Economic Dynamics & Control. –1999. –Vol. 23. –P. 1099–1131.
96. Van Den Broek W. A. Moving horizon control in dynamic games // Journal of Economic Dynamics & Control. –2002. –Vol. 26. –P. 937–961.
97. Sethi S., Sorger G. A Theory of Rolling Horizon Decision Making // Annals of Operations Research. –1991. –Vol. 29. –P. 387–416.
98. Vasin A., Divtsova A. The repeated game modelling an agreement on protection of the environment // In VIII Moscow International Conference on Operations Research (ORM2018), MAKS Press Moscow. –2018. –Vol. 1. –P. 261–263.
99. Vasin A., Divtsova A. Game-theoretic model of agreement on limitation of transboundary atmospheric pollution // Mat. Teor. Igr Pril. –2017. –Vol. 9, no. 1. –P. 27–44.
100. Petrosian O. L., Gromova E. V., Pogozev S. V. Strong time-consistent subset of core in cooperative differential games with finite time horizon // Mathematical theory of games and its applications. –2016. –Vol. 8, no. 4. –P. 79–106.
101. Petrosian O. L., Gromova E. V., Pogozev S. V. Strong time-consistent subset of core in cooperative differential games with finite time horizon // Automation and Remote Control. –2018. –Vol. 79, no. 10. –P. 1912–1928.
102. Wolf D. A., Zakharov V. V., Petrosian O. L. On the existence of IDP-core in cooperative differential games // Mathematical theory of games and its applications. –2017. –Vol. 9, no. 4. –P. 18–38.

103. Zakharov V. V., O-Hun K. Linear programming approach in cooperative games // Korean Math. Soc. –1997. –Vol. 34, no. 2. –P. 423–435.
104. Petrosian O. L., Zakharov V. V. IDP-Core: Novel Cooperative Solution for Differential Games // Mathematics. –2020. –Vol. 8, no. 5.
105. Petrosian O. L., Pogozhev S. V. Looking Forward Approach with Random Horizon in Cooperative Differential Games // Contributions to Game Theory and Management. –2020. –Vol. 13. –P. 360–387.
106. Petrosian O. L., Kuchkarov I. I. About the Looking Forward Approach in Cooperative Differential Games with Transferable Utility // Frontiers of Dynamic Games: Game Theory and Management. –2019. –P. 175–208.
107. Petrosian O. L., Nastych M. A., Li Y. The Looking Forward Approach in a Differential Game Model of the Oil Market with Non-transferable Utility // Frontiers of Dynamic Games. –2020. –P. 215–244.
108. Gromova E., Petrosian O. Control of Informational Horizon for Cooperative Differential Game of Pollution Control // 2016 International Conference Stability and Oscillations of Nonlinear Control Systems (Pyatnitskiy's Conference). –2016. –P. 1–4.
109. Petrosian O. L., Gromova E. V. Cooperative Differential Games with Dynamic Updating // IFAC-PapersOnLine. –2018. –Vol. 51, no. 32. –P. 413–417.
110. Petrosian O. L., Nastych M. A., Volf D. A. Non-cooperative Differential Game Model of Oil Market with Looking Forward Approach // Frontiers of Dynamic Games. –2018. –Vol. 25. –P. 189–202.
111. Petrosian O. L., Nastych M. A., Volf D. A. Differential game of oil market with moving informational horizon and non-transferable utility. // Proceedings of the Constructive Nonsmooth Analysis and Related Topics (dedicated to the memory of VF Demyanov) (CNSA), Saint-Petersburg, Russia, 22-27 May. –2017. –P. 1–4.

112. Petrosian O. L. Looking Forward Approach in Cooperative Differential Games with infinite-horizon // Vestnik S.-Petersburg Univ. Ser. 10. Prikl. Mat. Inform. Prots. Upr. –2016. –no. 4. –P. 18–30.
113. Petrosian O. L., Shi L., et al. Moving Information Horizon Approach for Dynamic Game Models // Mathematics. –2019. –Vol. 7, no. 12. –P. 1–31.
114. Shi L., Petrosian O. L. A Dynamic Oligopoly Marketing Model of Advertising // Contributions to Game Theory and Management. –2018. –Vol. 11. –P. 207–223.
115. Shi L., Petrosian O. L., Boiko A. V. Looking Forward Approach for dynamic cooperative advertising game model. // Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes. –2019. –Vol. 15, no. 2. –P. 221–234.
116. Yeung D. W. K., Petrosian O. L. Cooperative Stochastic Differential Games with Information Adaptation // International Conference on Communication and Electronic Information Engineering (CEIE 2016). –2016. –P. 375–381.
117. Kuchkarov I. I., Petrosian O. L. Open-Loop Based Strategies for Autonomous Linear Quadratic Game Models with Continuous Updating // Lecture Notes in Computer Science. –2020. –Vol. 12095. –P. 212–230.
118. Petrosian O. L., Tur A. V., Wang Z. Cooperative differential games with continuous updating using Hamilton–Jacobi–Bellman equation // Optimization Methods and Software. –2020. –Vol. 1275. –P. 256–270.
119. Wang Z., Petrosian O. L. On class of non-transferable utility cooperative differential games with continuous updating // Journal of Dynamics and Games. –2020. –Vol. 7, no. 4. –P. 291–302.
120. Petrosian O. L., Tur A. V., Zhou J. Pontryagin’s Maximum Principle for Non-cooperative Differential Games with Continuous Updating // Communications in Computer and Information Science. –2020. –Vol. 1275. –P. 256–270.

121. Kuchkarov I. I. Feedback based strategies for autonomous linear quadratic cooperative differential games with continuous updating // Contributions to Game Theory and Management. –2020. –Vol. 13. –P. 244–251.
122. Kuchkarov I. I., Petrosian O. L. Non-autonomous Linear Quadratic Non-cooperative Differential Games with Continuous Updating (to be published) // Journal of Global Optimization. –2021.
123. Zhou J., Tur A. V., et al. Transferable Utility Cooperative Differential Games with Continuous Updating using Pontryagin Maximum Principle // Mathematics. –2020. –Vol. 9, no. 2.
124. Petrosian O. L., Tikhomirov D. S., Gao H. Differential Game Model of Resource Extraction with Continuous and Dynamic Updating (to be published) // Journal of Systems Science and Complexity. –2021.
125. Petrosian O. L., Inga J., et al. Optimal Control and Inverse Optimal Control with Continuous Updating for Human Behavior Modeling // IFAC-PapersOnLine. –2020. –Vol. 53, no. 2. –P. 6670–6677.
126. Petrosian O. L., Tikhomirov D. S., Kuchkarov I. I. et al. About One Differential Game Model with Dynamic Updating // Automation and Remote Control. –2020. –no. 81. –P. 1733–1750.
127. Petrosian O. L., Tur A. V., et al. Cooperative differential game model with continuous updating // Mat. Teor. Igr Pril. –2020. –Vol. 12, no. 2. –P. 82–109.
128. Petrosian O. L., Tikhomirov D. S., et al. About one differential game model with dynamic updating // Mat. Teor. Igr Pril. –2019. –Vol. 11, no. 2. –P. 40–67.
129. Yeung D. W. K., Petrosian O. L. Infinite Horizon Dynamic Games: A New Approach via Information Updating // International Game Theory Review. –2017. –Vol. 19, no. 4. –P. 1–23.
130. Rufus I. Differential Games. –New York : John Wiley and Sons, 1965.

131. Fleming W. H. The convergence problem for differential games // Adv. in game theory, Ann. of Math. Studies. –1964. –P. 175–195.
132. Krasovsky N. N. Control of Dynamical System. Problem of minimum guaranteed result. –Moscow : Science, 1985.
133. Kononenko A. F. On equilibrium positional strategies in non-antagonistic differential games // Reports of the USSR Academy of Sciences. –1976. –T. 231, № 2. –C. 285–288.
134. Lukyanenko A., Gurtov A. Towards behavioral control in multi-player network games // 2009 International Conference on Game Theory for Networks. –2009. –P. 1–8.
135. Filippov A. F. Introduction to the theory of differential equations (in Russian). –Moscow : Editorial URSS, 2004.
136. Gillies D. B. Some theorems on n person games // Ph.D. thesis. –1953.
137. Edgeworth F. Y. Mathematical physics. // Kegan Paul, London. –1881.
138. Nikolaevskiy I., Lukyanenko A., Gurtov A. Nash Bargaining Solution Allocation is Not Suitable for Datacenter Jobs // International Game Theory Review. –2017. –Vol. 19, no. 1. –P. 1–22.
139. Kleimenov A. F. To the cooperative theory of non-coalition positional games // Reports of the USSR Academy of Sciences, Moscow. –1990.
140. Kleimenov A. F. Cooperative solutions in the position differential game of many individuals with continuous payment functions // Applied Mathematics and Mechanics. –1990. –Vol. 54, no. 3. –P. 389–394.
141. Yeung D. W. K., Petrosyan L. A. Subgame Consistent Economic Optimization: An Advanced Cooperative Dynamic Game Analysis // Springer, New York. –2012.
142. Petrosyan L. A., Danilov N. N. Cooperative differential games and their applications. –Publishing house of Tomsk University, 1985.

143. Tolwinski B., Haurie A., Leitmann G. Cooperative equilibria in differential games // Journal of Mathematical Analysis and Applications. –1986. –Vol. 119, no. 1. –P. 182–202.
144. Zaccour G. Time Consistency in Cooperative Differential Games: A Tutorial // INFOR: Information Systems and Operational Research. –2008. –Vol. 46, no. 1. –P. 81–92.
145. Gao H., Petrosian L. A., et al. Cooperation in two-stage games on undirected networks // Journal of Systems Science and Complexity. –2017. –Vol. 30(3). –P. 680–693.
146. Gromova E. V, Petrosyan L. A. Strongly dynamically stable cooperative solution in one differential game of harmful emissions management // Managing large systems. –2015. –Vol. 55. –P. 140–159.
147. Vorob'ev N. N. Game Theory. Lectures for Economists and Systems Scientists. –1977.
148. Scarf H. E. The core of an n person game // *Economica*. –1967. –Vol. 1, no. 35. –P. 50–69.
149. Billera L. J. Some theorems on the core of n person game // *SIAM Journal of Applied Mathematics*. –1970. –Vol. 18, no. 3. –P. 567–579.
150. Shapley L. S. On balanced games without side payments // *Mathematical programming*. –1972. –P. 261–290.
151. Bondareva O. N. Some applications of linear programming methods to the theory of cooperative games // *Problems of Cybernetics*. –1963. –№ 10. –C. 119–140.
152. Shapley L. S. On balanced sets and cores // *Naval Research Logistic Quarterly*. –1967. –Vol. 14. –P. 453–460.
153. Zakharov V. V., Dementieva M. Multistage Cooperative Games and Problem of Time Consistency // *International Game Theory Review*. –2004. –Vol. 6, no. 1. –P. 157–170.

154. Zakharov V. V., Akimova A. N. Geometric Properties of the Core, Subcore, Nucleolus. –2002. –Vol. 8. –P. 281–289.
155. Owen G. Game theory. –New York : Academic Press, 1982.
156. Von Neumann J., Morgenstern O. Theory of Games and Economic Behavior. –Princeton : Princeton University Press, 1970.
157. Haurie A. A note on nonzero-sum differential games with bargaining solutions // Journal of Optimization Theory and Applications. –1976. –Vol. 18, no. 1. –P. 31–39.
158. Petrosyan L. A. Time consistency of solutions for differential games with many participants // Vestnik of Leningrad University. Ser. 1: Mathematics, Mechanics, Astronomy. –1977. –№ 4. –C. 46–52.
159. Petrosyan L. A. Strong time consistent differential optimality principles // Vestnik of Leningrad University. Ser. 1: Mathematics, Mechanics, Astronomy. –1993. –№ 4. –C. 35–40.
160. Petrosyan L. A., Danilov N. N. Time consistency of solutions for non-antagonistic differential games with transferable payoffs // Vestnik of Leningrad University. Ser. 1: Mathematics, Mechanics, Astronomy. –1979. –№ 1. –C. 52–59.
161. Petrosyan L. A. Strongly time-consistent differential optimality principles // Vestnik St. Petersburg Univ. Math. Series 1. Mathematics. Mechanics. Astronomy. –1993. –Vol. 26. –P. 40–46.
162. Petrosyan L. A., Danilov N. N. Cooperative differential games and their applications. –1985.
163. Parilina E. M., Zaccour G. Node-Consistent Shapley Value for Games Played over Event Trees with Random Terminal Time // Journal of Optimization Theory and Applications. –2017. –Vol. 175, no. 1. –P. 236–254.
164. Parilina E. M., Zaccour G. Node-consistent core for games played over event trees // Automatica. –2015. –Vol. 53. –P. 304–311.

165. Yeung D., Petrosyan L. A. Subgame Consistent Solutions of a Cooperative Stochastic Differential Game with Nontransferable Payoffs // Journal of optimization theory and applications. –2005. –Vol. 124, no. 3. –P. 701–724.
166. Petrosyan L. A., Yeung D. A Time-Consistent Solution Formula for Bargaining Problem in Differential Games // Journal of optimization theory and applications. –2014. –Vol. 16, no. 4. –P. 1–11.
167. Petrosyan L. A., Danilov N. N. Stability of solutions in non-zero sum differential games with transferable payoffs // Vestnik Leningrad State University. Series 1. Mathematics. Mechanics. Astronomy. –1979. –P. 52–59.
168. Dockner E., Jorgensen S., et al. Differential Games in Economics and Management Science // Cambridge University Press, Cambridge. –2001.
169. Yeung D. W. K., Petrosyan L. A. Subgame Consistent Economic Optimization: An Advanced Cooperative Dynamic Game Analysis // Springer Science & Business Media: Berlin/Heidelberg, Germany. –2011.
170. Owen G. Game Theory. –Emerald Group Publishing Limited: Bingley, UK, 2013.
171. Nash J. F. Non-cooperative games // Ann. Math. –1951. –no. 54. –P. 286–295.
172. Yeung D. W. K. On differential games with a feedback Nash equilibrium. // J. Optim. Theory Appl. –1994. –no. 82. –P. 181–188.
173. Bellman R. Dynamic Programming. –2003.
174. Chander P., Tulkens H. A core-theoretic solution for the design of cooperative agreements on transfrontier pollution. // Public Goods, Environmental Externalities and Fiscal Competition. –2006. –P. 176–193.
175. Petrosjan L. A. Differential Pursuit Games // World Scientific Publ. Co. Pte. Ltd.: Singapore. –1993.

176. Yeung D. W. K., Petrosjan L. A. Cooperative Stochastic Differential Games // Springer Science & Business Media: Berlin/Heidelberg, Germany. –2006.
177. Petrosyan L. A., Danilov N. N. Stability of solutions in non-zero sum differential games with transferable payoffs. // Vestn. Leningr. Univ. –1979. –no. 1. –P. 52–59.
178. Yeung D. W. K., Petrosyan L. A. Subgame consistent cooperative solution of dynamic games with random horizon. // J. Optim. Theory Appl. –2011. –no. 150. –P. 78–97.
179. Prasad A., Sethi S. P., Naik P. A. Optimal Control of an Oligopoly Model of Advertising. // IFAC Proc. Vol. –2009. –no. 42. –P. 91–96.
180. Basar T., Olsder G. J. Dynamic Noncooperative Game Theory // London, Academic Press. –1995.
181. Krasovskii N. N., Kotel'nikova A. N. On a differential interception game // Proceedings of the Steklov Institute of Mathematics. –2010. –no. 268. –P. 161–206.
182. Bellman R. Dynamic Programming // Princeton. –1957.
183. Subbotin A. I. Generalization of the main equation of differential game theory // Journal of Optimization Theory and Applications. –1984. –Vol. 43. –P. 103–133.
184. Subbotin A. I. Generalized Solutions of First Order PDEs // Basel, Birkhauser. –1995.
185. Chander P., Tulkens H. A core-theoretic solution for the design of cooperative agreements on transfrontier pollution // International Tax and Public Finance. –1995. –Vol. 2, no. 2. –P. 279–293.
186. Chander P. The gamma-core and coalition formation // International Journal of Game Theory. –2007. –Vol. 35, no. 4. –P. 539–556.

187. Petrosyan L. A., Danilov N. N. Stability of solutions in non-zero sum differential games with transferable payoffs // Vestnik Leningrad State University. –1979. –Vol. 1. –P. 52–59.
188. Yeung D. W. K., Petrosian L. A. Subgame-consistent Economic Optimization // New York, Springer. –2012.
189. Yeung D. W. K., Petrosyan L. A. Subgame-consistent Economic Optimization // Springer. –2012.
190. Petrosyan L. A., Zaccour G. Time-consistent Shapley value allocation of pollution cost reduction // J. Econ. Dyn. Control. –2003. –Vol. 3. –P. 381–398.
191. Petrosyan L. A., Danilov N. N. Stability of solutions in non-zero sum differential games with transferable payoffs // Vestn. Leningr. Univ. –1979. –Vol. 1. –P. 52–59.
192. Petrosyan L. A. Time-consistency of solutions in multi-player differential games // Vestnik Leningrad State University. –1977. –T. 4. –C. 46–52.
193. Jorgensen S., Yeung D. W. K. Inter- and intragenerational renewable resource extraction // Annals of Operations Research. –1999. –P. 275–289.
194. Shevkoplyas E. V. Time-consistency Problem Under Condition of a Random Game Duration in Resource Extraction // Contributions to Game Theory and Management. –2009. –Vol. 2. –P. 461–473.
195. Shevkoplyas E. V. Stable cooperation in differential games with random duration // Mat. Teor. Igr Pril. –2010. –Vol. 2, no. 3. –P. 79–105.
196. Shevkoplyas E. V. The Shapley value in cooperative differential games with random duration // Ann. Dyn. Games. –2011.
197. Shevkoplyas E. V. The Hamilton-Jacobi-Bellman Equation for a Class of Differential Games with Random Duration // Automation and Remote Control. –2014. –Vol. 75. –P. 959–970.

198. Petrosjan L. A., Shevkoplyas E. V. Cooperative solutions for games with random duration // *Game Theory Appl.* –2003. –Vol. 9. –P. 125–139.
199. Kostyunin S. Yu., Shevkoplyas E. V. On simplification of integral payoff in differential games with random duration // *Vestnik S.-Petersburg Univ. Ser. 10. Prikl. Mat. Inform. Prots. Upr.* –2011. –Vol. 4. –P. 47–56.
200. Eiben A. E., Smith J. E. *Introduction to Evolutionary Computing* // Berlin, Springer. –2003.
201. Dahl C., Yucel M. Testing alternative hypotheses of oil producer behavior // *The Energy Journal.* –1991. –Vol. 12, no. 4. –P. 117–138.
202. Polasky S. Do oil producers act as 'oil'igopolists? // *Journal of Environmental Economics and Management.* –1992. –Vol. 23, no. 3. –P. 216–247.
203. Danielsen A. L., Kim S. OPEC stability: An empirical assessment // *Energy Economics.* –1998. –Vol. 10, no. 3. –P. 174–184.
204. Smith J. L. Inscrutable OPEC? Behavioral tests of the cartel hypothesis // *The Energy Journal.* –2005. –Vol. 26, no. 1. –P. 51–82.
205. Youhanna S. J. A Note on Modelling Opec Behavior 1983-1989: A Test of the Cartel and Competitive Hypotheses // *The American Economist.* –1994. –Vol. 38, no. 2. –P. 78–84.
206. Gulen S. G. Is OPEC a cartel? Evidence from cointegration and causality tests // *The Energy Journal.* –1996. –Vol. 17, no. 2. –P. 43–57.
207. Molchanov P. A statistical analysis of OPEC quota violations // *Economics.* Duke University, Durham, NC. –2003. –P. 1–31.
208. Bockem S. Cartel formation and oligopoly structure: a new assessment of the crude oil market // *Applied Economics.* –2004. –Vol. 36, no. 12. –P. 1355–1369.
209. Plaut S. E. OPEC is Not a Cartel // *Challenge.* –1981. –Vol. 24, no. 5. –P. 18–24.

210. Al-Sultan A. M. Alternative Models for OPEC Behavior // Journal of Energy and Development. –1995. –Vol. 18, no. 2. –P. 263–281.
211. Moran T. H. Managing an oligopoly of would-be sovereigns: the dynamics of joint control and self-control in the international oil industry past, present, and future // International Organization. –1987. –Vol. 41, no. 4. –P. 575–607.
212. Krasner S. D. The great oil sheikdown // Foreign Policy. –1973. –no. 13. –P. 123–138.
213. Ludkovski M., Sircar R. Differential Games // Commodities, Energy and Environmental Finance. Fields Institute Communications. –2015. –Vol. 74. –P. 317–333.
214. Salant S. Exhaustible Resources and Industrial Structure: A Nash-Cournot Approach to the World Oil Market // The Journal of Political Economy. –1976. –Vol. 84. –P. 1079–1094.
215. Berger K., Hoel M., et al. The Oil Market as an Oligopoly. Discussion paper // Central Bureau Statistics. –1988. –Vol. 32. –P. 1–19.
216. Cellini R., Lambertini L. A differential oligopoly game with differentiated goods and sticky prices // European Journal of Operational Research. –2007. –Vol. 176. –P. 1131–1144.
217. Basar T., Olsder G. J. Dynamic Noncooperative Game Theory. –Philadelphia, Pennsylvania : SIAM, 1999.
218. Fershtman C., Nitzan S. Dynamic voluntary provision of public goods // European Economic Review. –1991. –Vol. 35, no. 5. –P. 1057–1067.
219. Wirl F. Dynamic voluntary provision of public goods: extension for nonlinear strategies // European Journal of Political Economy. –1996. –Vol. 12, no. 3. –P. 555–560.
220. Basar T., Olsder G. J. Dynamic Noncooperative Game Theory // SIAM: Philadelphia, United States. –1999. –Vol. 23.

221. Leitmann G., Schmitendorf W. E. Some Sufficiency Conditions for Pareto Optimal Control // ASME Journal of Dynamical Systems, Measurement, and Control. –1973. –Vol. 95, no. G. –P. 356–361.
222. Jørgensen S. A Differential Games Solution to a Logarithmic Advertising Model // Journal of the Operational Research Society. –1982. –no. 33. –P. 425–432.
223. Jørgensen S. Sufficiency and game structure in Nash open-loop differential games // Journal of Optimization Theory and Applications. –1986. –Vol. 1, no. 50. –P. 189–193.
224. Engwerda J. LQ Dynamic Optimization and Differential Games. –New York : Willey, 2005.
225. Eisele T. Nonexistence and nonuniqueness of open-loop equilibria in linear-quadratic differential games // Journal of Optimization Theory and Applications. –1982. –Aug. –Vol. 37, no. 4. –P. 443–468. –Access mode: <https://doi.org/10.1007/BF00934951>.
226. Van Aarle B., Bovenberg A., Raith M. Monetary and Fiscal Policy Interaction and Government Debt Stabilization // Journal of Economics. –1995. –06. –Vol. 62. –P. 111–140.
227. Polyanin A., Zaitsev V. Handbook of Exact Solutions for Ordinary Differential Equations. –Boca Raton : CRC Press, 2002. –10. –ISBN: 9781584882978.
228. Petrosyan L. A., Zaccour G. Time-consistent Shapley value allocation of pollution cost reduction // J. Econ. Dyn. Control. –2003. –Vol. 27, no. 3. –P. 381–398.
229. Shapley L. A value for n-person games // In: H. W. Kuhn and A. W. Tucker, Eds., Contributions to the Theory of Games II, Annals of Mathematics Studies. –1953. –Vol. 28. –P. 307–317.
230. Petrosjan L. A., Zaccour G. Time-consistent Shapley value allocation of pollution cost reduction // Journal of Economic Dynamics & Control. –2003. –Vol. 27. –P. 381–398.

231. Başar T. Dynamic Noncooperative Game Theory, 2nd Edition. // Siam. –1999. –Vol. 23.
232. Leitmann G., Schmitendorf W. Some Sufficiency Conditions for Pareto-Optimal Control // Journal of Dynamic Systems, Measurement, and Control. –1973. –Vol. 95. –P. 356–361.
233. Long N. V. Pollution control: a differential game approach // Annals of Operations Research. –1992. –Vol. 37. –P. 283–296.
234. Yeung D. W. K., Petrosyan L. A. et al. Subgame Consistent Cooperation // Theory and Decision Library C. –2016.
235. Mombaur K., Truong A., Laumond J. P. From human to humanoid locomotion—an inverse optimal control approach // Autonomous Robots. –2010. –Apr. –Vol. 28, no. 3. –P. 369–383.
236. Flad M., Otten J., et al. Steering driver assistance system: A systematic cooperative shared control design approach // 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC). –2014. –. –P. 3585–3592.
237. Cole D. J., Pick A. J., Odhams A. M. C. Predictive and linear quadratic methods for potential application to modelling driver steering control // Vehicle System Dynamics. –2006. –Mar. –Vol. 44, no. 3. –P. 259–284.

SAINT PETERSBURG STATE UNIVERSITY

Printed as a manuscript

Petrosian Ovanes Leonovich

**DYNAMIC AND CONTINUOUS
INFORMATION UPDATING IN MODELS OF
CONFLICT CONTROL**

Specialization 2.3.1.

System analysis, information control and processing

Thesis submitted in conformity with the requirements
for the degree of doctor of physico-mathematical sciences

Translation from Russian

Saint Petersburg – 2021

Contents

INTRODUCTION	6
Chapter 1 IDP-core: Novel Cooperative Solution for Differential Games	24
1.1 Problem Statement and Preliminary Information	25
1.1.1 Differential Game Model	25
1.1.2 Cooperative Differential Game Model	26
1.1.3 Core	29
1.1.4 Non-emptiness of Core in Static Games	30
1.1.5 Time-consistency of Cooperative Solution and Imputation Distribution Procedure	31
1.2 IDP-core and Dominance of Imputation Distribution Procedures	35
1.2.1 Dominance of Imputation Distribution Procedures	35
1.2.2 IDP-core	36
1.2.3 Properties of IDP-core	42
1.3 Nonemptiness of IDP-core using Linear Programming Approach	44
1.4 Differential Game Model of Resource Extraction	45
1.4.1 Cooperative Strategies and Cooperative Trajectory	46
1.4.2 Characteristic Function	48
1.4.3 IDP-core	50
1.4.4 Non-emptiness of IDP-core	51
1.4.5 Core and IDP-core	51
Chapter 2 Class of Discrete Games with Dynamic Updating	54
2.1 Discrete Game Model with Dynamic Updating	54
2.1.1 Original Discrete Game Model	54
2.1.2 Truncated Subgame	55
2.2 Noncooperative Discrete Game with Dynamic Updating	56

2.3	Cooperative Discrete Game with Dynamic Updating	58
2.3.1	Resulting Cooperative Solution and Related Theorems	61
2.3.2	Properties of Resulting Cooperative Solution	64
2.4	Random Information Horizon Case	71
2.5	Dynamic Oligopoly Marketing Model of Advertising	74
2.5.1	Original Game Model	74
2.5.2	Noncooperative Outcomes in Truncated Subgame	76
2.5.3	Cooperative Outcomes in Truncated Subgame	81
2.5.4	Characteristic Function in Truncated Subgame	86
2.5.5	Numerical Simulation for Dynamic Advertising Game Model with Updating	87
Chapter 3	Class of Differential Games with Dynamic Updating	95
3.1	On General Class of Cooperative Differential Games with Dynamic Updating	96
3.1.1	Original Differential Game Model	96
3.1.2	Truncated Subgame	97
3.1.3	Concept of Cooperative Solution with Dynamic Updating	101
3.1.4	Construction of Characteristic Function in Game with Dy- namic Updating	104
3.1.5	Relationship of Solutions in Truncated Subgames and Result- ing Solutions	108
3.1.6	Cooperative Limited Resource Extraction Game with Dy- namic Updating	117
3.2	Cooperative Differential Games with Dynamic Updating and Uncer- tain Stochastic Dynamics	127
3.2.1	Original Game Model and Assumptions	127
3.2.2	Combined Truncated Subgame	128
3.2.3	Concept of Combined Cooperative Solution	133
3.2.4	Cooperative Differential Resource Extraction Game with Dy- namic Updating and Uncertain Stochastic Dynamics	136
3.3	Cooperative Differential Games with Dynamic Updating and Random Horizon	144
3.3.1	Random Truncated Subgame	144

3.3.2	Cooperative Solution Concept for Game with Dynamic Updating and Random Horizon	152
3.3.3	Random Horizon Dynamic Updating in Cooperative Extraction Differential Game	156
3.4	Differential Game Model of Oil Market with Dynamic Updating . . .	171
3.4.1	Introduction	171
3.4.2	Non-cooperative Game Model	174
3.4.3	Partially Cooperative Game Model	181
3.4.4	Historical Cooperative Agreement	190
Chapter 4	Class of Differential Games with Continuous Updating	199
4.1	Differential Game Model with Continuous Updating	199
4.1.1	Original Differential Game Model	199
4.1.2	Game Model with Continuous Updating	200
4.2	Non-cooperative Differential Game Model with Continuous Updating	202
4.2.1	Nash Equilibrium in Game with Continuous Updating	202
4.2.2	Hamilton-Jacobi-Bellman Equations with Continuous Updating	204
4.2.3	Pontryagin's Maximum Principle with Continuous Updating .	211
4.2.4	Autonomous Linear-Quadratic Case with Continuous Updating	219
4.2.5	Non-autonomous Linear-Quadratic Case with Continuous Updating	233
4.3	Cooperative Differential Game Model with Continuous Updating . .	259
4.3.1	Cooperative Game Model with Continuous Updating	259
4.3.2	Hamilton-Jacobi-Bellman Equations for Cooperative Game Model with Continuous Updating	272
4.3.3	Pontryagin's Maximum Principle for Cooperative Differential Game Model with Continuous Updating	282
4.3.4	Linear-Quadratic Case of Cooperative Differential Game with Continuous Updating	304
4.4	On Class of Non-Transferable Utility Differential Games with Continuous Updating	318
4.4.1	Pareto Optimal Strategies with Continuous Updating	318
4.4.2	Hamilton-Jacobi-Bellman Equation with Continuous Updating	319
4.4.3	Subgame Consistency with Continuous Updating	321
4.4.4	Differential Game of Non-renewable Resource Extraction . . .	322

4.5	Differential Game Model of Resource Extraction with Continuous and Dynamic Updating	329
4.5.1	Original Game Model of Resource Extraction	330
4.5.2	Game Model with Dynamic Updating	335
4.5.3	Game Model with Continuous Updating	338
4.5.4	Numerical Simulation	345
Chapter 5 Engineering Applications of Continuous Approach to Inverse Optimal Control Problems		349
5.1	Optimal Control with Continuous Updating	349
5.1.1	Original Optimal Control Problem	350
5.1.2	Optimal Control Problem with Continuous Updating	350
5.1.3	Optimal Control with Continuous Updating	351
5.2	Linear-Quadratic Optimal Control with Continuous Updating	352
5.2.1	Problem Formulation in Linear-Quadratic Case with Continuous Updating	353
5.2.2	Optimal Control with Continuous Updating for LQP	353
5.3	Inverse Optimal Control Problem with Continuous Updating	355
5.4	Simulation Results	356
5.4.1	Single-track and Steering Model	356
5.4.2	Optimal Control with Continuous Updating	358
5.4.3	Inverse Optimal Control Problem	360
5.4.4	Discussion	362
5.5	Conclusion	362
Conclusion		363
References		366

INTRODUCTION

Relevance of Thesis Topic

The main tasks of modern game theory are the design and analysis of the principles and concepts of optimal behavior of participants in various conflict controlled processes. Conflicts that actually take place develop in time, therefore, dynamic models are gaining special relevance. Differential games are convenient mathematical models for describing conflict-controlled processes in economics, ecology, management and other areas. The theory of differential games emerged as a separate branch of mathematics in the 1950s. One of the first works in the field of differential games is considered to be the work of R. Isaacs [1], in which the problem of intercepting an airplane was formulated in terms of states and controls guided missile, and also derived the fundamental equation for defining a solution. The contribution of R. Isaacs, together with the classic research of R. Bellman [2] created the basis for using the results of optimal control theory in conflict control problems with several participants. The first interesting results in the theory of differential games were obtained by L. Berkovits [3], G. Leitman [4], W. Fleming [5], A. Friedman [6] and others. A significant contribution to the further development of differential games was made by Russian scientists L.S. Pontryagin [7; 8], L.A. Petrosyan [9; 10], N.N. Krasovskiy [11; 12], B.N. Pshenichny [13], whose work was mainly related to differential pursuit games. The most important results in the field of substantiation and methods for finding solutions of antagonistic differential games were obtained in the works of N. N. and A.I. Subbotin [14; 15; 16]. In parallel, the theory began to develop nonantagonistic differential games in which the Nash equilibrium was used as the optimality principle [17]. Especially should to note the work of domestic scientists who have made a great contribution to the development nonantagonistic differential games: E. M. Vaisbord, R.V. Gamkrelidze, N.L. Grigorenko, V.I. Zhukovskiy, A.F. Kleimenov, A.F. Kononenko, A.V. Kryazhinskiy, A.B. Kurzhaniskii, V.N. Lagunova, N. Yu.

Lukoyanov, S.S. Kumkov, O. A. Malafeev, Mishchenko E. F., V. S. Patsko, N. N. Petrov, N.N. Subbotin, Tynyansky N.T., Chikriya A.A., Chistyakov S.V., Chentsov A.G. [18; 19; 20; 21; 22; 23; 24; 25; 26; 27; 28; 29; 30; 31; 32; 33; 34; 35; 36; 37; 38; 39; 40; 41; 42; 43; 44; 45; 46] and many others. Later, works using methods of differential games appeared and in the field of modeling conflict-driven economic processes, including in the tasks of environmental policy, optimal exploitation of natural resources, etc. (see, for example, [47; 48; 49]). This the area is developing quite quickly, a detailed analysis of the indicated works can be found in [50] (see also [51]). We should especially note the works of L.A. Petrosyan, V.V. Zakharov, N.A. Zenkevich, V.V. Mazalov, A.N. Rettiyeva, S. Jorgensen, E. Dokner, N. Long, G. Sorger, J. Zakkur, J. Kravchik, J. Filar and others [47; 52; 53; 54; 55; 56; 57; 58; 59; 60; 61; 62; 63; 64; 65; 66; 67; 68; 69; 70; 71; 72; 73; 74; 75], devoted to the use of the game-theoretic approach to solve the problem of environmental protection.

Most real-life conflict control processes evolve dynamically or continuously in time, and their participants continuously receive updated information and adapt. For such processes, an approach has been proposed to construct more realistic models, namely, the approach of dynamic updating or continuous updating. Game theory is a classical and fundamental tool that can be used to model the behavior of conflict processes with several participants. In particular, considering the dynamic processes that evolve in time, this is the theory of dynamic and differential games.

Classical dynamic and differential game models assume that the game structure does not change in the interval on which the game is defined, or that the players have complete information about the change of the game structure. However, if we consider a long-term process, these assumptions cannot cope with reality. If we want to model the behavior of players in a real-life scenario, it is necessary to take into account the specific behavior of player in the following sense: usually, even if the players have long-term information about the process, they would rather use short-term information to decide future actions. In the class of games with dynamic and continuous updating, it is supposed that the players have or use only information about motion equations and payoff functions defined on the interval with a length equal to the value of the information horizon. The information horizon defines how far the players can or want to forecast their actions. The information about the motion equations and payoff functions is updated as the current time evolves. In order to define the best possible behavior of players in this type of dynamic or

differential games, it is necessary to use a special approach, which is a subject of the Thesis. This approach has been developed in the papers of the thesis author and his coauthors.

Most real-life conflict control processes evolve dynamically or continuously in time, and their participants continuously receive updated information and adapt. For such processes, an approach was proposed to construct more realistic models, namely, games with dynamic updating [76; 77] and games with continuous updating [78; 79]. The fundamental models considered previously in the theory of differential games are related to the problems defined on: a fixed time interval (players have all the information on a closed time interval) [26], an infinite time interval with discounting (players have information on an infinite time interval) [80], and a random time interval (players have information on a given time interval, but the terminating instant is a random variable) [81]. Furthermore, one of the first works in the theory of differential games was devoted to the pursuer–evader game (the player’s payoff depends on when the opponent is caught) [82]. In all the above models and suggested solutions, it is assumed that players at the beginning of the game use all the information about the dynamics of the game (motion equations) and the preferences of players (payoff functions). However, this approach neglects the fact that, in many real-life processes, at the initial instant the players do not use all the information about the game. Thus, existing approaches cannot be directly used to construct a sufficiently large range of real-life game-theoretic models.

In game models with dynamic updating, it is assumed that players:

- 1 have information about the motion equations and payoff functions on a truncated time interval of a length \bar{T} , which is called the information horizon;
- 2 receive updated information about the motion equations and payoff functions at fixed time instants $t_j = t_0 + j\Delta t$, $j = 0, \dots, l$, $l = \frac{T-t_0}{\Delta t}$ and, consequently, dynamically adapt to this information.

In game models with continuous updating, it is assumed that players:

- 1 have information about the motion equations and payoff functions on a truncated time interval of a length \bar{T} , which is called the information horizon;
- 2 continuously receive updated information about the motion equations and payoff functions and, consequently, continuously adapt to this information.

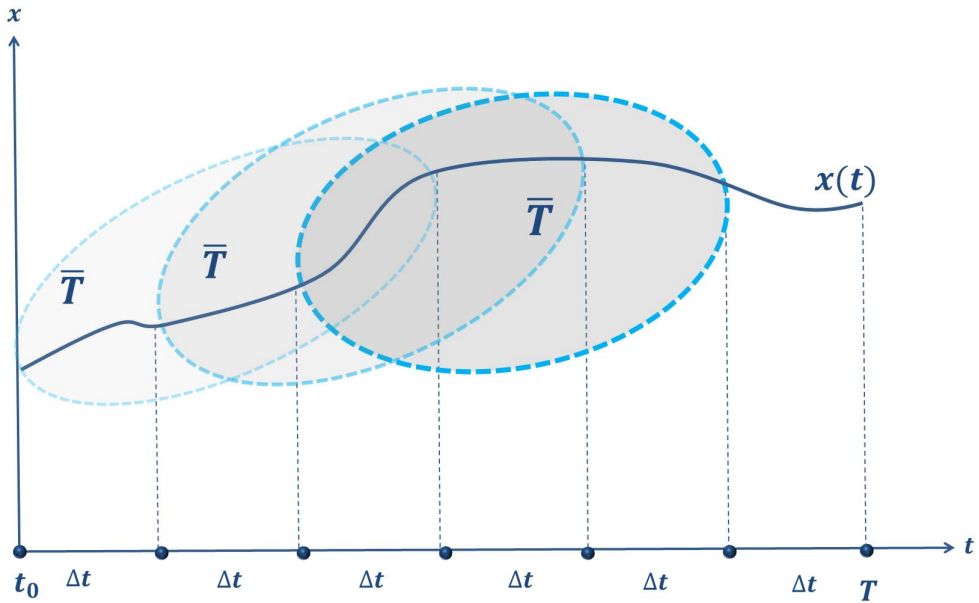


Figure 1: Each blue oval shows the information available to players on the interval $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, namely $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$, $l = \frac{T-t_0}{\Delta t}$.

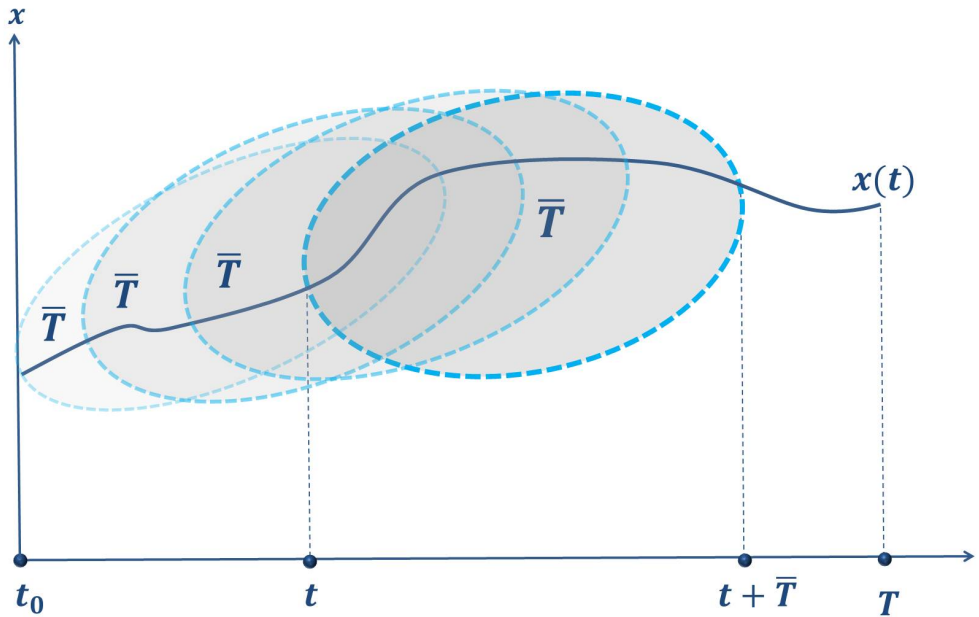


Figure 2: Each blue oval shows the information available to players at instant t , namely $[t, t + \bar{T}]$, where \bar{T} is the time horizon.

Obviously, it is difficult to obtain Nash equilibrium or cooperative strategies due to the lack of fundamental approaches to differential games and control problems with dynamic and continuous updating. Classical methods, such as dynamic programming and Hamilton-Jacobi-Bellman equation [2] or Pontryagin's Maximum Principle [83], do not allow to directly construct Nash equilibrium or cooperative strategies in problems with information updating. Taking into account

the assumptions described above, the two main problems arise:

- 1 How to define a solution concept similar to Nash equilibrium, cooperative and Pareto optimal strategies, a characteristic function, and a cooperative solution for the class of games with continuous updating?
- 2 How to derive the corresponding optimality conditions for Nash equilibrium, cooperative and Pareto optimal strategies, and the characteristic function?

Besides the problem of constructing in some sense optimal strategies, the corresponding trajectory for the class of games with dynamic and continuous updating, there are problems related to examining the time-consistency and subgame consistency property of a cooperative solution.

Note that there exists a wide range of real-life conflict control processes that can be modeled using the approach of dynamic and continuous updating. Therefore, it is necessary to develop this approach for a class of cooperative and noncooperative games, special cases of linear-quadratic games, and particular dynamic and differential game models to show the significance of the proposed approach. Also, it is important to use the proposed approach in engineering, environmental and socio-economic control problems to make it closer to applications. All these considerations show the relevance of the Thesis topic.

Degree of Problem Examination in Literature

The class of differential games with dynamic and continuous updating has some similarities with Model Predictive Control (MPC) theory, which is worked out within the framework of numerical optimal control [84; 85; 86; 87]. In the MPC approach, the current control action is achieved by solving a finite-horizon open-loop optimal control problem at each sampling instant. For linear systems, there exists a solution in explicit form, [88; 89]. However, in general, the MPC approach demands the solution of several optimization problems. Another related series of papers corresponds to the class of stabilizing control [90; 91; 92; 93]; here, similar approaches were considered for the class of linear-quadratic optimal control problems. However, in the current work and papers on the continuous updating approach, the main goal is different: to model players' behavior when information about the game is updated continuously in time.

The following list of papers are more closely related to the continuous updating approach: [94; 95; 96; 97; 98; 99]. The paper [96] was devoted to the moving horizon approach for dynamic game models. A new solution concept based on moving horizon control was introduced for nonzero-sum infinite-horizon differential games; the linear-quadratic case with open-loop and closed-loop strategies was considered. Another paper [94] was devoted to the equilibrium concept for both discrete- and continuous-time dynamic games and under varying (symmetrical and asymmetrical) modes of play. In the paper [95], a time-varying macroeconomic model was considered in which some of the parameters fluctuate exogenously, according to a Markov chain. The paper [97] developed a theoretical framework for the common business practice of rolling horizon decision making. The main idea of the presented approach is that the usefulness of rolling horizon methods is, to a great extent, implied by the fact that forecasting the future is a costly activity. The issue similar to the continuous updating approach was investigated in the papers [98] and [99], where the authors considered repeated games with sliding planning horizons.

Thesis Content

The thesis is devoted to developing theoretical models of conflict control (games with dynamic and continuous updating). Chapter 1 focuses on constructing and studying a new cooperative solution (the IDP-core) for differential games with dynamic and continuous updating. The IDP-core is defined based on the axioms of time-consistency and IDP-nondominance. The strong time-consistency of this solution is proved. An explicit formula for constructing the IDP-core is given. A numerical algorithm for analyzing the non-emptiness property is presented. Chapters 2 and 3 cover the application of the dynamic updating approach to dynamic and differential games. Chapter 4 is devoted to studying and applying the continuous updating approach to the class of differential games. In general, Chapters 2–4 deal with both cooperative and non-cooperative statements. For the first time, a wide range of optimality conditions for cooperative and Nash equilibrium strategies in the classes of open-loop and closed-loop strategies is presented. Methods for constructing the corresponding cooperative and equilibrium trajectory with dynamic and continuous updating are developed. In addition, an algorithm for constructing a characteristic function and a cooperative solution with dynamic and continuous

updating is presented for the cooperative game model. A connection between the cooperative solutions defined on truncated intervals determined by the information horizon is established. Other results obtained concern the relationship between the games with dynamic and continuous updating. The convergence of the strategies and trajectories with dynamic updating to the corresponding strategies and trajectories with continuous updating is proved. Chapter 5 is devoted to the inverse optimal control problem with continuous updating. As an application, the continuous updating approach is used for the inverse optimal control problem of the driver assistance system.

Chapter 1. In the paper [100] (later in [101]) the author of the thesis and his coauthors introduced the notion of a strong time-consistent subset of the Core. The authors constructed a new cooperative solution using the geometric approach and proved that it was a subset of Core and possessed a strong time-consistency property. Later on, this solution was called the IDP-core, and it can be constructed using a system of linear constraints for imputation distribution procedures. These conditions are defined for each time instant of a differential game. From the non-emptiness of a set described by these constraints, the non-emptiness of the corresponding set of IDPs at each time instant, it follows that the IDP-core is not empty. In the paper [102] the technique proposed in [103] was applied to study the non-emptiness of the IDP-core for each time instant. If for each time instant it is non-empty, we conclude that the IDP-core is non-empty. The obtained results can be used to construct the IDP-core and verify its non-emptiness as a numerical example. Also, a special case of this approach is presented for 3-player differential games. It is possible to analytically construct conditions for the non-emptiness of the IDP-core depending on the characteristic function. Furthermore, it is possible to define an analytical formula for IDP-core selectors, particularly the formula for imputation distribution procedures of IDP-core selectors. Later on, in the paper [104] the IDP-core was studied as a cooperative solution for differential games constructed solely using the axiom properties of time-consistency and IDP dominance. The publications cited above describe a fairly new approach to building time-consistent cooperative solutions. This approach uses time-consistency as the main axiomatic property to define a cooperative solution. This approach was the subject of research in [104]. Note that using time-consistency as an axiom for dynamic cooperative games, social choice theory, and mechanism design is promising. Another important property

discussed in [104] is called Imputation Distribution Procedure dominance (IDP-dominance). According to this property, the corresponding cooperative solution is constructed using the imputation distribution procedures that are undominated. We will say that an original IDP is undominated if there does not exist another IDP, a coalition S and a time instant such that the instantaneous payments corresponding to this IDP are higher for the players from S at this instant than their instantaneous payments corresponding to the original IDP.

The IDP-core can be effectively used especially for the class of games with dynamic and continuous updating. When constructing a cooperative solution in games with updating, the imputation distribution procedure corresponding to the cooperative solution defined on the time interval of the information horizon is applied at each step or time of modeling. Then, the IDP is employed at each step to construct the cooperative solution for the entire time interval of the game. However, the IDP-core is initially built using the IDP, which reduces the number of computational steps for obtaining the cooperative solution with dynamic updating. In the paper [105], the IDP-core was used for the differential game model with dynamic updating and a random horizon. In the paper [106], the connection between the IDP-core defined on the information horizon and the IDP-core defined for the differential game with dynamic updating was established.

Chapters 2 and 3. The class of games with dynamic updating was the first in the stream of works related to information updating. This class of games has mostly been studied in the papers of the author of the thesis and his coauthors, [76; 77; 105; 106; 107; 108; 109; 110; 111; 112; 113; 114; 115; 116]. These works laid the foundation for further study in the class of games with dynamic updating and assumed that the information about the motion equations and payoff functions was updated at discrete time instants, and the interval on which players knew the information was defined by the value of the information horizon.

The first work devoted to this class of games is [76]. This paper constructed a model of a cooperative differential game with a prescribed duration and dynamic updating. It introduced the concepts of a truncated subgame, resulting cooperative strategies, a conditionally cooperative trajectory, and a resulting cooperative solution. A theorem was proved, showing that an arbitrary resulting cooperative solution is Δt -time-consistent in this class of games. Games with dynamic updating, a stochastic forecast, and dynamic adaptation were presented in [77]. In the paper

[112], the approach presented above was applied to game models with an infinite horizon. The paper [105] was devoted to a special class of differential games with dynamic updating and a random horizon. The cooperative setting was considered, a novel cooperative solution (IDP-core) was used, and a new updating procedure for parameters of random horizon was presented. The paper [106] presented an extensive description of cooperative differential games with dynamic updating. An approach for constructing cooperative strategies, the corresponding trajectory, characteristic function, and related cooperative solution with dynamic updating were presented. The connection between the cooperative solutions on a truncated time interval and on the whole time interval was proved. The paper [108] studied the dependence of players' payoffs on the value of the information horizon. The paper [109] studied the properties of cooperative solution in the class of game models with dynamic updating. The paper [116] was devoted to constructing a special class of Hamilton-Jacobi-Bellman equations for a noncooperative dynamic game model defining the various types of information structure. The results obtained can be used for constructing models where players use different information structures. The papers [107; 110; 111] applied game models with dynamic updating to the oligopoly model of an oil market. Numerical modeling was performed in Matlab using the data on Brent and Light oil prices. The paper [113] studied the class of dynamic games with information updating, where cooperative and noncooperative settings were considered. The papers [114] and [115] applied a discrete-time dynamic updating model to advertising.

Chapter 4. The class of differential games with continuous updating was considered in the papers [78; 79; 117; 118; 119; 120; 121; 122; 123; 124], assuming that the updating process evolves continuously in time.

For the first time, this class of games was considered in the papers [78] and [79]. In [78], the system of Hamilton-Jacobi-Bellman equations was derived for the feedback Nash equilibrium with continuous updating. The paper [79] was devoted to the class of autonomous linear-quadratic differential games with continuous updating, where the feedback strategies were considered. The convergence of Nash equilibrium strategies and the related trajectory with dynamic updating and continuous updating was proved. Later on, in the paper [118], the class of cooperative differential games with transferable utility using Hamilton-Jacobi-Bellman equations was considered. The construction of characteristic function with

continuous updating and several related theorems were presented, and the property of strong time-consistency of the cooperative solution with continuous updating was proved. Another result related to Hamilton-Jacobi-Bellman equations with continuous updating was obtained for the class of cooperative differential games with nontransferable utility [119]. The paper [124] was devoted to a detailed study of a resource extraction game model for the cooperative and noncooperative game models. In the paper [120], an explicit form of the Nash equilibrium for the differential game with continuous updating was derived using Pontryagin's maximum principle. The paper [123] studied the cooperative setting using Pontryagin's Maximum principle. More papers related to the class of linear-quadratic differential games with continuous updating were devoted to open-loop Nash equilibrium strategies [117] and the cooperative setting in the characteristic function form with continuous updating [121]. The latest paper related to the linear-quadratic game with continuous updating [122] studied the non-autonomous case, where the updating model itself depends upon the current time t . Moreover, the convergence results were also obtained for the non-autonomous case.

Chapter 5. Other potentially important results are related to the formulation and solution of the inverse optimal control problem with continuous updating. Applied inverse optimal controls are becoming more and more important, as humans increasingly interact with various kinds of software and engineering interfaces. An important point to ensure an optimal response of an engineering software system to human actions is to determine the type and goals of the system user. Here, one of the approaches can be solving an inverse optimal control problem, in which the type of a human and his goals are defined by determining the integrand of his objective function. In this direction, the first paper [125] was devoted to applying the continuous updating approach to the class of inverse optimal control problems. The approach can be used for defining the driver's behavior model necessary in driver assistance systems. The purpose of this study is to determine the driver's behavioral profile for building an adaptive vehicle control assistance system in various driving modes (fuel economy, standard, sport). The approach of continuous updating allows for more accurately modeling human behavior by assuming the continuous updating of information about the dynamic system and the objective function.

Purpose and Tasks of Thesis

Purpose of the thesis is the development of the class of conflict control

models (differential and dynamic games with continuous and dynamic updating), construction of optimality conditions for non-cooperative and cooperative strategies with continuous and dynamic updating, the algorithm of construction of related trajectory, characteristic function, and cooperative solution with continuous updating. Most importantly, it is related to the study of special properties of cooperative solutions with continuous updating, including time-consistency and subgame consistency. Moreover, the purpose includes researching the correspondence between the dynamic updating and continuous updating cases by proving the convergence of the optimal strategies and corresponding trajectories and by studying the correspondence between the cooperative solutions on the updating interval and in the whole game with dynamic or continuous updating. Also the purpose includes the development of an application for the continuous updating approach and present specific simulation results.

The purpose mentioned above is achieved by solving the following **tasks of the thesis**:

- to develop the new class of conflict control models (noncooperative and cooperative dynamic and differential games with dynamic and continuous updating);
- to develop techniques for constructing Nash equilibrium, Pareto-optimal, and cooperative strategies as well as the characteristic function and cooperative solution with dynamic and continuous updating;
- to obtain optimality conditions for the cooperative and Nash equilibrium strategies with dynamic and continuous updating;
- to develop a new cooperative solution for the class of games with dynamic and continuous updating;
- to study the convergence of the solutions with dynamic updating to the solutions with continuous updating;
- to develop algorithms for constructing the characteristic function and cooperative solution with dynamic and continuous updating;
- to study the properties of δt strong time-consistency and strong time-consistency of cooperative solutions with dynamic and continuous updating;

- to examine the inverse optimal control problem with continuous updating.

Theoretical Novelty of Thesis

The scientific novelty of the thesis includes the following: the development of the theory of new conflict control models (dynamic and differential games with dynamic and continuous updating); the creation and adaptation of key optimality conditions for cooperative and equilibrium strategies in this class of games; the creation and study of special cooperative solutions for the class of games with dynamic and continuous updating; the study of the properties of cooperative and noncooperative solutions with dynamic and continuous updating; the study of the properties of time consistency and subgame consistency of cooperative solutions with continuous updating; the study of the convergence of cooperative and noncooperative solutions with dynamic and continuous updating. Also, the novelty of the thesis is the construction and examination of a certain class of inverse optimal control problems with continuous updating.

Theoretical and Practical Significance

Theoretical significance of the work is that it develops and describes not only the class of the games with continuous and dynamic updating but also presents the optimality conditions in the form of Hamilton-Jacobi-Bellman equations, Pontryagin's maximum principle, and conditions for linear-quadratic (closed-loop and open-loop based) Nash equilibrium and cooperative strategies with continuous and dynamic updating. The use of classical approaches is limited due to the nature and assumptions of continuous updating. The next theoretical results are related to studying the properties of time-consistency and subgame consistency of the cooperative solution with continuous and dynamic updating. It is proved that the cooperative solution with continuous (dynamic) updating is always strong time-consistent (Δt time-consistent) and subgame consistent (Δt subgame consistent) depending on the cooperation type. Other theoretical results are related to proving the connection between the solutions defined in subgames and the solutions defined on the whole interval, i.e., the solutions with continuous updating. The uniform and pointwise convergence of the optimal strategies with dynamic updating and the corresponding trajectories to those with continuous updating is proved for a special case of the linear-quadratic game model.

The theoretical significance of the work has been confirmed by the relevant projects supported by scientific foundations and St. Petersburg State University. The author was the head of

- project no. 18-71-00081 "Development of methodological tools for the mathematical theory of dynamically stable agreements for digital intelligent systems for formation of cooperation agreement as a step towards artificial intelligence in this field" (Russian Science Foundation)

and a participant of

- project no. 18-00-0072 COMFI "Development of game-theoretic methods for the analysis of conflicts and cooperation in a changing environment" (the Russian Foundation for Basic Research),
- project no. 9.38.205.2014 "New constructive approaches in nonsmooth analysis and nondifferentiable optimization and their applications" (St. Petersburg State University), and
- project no. 9.38.76.2011 "Development of the theory and constructive methods of nonsmooth analysis and their application to solving problems of mathematical modeling" (St. Petersburg State University).

Practical significance of the work is determined by the fact that the real-life conflict-controlled processes evolve continuously in time, and the participants of these processes usually are not or cannot use full information about it. Therefore, it is important to introduce dynamic and continuous game models with information updating relevant for modeling engineering, ecological and socio-economic systems. Another important practical contribution of the Thesis is the construction and research of a class of inverse optimal control problems with continuous updating. These problems can be used to analyze human's profile in human-machine engineering systems. The results are illustrated on the driver assistance system model and applied to the real driving data from a simulator located in the Institute of Control Systems, Karlsruhe Institute of Technology.

Research Methods

In the thesis, the author uses the **research methods** of dynamic and differential game theory (the construction of dynamic and differential games, subgames), noncooperative game theory (Nash equilibrium strategy profile, related trajectory), cooperative game theory (cooperative strategies, characteristic function, the time- and subgame consistency of cooperative solutions, the imputation distribution procedure), optimal control theory (dynamic programming, Pontryagin's maximum principle, and the linear-quadratic special case), mathematical analysis (the convergence of cooperative and noncooperative strategies and trajectories), optimization and probability theories (the distributions of random variables, stochastic processes).

Results Submitted for Defense

Main results obtained in the thesis and taken out to defense are as follows:

- New classes of conflict control models (dynamic games with information updating, differential games with dynamic updating, and differential games with continuous updating) have been described and studied. The cooperative transferable utility, non-transferable utility, and noncooperative statements have been examined. Techniques for constructing Nash equilibrium, Pareto-optimal, and cooperative strategies, as well as the characteristic function and the cooperative solution with dynamic and continuous updating have been presented.
- New optimality conditions for the cooperative and Nash equilibrium strategies, both in the class of open-loop and closed-loop strategies, have been presented in the form of Hamilton-Jacobi-Bellman equations and Pontryagin's Maximum principle with continuous updating. Conditions for a special case of linear-quadratic autonomous and nonautonomous continuous updating models have been presented, and related theorems have been proven. These optimality conditions have been adopted for solving classical differential game models in the case of continuous updating.
- New results have been established on the connection between the solutions of the game models with dynamic and continuous updating in the linear-quadratic autonomous and nonautonomous cases and one special classical differential game model of resource extraction. The convergence of the cooperative and Nash equilibrium strategies and related trajectories with

dynamic updating to the corresponding strategies and trajectories with continuous updating has been proven.

- New algorithms for constructing the characteristic function and cooperative solution with dynamic and continuous updating have been developed. The connection between the cooperative solutions defined on the truncated intervals given by the information horizon has been proven. The properties of δt strong time-consistency and strong time-consistency have been established for arbitrary cooperative solutions with dynamic and continuous updating, respectively.
- The new cooperative solution (the IDP-core) for differential games with dynamic and continuous updating has been defined based on the axioms of time-consistency and IDP-nondominance. The strong time-consistency property of this solution has been proven. An explicit formula for constructing the IDP-core has been presented and proven. A numerical algorithm to analyze the non-emptiness property has been presented. The IDP-core has been applied to the differential game model with dynamic updating.
- A new inverse optimal control problem with continuous updating has been formulated and solved. The continuous updating approach has been applied to the inverse optimal control problem of a driver assistance system.

Approval of Obtained Results

The results presented in the thesis were reported and **approved** at the following international conferences and seminars (with a sufficient number of foreign participants):

- CPS2012–2013,2015: XLIII, XLIV, XLVI international conferences "Control Processes and Stability", Faculty of Applied Mathematics and Control Processes of St. Petersburg State University, St. Petersburg (Russia);
- GTM2013–2020: 7th, 8th, 9th, 10th, 11th, 12th, 13th, 14th International Conferences on Game Theory and Management, St. Petersburg (Russia);

- The Second Russian Economic Congress, Suzdal (Russia), February 18–22, 2013;
- SING11 - GTM2015: European Meeting on Game Theory, St. Petersburg (Russia) , July 8–10, 2015;
- SING14,15: European Meetings on Game Theory, Bayreuth (Germany), 2018, and Turku (Finland), 2019;
- SCP2015: Stability and Control Processes Conference, in memory of Prof. Vladimir Zubov, St. Petersburg (Russia), October 5–9, 2015;
- SCP2016, 2020: XV International conference “Stability and Oscillations of Nonlinear Control Systems” (Pyatnitskiy’s conference), Moscow (Russia), 2016 and 2020 (online);
- GAM-SOC2015: Workshop on Game Theory and Social Choice, Budapest (Hungary), December 11–12, 2015;
- Constructive Nonsmooth Analysis and Related Topics (dedicated to the memory of V.F. Demyanov), St. Petersburg (Russia), May 22–27, 2017;
- ISDG18: 18th International Symposium on Dynamic Games and Applications, Grenoble (France), July 9–12, 2018;
- ISDG-GTM2019: International Meeting on Game Theory, St. Petersburg (Russia), July 3–5, 2019;
- MOTOR2019, 2020: Mathematical Optimization Theory and Operations Research, Yekaterinburg (Russia), 2019, and Novosibirsk (Russia), 2020 (online);
- IFAC2020: World Congress on Optimal Control, Berlin (Germany), July 12–17, 2020 (online);
- The 31st Stony Brook International Conference on Game Theory. Workshop on Differential Games. NOTE: the author of the thesis organized the workshop and made a presentation ([http : //www.gtcenter.org/?page = Workshops.html](http://www.gtcenter.org/?page=Workshops.html)), Stony Brook, New York (USA), July 20–24, 2020 (online);

Publications. The author published over 38 scientific papers, including 26 papers in the periodicals indexed by The Web of Science or Scopus databases. Among them, 27 papers are on the thesis topic [76; 77; 78; 79; 100; 101; 102; 104; 106; 107; 108; 109; 110; 112; 113; 114; 115; 116; 117; 118; 119; 120; 125; 126; 127; 128; 129], including 6 papers in the periodicals from the list of peer-reviewed journals recommended by the Higher Attestation Commission of the Russian Federation [100; 102; 112; 115; 127; 128], and 22 papers in the periodicals indexed by The Web of Science or Scopus databases [76; 77; 78; 79; 101; 104; 106; 107; 108; 109; 110; 112; 113; 114; 115; 117; 118; 119; 120; 125; 126; 129]. The main scientific results submitted for defense were published in the following peer-reviewed periodicals: [76; 77; 78; 79; 105; 106; 107; 108; 109; 110; 111; 112; 113; 114; 115; 116; 117; 118; 119; 120; 123; 126; 127; 128; 129] — item 1; [78; 79; 117; 118; 119; 120; 123; 125; 129] — item 2; [79] — item 3; [76; 77; 106; 109; 112; 113; 117; 118; 119; 123; 126; 127; 128] — item 4; [100; 101; 102; 104] — item 5; [125] — item 6.

The thesis consists of the Introduction, five chapters, the conclusion, the list of designations, the list of references, and appendices. The definitions, statements, theorems, examples, and remarks are numbered by the chapter and section where they are located. The corollaries are numbered by the theorems to which they refer. The volume of the work is 387 pages; the bibliography includes 237 titles; there are 100 figures, 8 tables and 0 appendices in the work.

Acknowledgments

The author **acknowledges** and expresses the deepest sense of gratitude to D.Sc., head of the department of Mathematical Modeling of Energy Systems SPbU, Prof. V. V. Zakharov for the long and patient work with the author, the support of whom helped substantially to finish the thesis.

The author considers it his pleasant duty to thank all the professors from the Faculty of Applied Mathematics and Control Processes (St. Petersburg State University) for being a scientific family who raised, hopefully, a scientist worthy of Zubov's heritage. The author is thankful to: Dr. M. Flad and J. Inga for giving the idea for applying the continuous updating approach and introducing to inverse optimal control problems; Prof. H. Gao for giving great scientific targets and supporting in nonscientific ventures; D. Chai for pushing hardly the nonscientific work and introducing to other scientific directions of interest at the current time; H.

Wei for friendly attitude, trust and guideness in scientific and personal development, for good example of a leader and a good person; Dr. Sci., Prof. A. U. Krylatov for giving a great example of an unstoppable spirit and supporting with life advice during the early time of the author's academic career; associate Prof. I. S. Blekanov for his openness and humanity, interesting academic discussions and joint ventures during latest time of thesis writing period; Dr. Li Yin for being a great friend, helping in various joint ventures and selflessly helping on the author's way; all dearest Phd students and undergraduates of the author who have been working hard with him side by side for years; L. Shi for his endless energy; Z. Wang for her strong spirit and will to become a PhD; J. Zou for his hard work, understanding, and the faith to his scientific supervisor; I. Kuchkarov for his hard work side by side for several years on the continuous updating approach; and others with whom the author has started to work recently and did not mention here. The author thanks G. V. Yanchenkova for being always on his side and protecting him during school time. The author deeply thanks his family for not giving up the hope for his son, for their constant help, support, advice in every venture, and a nontrivial interesting, and unforgettable life. Especially the author deeply thanks his mother Nina Y. Mitina for her endless care, for giving life and being a lifelong example of a beautiful and perfect hard-working woman. The author deeply thanks his father Leon A. Petrosyan whose unshakable wisdom and important advice always helped to resolve most of the problems on his way, for being the one with whom it was possible to discuss all the life and scientific problems and finally for the great example of a best man and a best scientist. Naturally, the author thanks his brother Evgeny V. Mitin and sister Naira L. Kazakova-Frehse for their special places in the author's life, for their advice, help, and great support in different and difficult periods of life. The author thanks the whole family of his sister, her husband Torsten Frehse, son Karl Frehse for being good friends, and finally, her daughter Tatjana V. Kazakova with whom the author has always taken an example in life.

Chapter 1

IDP-core: Novel Cooperative Solution for Differential Games

This chapter focuses on a new cooperative solution for dynamic and differential games, namely the IDP-core. A new approach to constructing solutions of dynamic and differential games is proposed, in which time-consistency is used as the main axiom for a cooperative solution. Another important property, and a new approach to constructing the IDP-core, is the property of IDP-dominance: it allows choosing undominated imputation distribution procedures to obtain a cooperative solution or a set of imputations. This approach shows the potential of time-consistency as a fundamental axiom in different application areas, such as social choice and mechanism design. The general procedure for determining a cooperative solution is also new: the IDP-core is built using imputation distribution procedures and not the imputation themselves. This feature makes the IDP-core a fruitful cooperative solution for games with dynamic and continuous updating.

When constructing a cooperative solution in games with updating, the imputation distribution procedure corresponding to the cooperative solution defined on the time interval of the information horizon is applied at each step or time of modeling. Then, the IDP is employed at each step to construct the cooperative solution for the entire time interval of the game. However, the IDP-core is initially built using the IDP, which reduces the number of computational steps for obtaining the cooperative solution with dynamic updating. In Chapter 3, the results of using the IDP-core for differential games with dynamic updating are presented [105], [106]. In the paper [105], the IDP-core was used for the differential game model with dynamic updating and a random horizon. In the paper [106], the connection between the IDP-core defined on the information horizon and the IDP-core defined for the differential

game with dynamic updating was established.

1.1 Problem Statement and Preliminary Information

1.1.1 Differential Game Model

The theory of differential games was developed as a separate class of applied mathematics in the 1950s. One of the first works in the field of differential games is the work of R. Isaacs [130] in which the notions of state, controls, and the problem of aircraft interception by a guided missile were formulated, and a fundamental equation for defining the solution derived. Other related works describing the class of zero-sum differential games, see [3; 82; 83; 131; 132]. To model situations where several participants (players) have different goals or payoff functions (and, therefore, act individually), noncooperative differential games were developed. As an optimality principle in noncooperative differential games, the Nash equilibrium in open-loop or closed-loop form is commonly used; see [26; 41; 80; 133]. Interesting applications of the theory of differential games can be found in the paper [134].

In this section, the general description of the differential game model is given. The main concepts of this model are the type of model, the payoff functions of players, motion equations, and the cooperative solution concept. The type of the game model reflects what we intend to do with the model. In this section we consider the cooperative game model. Here we need to define how to allocate the joint cooperative payoff among the players. The payoff functions of players define their objectives depending on the state of the game and strategies and are calculated on some specific time interval (in our case, a closed time interval). Motion equations define how the state of the game changes according to the strategies of players. In the case of the cooperative game model, the solution concept defines the exact type of imputation set that will be used to allocate the joint payoff among the players.

Consider an n -player differential game $\Gamma(x_0, T - t_0)$ with prescribed duration $T - t_0$ and initial condition x_0 . The game dynamics are defined by the system of differential equations

$$\begin{aligned} \dot{x} &= f(x, u_1, \dots, u_n), \quad x \in R^n, \quad u_i \in U_i \subset \text{comp}R^k, \quad t \in [t_0, T], \quad i = \overline{1, n}, \\ x(t_0) &= x_0, \end{aligned} \quad (1.1)$$

where the open-loop control $u_i(t)$ satisfying the system (1.1) is a strategy of player i , and $\text{comp}R^k$ is the compact set in the k -dimensional real space (k is an integer).

Let $N = \{1, \dots, n\}$ be the set of players. The payoff of player i is defined in the following way:

$$K_i(x_0, T - t_0; u_1, \dots, u_n) = \int_{t_0}^T h_i(x(\tau), u_1(\tau), \dots, u_n(\tau)) d\tau, \quad i = \overline{1, n}, \quad (1.2)$$

where $h_i(x, u_1, \dots, u_n) \geq 0$, $i = \overline{1, n}$ and $f(x, u_1, \dots, u_n)$ are integrable functions, and $x(t)$ is the solution of the system (1.1) with controls $u(t) = (u_1(t), \dots, u_n(t))$ involved. The strategy profile $u(t, x) = (u_1(t, x), \dots, u_n(t, x))$ is called admissible if the problem (1.1) has a unique and continuable solution on the interval $[t_0, T]$. In the thesis, the conditions of existence, uniqueness and continuability of A. F. Filippov [135] are used:

- 1 the right-hand side $f(t, x, u)$ of the motion equations (1.1) is continuous on the set $[t_0, T] \times X \times U_1 \times \dots \times U_n$,
- 2 the right-hand side $f(t, x, u)$ of the motion equations satisfies the Lipschitz conditions for x with a constant $k_1 > 0$ uniformly in u :

$$\|f(t, x', u) - f(t, x'', u)\| \leq k_1 \|x' - x''\|, \quad \forall t \in [t_0, T], \quad x', x'' \in X, u \in U,$$

- 3 there exists a constant k_2 such that the function $f(t, x, u)$ satisfies the condition

$$\|f(t, x, u)\| \leq k_2(1 + \|x\|), \quad \forall t \in [t_0, T], \quad x \in X, u \in U,$$

- 4 for any $t \in [t_0, T]$ and $x \in X$, the set

$$G(x) = \{f(t, x, u) | u \in U\}$$

is a convex and compact set from R^l .

1.1.2 Cooperative Differential Game Model

The theory of cooperative games examines how optimal parameters of cooperative and strategic agreements are to be determined. The main problem in the theory of

cooperative games with transferable utilities is to determine the allocation procedure for total payoff in cases when all players cooperate. A rule of how to allocate cooperative payoff among the players is called an imputation. In the theory of cooperative games with non-transferable utilities, the main problem is to define the agreement on strategies or a game outcome favorable to all players.

Within classical cooperative game theory with transferable utilities, numerous cooperative solutions or allocation rules were studied. One of them is the Core. The concept of Core was proposed by D. Gillis [136], which is a generalization of the contractual Edgeworth curve [137]. Edgeworth described a market with two products and two participants; here the Core is defined as a part of the Pareto front. The Core is the set of undominated imputations, each of which can be used as a game solution. An interesting discussion of the possibility of using cooperative solutions for applied problems is discussed in the paper [138].

It is also important to construct cooperative solutions for a class of dynamic and differential games. Solutions for such models can be used for modeling cooperative and strategic agreements where conditions are defined over a long time interval [139; 140]. A comprehensive description of dynamic cooperative games is presented in [141]. More cooperative differential and dynamic game models are considered in the papers [9; 142; 143; 144; 145].

In the cooperative differential game model with transferable utility, there are two problems:

- 1 Determination of a strategy set for players that maximizes the sum of their payoffs or determination of strategies corresponding to the cooperative behavior. These strategies $u^* = (u_1^*, \dots, u_n^*)$ are called optimal; the corresponding trajectory is called the cooperative trajectory and denoted by $x^*(t)$.
- 2 Determination of an allocation rule for the maximum joint payoff of players corresponding to the optimal strategies $u^*(t)$ and determination of the optimal trajectory $x^*(t)$. Namely, the determination of a cooperative solution as a subset of the imputation set.

Let $u^* = (u_1^*, \dots, u_n^*)$ be the vector of optimal strategies (open-loop controls)

for players, i.e., a set of controls that maximizes the joint payoff of players:

$$u^* = (u_1^*, \dots, u_n^*) = \arg \max_{u_1, \dots, u_n} \sum_{i=1}^n K_i(x_0, T - t_0; u_1, \dots, u_n). \quad (1.3)$$

Suppose that the maximum in (1.3) is achieved on the set of admissible strategies.

To determine how to allocate the maximum total payoff among players, it is necessary to define the notion of the characteristic function of coalition $S \subseteq N$. The characteristic function shows the strength of a coalition and thus allows the contribution of players to each coalition to be taken into account.

Suppose that in the game $\Gamma(x_0, T - t_0)$, a characteristic function $V(S; x_0, T - t_0)$, $S \subseteq N$, is constructed in any relevant way (for example, like in [146]). Assume that the superadditivity conditions are satisfied:

$$\begin{aligned} V(S_1 \cup S_2; x_0, T - t_0) &\geq V(S_1; x_0, T - t_0) + V(S_2; x_0, T - t_0), \\ \forall S_1, S_2 \subseteq N, S_1 \cap S_2 &= \emptyset. \end{aligned}$$

We denote by $L(x_0, T - t_0)$ the set of imputations [147] in the game $\Gamma(x_0, T - t_0)$:

$$\begin{aligned} L(x_0, T - t_0) = \left\{ \xi(x_0, T - t_0) = (\xi_1(x_0, T - t_0), \dots, \xi_n(x_0, T - t_0)) : \right. \\ \sum_{i=1}^n \xi_i(x_0, T - t_0) = V(N; x_0, T - t_0), \\ \left. \xi_i(x_0, T - t_0) \geq V(\{i\}; x_0, T - t_0), i \in N \right\}, \end{aligned}$$

where $V(\{i\}; x_0, T - t_0)$ is a value of the characteristic function $V(S; x_0, T - t_0)$ for coalition $S = \{i\}$.

We denote by $M(x_0, T - t_0)$ an arbitrary cooperative solution or subset of the imputation set $L(x_0, T - t_0)$:

$$M(x_0, T - t_0) \subseteq L(x_0, T - t_0).$$

Suppose that at the beginning of game $\Gamma(x_0, T - t_0)$ at the instant t_0 , the players agreed to select a subset of $L(x_0, T - t_0)$ or some cooperative solution. However, suppose that at some instant \bar{t} , the players decided to reconsider the chosen cooperative solution, or decided to reconsider the allocation rule for their

cooperative payoff. To model their behavior, it is necessary to define the notion of a subgame $\Gamma(x^*(t), T-t)$ along the cooperative trajectory $x^*(t)$ starting at an instant $t \in [t_0, T]$.

For each subgame $\Gamma(x^*(t), T-t)$, $t \in [t_0, T]$, along the trajectory $x^*(t)$, we define the superadditive characteristic function $V(S; x^*(t), T-t)$, $S \subseteq N$ in the same way as it was done for the original game $\Gamma(x_0, T-t_0)$:

$\forall S, A \subseteq N, S \cap A = \emptyset :$

$$V(S \cup A; x^*(t), T-t) \geq V(S; x^*(t), T-t) + V(A; x^*(t), T-t). \quad (1.4)$$

It is also possible to define the notion of an imputation $\xi(x^*(t), T-t)$ for a subgame $\Gamma(x^*(t), T-t)$ along the cooperative trajectory $x^*(t)$, $t \in [t_0, T]$. The set of all possible imputations in the subgame $\Gamma(x^*(t), T-t)$ is denoted by $L(x^*(t), T-t)$, $t \in [t_0, T]$:

$$L(x^*(t), T-t) = \left\{ \xi(x^*(t), T-t) = (\xi_1(x^*(t), T-t), \dots, \xi_n(x^*(t), T-t)) : \right. \\ \left. \sum_{i=1}^n \xi_i(x^*(t), T-t) = V(N; x^*(t), T-t), \right. \\ \left. \xi_i(x^*(t), T-t) \geq V(\{i\}; x^*(t), T-t), i \in N \right\}. \quad (1.5)$$

The superaditivity property (1.4) for the characteristic function $V(S; x^*(t), T-t)$ guarantees the non-emptiness of the imputation set $L(x^*(t), T-t)$, $t \in [t_0, T]$. The cooperative solution of the subgame $\Gamma(x^*(t), T-t)$ is denoted by $M(x^*(t), T-t)$.

1.1.3 Core

In cooperative game theory, the main problem is "fair" allocation of the maximum joint payoff $V(N; x_0, T-t_0)$ among the players from the grand coalition $N = \{1, \dots, n\}$.

Suppose that players in the cooperative differential game $\Gamma(x_0, T-t_0)$ (the subgame $\Gamma(x^*(t), T-t)$, $t \in [t_0, T]$, along the cooperative trajectory $x^*(t)$) made an agreement on the allocation rule $\xi(x_0, T-t_0)$ (imputation $\xi(x^*(t), T-t)$), where none of the imputations dominates $\xi(x_0, T-t_0)$ ($\xi(x^*(t), T-t)$) [147]. Such an allocation rule is stable in the sense that there does not exist an imputation that would be better for each coalition at every time instant $t \in [t_0, T]$.

Definition 1.1.1 We call the set of undominated imputations in the cooperative differential game $\Gamma(x^*(t), T - t)$ by the Core and denote it by $C(x^*(t), T - t)$, $t \in [t_0, T]$.

The following theorem holds:

Theorem 1.1.1 An imputation $\xi(x^*(t), T - t)$ belongs to the Core $C(x^*(t), T - t)$ if and only if for all $S \subseteq N$, the following inequalities are satisfied:

$$V(S; x^*(t), T - t) \leq \sum_{i \in S} \xi_i(x^*(t), T - t), \quad t \in [t_0, T].$$

1.1.4 Non-emptiness of Core in Static Games

It is important to study the non-emptiness property of a cooperative solution, which is to determine the conditions under which the cooperative solution is not empty since its applicability depends on the wideness of the class of games to which this solution can be applied. G. Scarf [148] showed that the Core is not empty for the class of convex games in characteristic function form. The characteristic function is a function of a coalition or subset of players in the game that shows its profit. Generalization of Scarf's results can be found in the papers of L. Biller [149] and L. Shapley [150]. Necessary and sufficient conditions for the non-emptiness of Core were formulated by Bondareva [151] and Shapley [152], where the main role of proof is the concept of a balanced game. Unfortunately, based on this concept, it is impossible to apply a constructive method for choosing specific imputations from the Core. V. Zakharov in [103] proposed the necessary and sufficient conditions for the non-emptiness of Core, which simplify the test for a single-point solution (imputation), such as the Shapley value, the Banzhaf power index, and others, whenever they belong to the Core. In [153] and [154], based on this approach, geometric properties of several cooperative solutions were investigated. This approach implies that the non-emptiness property of Core can be formulated by a linear programming problem constructed using the values of the characteristic function.

These are the main results concerning the nonemptiness conditions of the Core in static games. Necessary and sufficient conditions for the non-emptiness of the Core were formulated by O. Bondareva [151] and L. Shapley [152]. These conditions are based on the concept of a balanced game, but the application of this approach for a specific game model is difficult.

In the paper [155], G. Owen showed that in the game (N, v) there exists a non-empty Core if and only if the optimal value of the linear programming problem

$$\begin{aligned} \sum_{i \in N} \xi_i &\longrightarrow \min \\ \sum_{i \in S} \xi_i &\geq v(S), \quad \forall S \subseteq N, S \neq \emptyset \end{aligned}$$

is equal to $v(N)$.

The papers [103; 153; 154] also involved linear programming problems for the Core's non-emptiness. Consider the following linear programming problem:

$$\begin{aligned} \sum_{i \in N} \xi_i &\longrightarrow \min \\ \sum_{i \in S} \xi_i &\geq v(S), \quad \forall S \subseteq N, S \neq N, \emptyset. \end{aligned} \tag{1.6}$$

Suppose that $\xi^0 = (\xi_1^0, \dots, \xi_n^0)$ is some arbitrary optimal solution of the linear programming problem (1.6). The set of all optimal solutions of the optimization problem (1.6) is denoted by $X^0(v)$. In [103] it was shown that the necessary and sufficient conditions of the Core's non-emptiness can be formalized in the following way:

Theorem 1.1.2 *The Core in a cooperative game with transferable utility (N, v) is nonempty if and only if the following inequality is satisfied:*

$$\sum_{i \in N} \xi_i^0 \leq v(N), \tag{1.7}$$

where $\xi^0 \in X^0(v)$ is the solution of the linear programming problem (1.6).

1.1.5 Time-consistency of Cooperative Solution and Imputation Distribution Procedure

The theory of cooperative differential games helps to devise cooperative agreement terms, namely: cooperative strategies, the corresponding trajectory, the joint payoff along the cooperative trajectory, allocation rules of the joint payoff among the players, and the time-consistency property of the solution. A natural approach for researching cooperative differential games is an attempt at transferring the

results of classical static cooperative theory [156] to the theory of differential games. However, to use the results of classical theory, it is also necessary to study the time-consistency and strong time-consistency properties of cooperative solutions. The time-consistency of a cooperative solution shows that it is not beneficial for the players to deviate from the chosen cooperative solution during the game. The time-consistency problem of Nash's bargaining solution in a dynamic bargaining game model was discussed in [157]. The use of time-inconsistent cooperative solutions in economics, ecology, and management makes these solutions unfeasible because players might find it profitable to reconsider the cooperative solution.

L. Petrosyan first gave a mathematical formulation for the concepts of time-consistency [158] and strong time-consistency in cooperative differential games [159]. To prevent a violation of the cooperative agreement terms or the time-consistency property, a payment scheme was proposed in [160], which is called the Imputation Distribution Procedure (IDP). The notion of strong time-consistency was formulated in [161]. The recent papers [145; 162; 163; 164] are devoted to the study of the time-consistency property of cooperative solutions. Lately, in the paper [165] the authors first introduced the notion of subgame consistency for dynamic cooperative differential games with non-transferable utility. In the same paper, the authors derive subgame consistent solutions for a class of cooperative stochastic differential games with non-transferable utility. Another technique for constructing a subgame consistent solution is presented in [166].

The main approach for solving the time inconsistency problem of a cooperative solution in a differential game is the imputation distribution procedure (IDP), proposed in [167]. In that paper, the imputation distribution procedure was defined as a vector function for a fixed imputation. In this section, we consider another approach that generalizes the notion of IDP.

Assume that IDPs in the cooperative differential game $\Gamma(x_0, T - t_0)$ are integrable vector functions that constitute some imputation from the imputation set:

$$\beta(t) : \int_{t_0}^T \beta(\tau) d\tau \in L(x_0, T - t_0) \quad (1.8)$$

or

$$\int_{t_0}^T \beta_i(\tau) d\tau \geq V(\{i\}; x_0, T - t_0), \quad i \in N,$$

$$\sum_{i \in N} \int_{t_0}^T \beta_i(\tau) d\tau = V(N; x_0, T - t_0).$$

Therefore, in the above definition, IDP is not based on the imputation itself but generates it. We define also the so-called corresponding IDP, the concept of which is close to the original definition of IDP in the paper [167].

Definition 1.1.2 *An integrable function $\beta(t) = (\beta_1(t), \dots, \beta_n(t))$, $t \in [t_0, T]$, is called a corresponding imputation distribution procedure (IDP) for $\xi(x_0, T - t_0) \in L(x_0, T - t_0)$ if the following equalities hold:*

$$\xi_i(x_0, T - t_0) = \int_{t_0}^T \beta_i(\tau) d\tau, \quad i \in N. \quad (1.9)$$

Actually, the corresponding IDP $\beta(t)$ depends on $\xi(x_0, T - t_0)$ and is not unique for this imputation. We can represent it in the form

$$\beta(t) = \beta(t, \xi(x_0, T - t_0))$$

or

$$\beta_i(t) = \beta_i(t, \xi_i(x_0, T - t_0)), \quad i \in N.$$

From (1.9) we have for $t \in [t_0, T]$, $i \in N$:

$$\xi_i(x_0, T - t_0) = \int_{t_0}^t \beta_i(\tau) d\tau + \int_t^T \beta_i(\tau) d\tau$$

or

$$\int_t^T \beta_i(\tau) d\tau = \xi_i(x_0, T - t_0) - \int_{t_0}^t \beta_i(\tau) d\tau.$$

That is, at an instant t the IDP divides the imputations in two parts: the payoffs to player i , which are obtained on the intervals $[t_0, t]$ and $(t, T]$.

Definition 1.1.3 *A cooperative solution $M(x_0, T - t_0)$ in the game $\Gamma(x_0, T - t_0)$ is called time-consistent if for each imputation $\xi(x_0, T - t_0) \in M(x_0, T - t_0)$, there*

exists a corresponding IDP $\beta(t) = \beta(t, \xi(x_0, T - t_0))$ such that

$$\int_t^T \beta(\tau) d\tau \in M(x^*(t), T - t), \quad t \in [t_0, T] \quad (1.10)$$

or equivalently,

$$\xi(x_0, T - t_0) - \int_{t_0}^t \beta(\tau) d\tau \in M(x^*(t), T - t), \quad t \in [t_0, T]. \quad (1.11)$$

Note that from condition (1.10) we have the following equality:

$$\sum_{i \in N} \int_t^T \beta_i(\tau) d\tau = V(N; x^*(t), T - t), \quad t \in [t_0, T]. \quad (1.12)$$

It is obvious that if $M(x^*(t), T - t) \neq \emptyset$ for $\forall t \in [t_0, T]$, then for any function $\xi(x^*(t), T - t) \in M(x^*(t), T - t)$ ($\xi(x^*(t_0), T - t_0) = \xi(x_0, T - t_0)$) differentiable with respect to t , the IDP $\beta(t)$ can be defined by

$$\begin{aligned} \beta(t) &= -\frac{d}{dt} \xi(x^*(t), T - t), \quad t \in [t_0, T], \quad i \in N, \\ \xi(x^*(t_0), T - t_0) &= \xi(x_0, T - t_0). \end{aligned} \quad (1.13)$$

Then the imputation $\xi(x_0, T - t_0)$ is defined by the formula

$$\xi(x_0, T - t_0) = \int_{t_0}^t \beta(\tau) d\tau + \xi(x^*(t), T - t), \quad t \in [t_0, T].$$

We define an imputation in the current cooperative game $\Gamma(x^*(t), T - t)$ with a characteristic function $V(S; x^*(t), T - t)$ that corresponds to a given IDP $\beta(t) = \beta(t, \xi(x_0, T - t_0))$ as

$$\xi(x^*(t), T - t) = \int_t^T \beta(\tau) d\tau. \quad (1.14)$$

From Definition 1.1.3 we have

$$\xi(x^*(t), T - t) \in M(x^*(t), T - t). \quad (1.15)$$

We will call the imputation (1.14) the dynamic imputation generated by the corresponding IDP $\beta(t) = \beta(t, \xi(x_0, T - t_0))$.

1.2 IDP-core and Dominance of Imputation Distribution Procedures

Consider the game evolution at an instant $t \in (t_0, T)$. Suppose that at the instant t_0 , players agreed to realize an imputation $\xi(x_0, T - t_0) = (\xi_1(x_0, T - t_0), \dots, \xi_n(x_0, T - t_0))$. Then, according to the corresponding IDP $\beta(t)$, until the instant t player $i \in N$ will obtain the payoff

$$\int_{t_0}^t \beta_i(\tau) d\tau.$$

However, for some players, the IDP $\beta(t, \xi(x_0, T - t_0))$ would not be beneficial if there exists another imputation distribution procedure $\bar{\beta}(t, \xi(x_0, T - t_0))$, according to which player i on the interval $[t_0, t]$ will obtain a higher payoff:

$$\int_{t_0}^t \bar{\beta}_i(\tau) d\tau > \int_{t_0}^t \beta_i(\tau) d\tau. \quad (1.16)$$

In this case, the IDP $\beta(t)$ may be considered as less beneficial for player i at least on the interval $[t_0, t]$. It is important to note that the notion of IDP-dominance can be applied to imputation distribution procedures not necessarily defined for a unique imputation. As the IDP defines how dynamic imputation is to be constructed, it also makes sense to consider the notion of IDP-dominance not only for a fixed imputation.

1.2.1 Dominance of Imputation Distribution Procedures

In this section, we consider the IDP $\beta(t)$ defined by formula (1.8). Suppose that the function $V(S; x^*(t), T - t)$, $S \subseteq N$, is continuously differentiable with respect to $t \in [t_0, T]$. We define the function $U(S; x^*(t), T - t)$ in the following way:

$$U(S; x^*(t), T - t) = -\frac{d}{dt} V(S; x^*(t), T - t), \quad t \in [t_0, T], \quad S \subseteq N. \quad (1.17)$$

Definition 1.2.1 *An IDP $\beta(t)$ dominates an IDP $\bar{\beta}(t)$ by coalition $S \subseteq N$ at an instant $\bar{t} \in [t_0, T]$ (this fact will be denoted by $\beta(t) \stackrel{S, \bar{t}}{>} \bar{\beta}(t)$) if the following*

inequalities hold:

$$\begin{aligned} \beta_i(\bar{t}) &> \bar{\beta}_i(\bar{t}), \quad i \in S, \\ \sum_{i \in S} \beta_i(\bar{t}) &\leq U(S; x^*(\bar{t}), T - \bar{t}). \end{aligned} \quad (1.18)$$

Definition 1.2.2 *An IDP $\beta(t)$ is undominated if at any $\bar{t} \in [t_0, T]$, there does not exist $\bar{\beta}(t)$ that dominates $\beta(t)$ by coalition $S \subseteq N$:*

$$\bar{\beta}(t) \stackrel{S, \bar{t}}{\nrightarrow} \beta(t), \quad \forall \bar{\beta}(t), S. \quad (1.19)$$

1.2.2 IDP-core

In the paper [101], the applicant and his coauthors first introduced and treated a subset of the imputation set in a cooperative differential game, which was named the Subcore. This subset was designed using a set of imputation distribution procedures satisfying the system of inequalities and equalities. This approach is not classical for the theory of differential games since it uses IDPs for imputations, not vice versa. Based on the Subcore, in the paper [102] the applicant and his coauthors redefined this notion for the dynamic case, named it the IDP-core, and formulated necessary and sufficient conditions of the existence of the IDP-core along the cooperative trajectory of the game. As demonstrated below, the solution concept for the IDP-core is defined by introducing the notion of IDP dominance and using the time consistency properties or axioms defined above. It is proved that the IDP-core has the necessary and sufficient conditions for a dynamic imputation when defined by the system of inequalities introduced in the papers [100; 101].

Suppose that players in the game $\Gamma(x_0, T - t_0)$ agreed on an allocation rule for the total payoff of the grand coalition N (an imputation $\xi(x_0, T - t_0)$) using the cooperative solution of the IDP-core:

Definition 1.2.3 *By the dynamic IDP – core($x^*(t), T - t$) along the cooperative trajectory $x^*(t)$, $t \in [t_0, T]$ (IDP – core($x_0, T - t_0$)), we call a solution of the cooperative differential game $\Gamma(x^*(t), T - t)$ ($\Gamma(x_0, T - t_0)$) that includes all time-consistent imputations generated by undominated IDPs $\beta(\tau)$, $\tau \in [t, T]$ (1.8), $t \in$*

$[t_0, T]$ $(\beta(t), t \in [t_0, T])$:

$$\begin{aligned} IDP - core(x^*(t), T - t) = \left\{ \xi(x^*(t), T - t) = \int_t^T \beta(\tau) d\tau : \right. \\ \left. \xi(x^*(t), T - t) \text{ and the corresponding } \beta(\tau) \text{ satisfies (1.10),} \right. \\ \left. \nexists \bar{\beta}(t), S, \bar{t} : \bar{\beta}(t) \underset{S, \bar{t}}{>} \beta(t) \right\}. \end{aligned} \quad (1.20)$$

Note that the $IDP - core(x^*(t), T - t)$ includes imputations from the Core $C(x^*(t), T - t)$ of the cooperative game $\Gamma(x_0, T - t_0)$ for which there exists a corresponding undominated IDP and this IDP generates a dynamic imputation belonging to $C(x^*(t), T - t)$ for each $t \in [t_0, T]$.

Theorem 1.2.1 *Let $C(x^*(t), T - t)$ be not empty for any $t \in [t_0, T]$. A dynamic imputation $\xi(x^*(t), T - t)$ in the cooperative differential game $\Gamma(x_0, T - t_0)$ belongs to the dynamic $IDP - core(x^*(t), T - t)$ if and only if for the corresponding $\beta(t) = \beta(t, \xi(x_0, T - t_0))$ the following conditions are satisfied $\forall t \in [t_0, T]$:*

$$\sum_{i \in S} \beta_i(t) \geq U(S; x^*(t), T - t), \quad \forall S \subset N, \quad (1.21)$$

$$\sum_{i \in N} \beta_i(t) = U(N; x^*(t), T - t). \quad (1.22)$$

Proof Sufficiency. Let conditions (1.21)-(1.22) hold for the corresponding IDP $\beta(t) = \beta(t, \xi(x_0, T - t_0))$ at any $t \in [t_0, T]$. By integrating (1.21)-(1.22) on the interval we obtain

$$\begin{aligned} \sum_{i \in S} \xi(x^*(t), T - t) \geq V(S; x^*(t), T - t), \quad \forall S \subset N, \\ \sum_{i \in N} \xi(x^*(t), T - t) = V(N; x^*(t), T - t). \end{aligned} \quad (1.23)$$

This means that

$$\xi(x_0, T - t_0) = \int_{t_0}^T \beta(t) dt \in C(x^*(t_0), T - t_0) = C(x_0, T - t_0).$$

Let us show that $\beta(t) = \beta(t, \xi(x_0, T - t_0))$ is undominated. It can be proved by

contradiction. Suppose that for some $\bar{S} \subset N$, there exist $\bar{t} \in [t_0, T]$ and $\bar{\beta}(t)$ such that:

$$\begin{aligned} \xi(x_0, T - t_0) &= \int_{t_0}^T \bar{\beta}(t) dt \in L(x_0, T - t_0), \\ \bar{\beta}_i(\bar{t}) &> \beta_i(\bar{t}), \quad i \in \bar{S}, \\ \sum_{i \in \bar{S}} \bar{\beta}_i(\bar{t}) &\leq U(\bar{S}; x^*(\bar{t}), T - \bar{t}). \end{aligned} \quad (1.24)$$

Therefore,

$$\sum_{i \in \bar{S}} \beta_i(\bar{t}) < U(\bar{S}; x^*(\bar{t}), T - \bar{t}).$$

This inequality contradicts (1.21). Thus, $\beta(t) = \beta(t, \xi(x_0, T - t_0))$ is undominated on the interval $[t_0, T]$. Note that $\beta(t)$ is undominated on any subinterval $[\tau, T]$, $\tau \in [t_0, T]$, in the subgame $\Gamma(x^*(\tau), T - \tau)$.

For $S = \{i\}$, condition (1.21) is represented in the form

$$\beta_i(t) \geq U(\{i\}; x^*(t), T - t), \quad i \in N.$$

Integrating these inequalities and equality (1.22) taking into account (1.9) and (1.17), we obtain:

$$\begin{aligned} \xi_i(x^*(t), T - t) &= \int_t^T \beta_i(\tau) d\tau \geq V(\{i\}; x^*(t), T - t), \quad i \in N, \\ \sum_{i \in N} \xi_i(x^*(t), T - t) &= \sum_{i \in N} \int_t^T \beta_i(\tau) d\tau = V(N; x^*(t), T - t). \end{aligned} \quad (1.25)$$

Thus, the dynamic payoff

$$\xi(x^*(t), T - t) = \int_t^T \beta(\tau) d\tau$$

is the imputation in the current game $\Gamma(x^*(t), T - t)$ generated by the corresponding IDP $\beta(t) = \beta(t, \xi(x_0, T - t_0))$.

Due to the non-emptiness of $C(x^*(t), T - t)$ for any $t \in [t_0, T]$, the IDP $\beta(t) =$

$\beta(t, \xi(x_0, T - t_0))$ satisfying (1.21)-(1.22) generates the payoff vector

$$\xi(x^*(t), T - t) = \int_t^T \beta(\tau) d\tau,$$

which belongs to $C(x^*(t), T - t)$. Thus, $\xi(x^*(t), T - t)$ satisfies (1.10), and therefore the imputation $\xi(x_0, T - t_0)$ is time-consistent and lies in $IDP - core(x_0, T - t_0)$.

Necessity. Let an imputation $\xi(x_0, T - t_0)$ in the cooperative differential game $\Gamma(x_0, T - t_0)$ belong to $IDP - core(x_0, T - t_0)$. Therefore, by Definition 1.2.3, the imputation $\xi(x_0, T - t_0)$ generated by the undominated IDP $\beta(t) = \beta(t, \xi(x_0, T - t_0))$ belongs to $C(x_0, T - t_0)$ and is time-consistent. Due to Definition 1.2.3, the time-consistency of the imputation $\xi(x_0, T - t_0)$ means that there exists an IDP $\beta(t)$ such that

$$\xi(x^*(t), T - t) = \int_t^T \beta(\tau) d\tau \in C(x^*(t), T - t), \quad t \in [t_0, T].$$

Let us show that this inclusion holds for the undominated $\beta(t) = \beta(t, \xi(x_0, T - t_0))$ corresponding to the imputation $\xi(x_0, T - t_0)$. Suppose that this is not true. Then there exists coalition $\bar{S} \subset N$ and some instant $\bar{t} \in [t_0, T]$ such that the following inequality holds:

$$\sum_{i \in \bar{S}} \int_{\bar{t}}^T \beta_i(t) dt < V(\bar{S}; x^*(\bar{t}), T - \bar{t}). \quad (1.26)$$

Note that $\beta(\bar{t}) = \beta(\bar{t}, \xi(x_0, T - t_0))$ belongs to the set of undominated imputations in a cooperative game with the characteristic function $U(S; x^*(\bar{t}), T - \bar{t})$, $S \subseteq N$. Therefore, as follows from Theorem 1.1.1, for $\bar{S} \subset N$ the following has to be fulfilled:

$$\sum_{i \in \bar{S}} \beta_i(\bar{t}) \geq U(\bar{S}; x^*(\bar{t}), T - \bar{t}). \quad (1.27)$$

Integrating this inequality on the interval $[\bar{t}, T]$, we obtain

$$\sum_{i \in \bar{S}} \int_{\bar{t}}^T \beta_i(t) dt \geq V(\bar{S}; x^*(\bar{t}), T - \bar{t}), \quad (1.28)$$

which contradicts (1.27). Thus, the theorem is proved. \square

Proposition 1.2.4 *If $C(x^*(t), T - t)$ is empty for some $t = \bar{t} \in [t_0, T]$, then $IDP - core(x^*(t_0), T - t_0)$ is empty.*

Proof Suppose that we can find a time-consistent imputation $\xi(x_0, T-t_0) \in \text{IDP-core}(x^*(t_0), T-t_0)$ and a corresponding undominated IDP $\beta(t)$ that satisfies (1.21) and (1.22). Integrating inequalities (1.21) and (1.22), we obtain:

$$\begin{aligned} \sum_{i \in S} \int_{\bar{t}}^T \beta_i(\tau) d\tau &\geq V(S; x^*(\bar{t}), T - \bar{t}), \quad S \subset N, \\ \sum_{i \in N} \int_{\bar{t}}^T \beta_i(\tau) d\tau &= V(N; x^*(\bar{t}), T - \bar{t}). \end{aligned} \quad (1.29)$$

Thus,

$$\xi(x^*(\bar{t}), T - \bar{t}) = \int_{\bar{t}}^T \beta(\tau) d\tau \in C(x^*(\bar{t}), T - \bar{t}). \quad (1.30)$$

This result contradicts the emptiness of $C(x^*(\bar{t}), T - \bar{t})$. Therefore, $\text{IDP-core}(x^*(t_0), T-t_0)$ is empty. The proposition is proved. \square

Proposition 1.2.5 *If $C(x^*(t), T-t)$ is not empty for any $t \in [t_0, T]$, then $\text{IDP-core}(x^*(t_0), T-t_0) = C(x^*(t_0), T-t_0)$.*

Proof Consider an imputation $\xi(x_0, T-t_0) \in C(x^*(t_0), T-t_0)$ that does not belong to $\text{IDP-core}(x^*(t_0), T-t_0)$. That is, $\xi(x_0, T-t_0)$ belongs to $C(x^*(t_0), T-t_0)$, but is time inconsistent. According to the time consistency definition, the imputation $\xi(x_0, T-t_0)$ is time inconsistent if there does not exist IDP $\beta(t) = \beta(t, \xi(x_0, T-t_0))$ such that at any $t \in [t_0, T]$, the dynamic imputation $\xi(x^*(t), T-t)$ generated by this IDP belongs to the Core $C(x^*(t), T-t)$.

Due to Theorem 1.2.1, if the corresponding IDP $\beta(t) = \beta(t, \xi(x_0, T-t_0))$ satisfies conditions (1.21) and (1.22), then $\xi(x_0, T-t_0)$ belongs to $\text{IDP-core}(x_0, T-t_0)$. That is, $\xi(x_0, T-t_0)$ is time-consistent by Definition 1.2.3, i.e., the following inclusion holds for any $t \in [t_0, T]$:

$$\xi(x^*(t), T-t) = \int_t^T \beta(\tau) d\tau \in C(x^*(t), T-t).$$

The proposition is proved. \square

Remark 1 *Proposition 1.2.5 states that if the Core is not empty, then the IDP-*

core and the Core coincide in the current game or equivalently, imputations from the Core coincide with the imputations from the IDP-core. The system of inequalities (1.21)-(1.22) allows extracting from the set of imputation distribution procedures a subset that provides time-consistency and the IDP-nondominance of all imputations from the Core. For other subsets of the IDP set, this kind of result is generally not true. Also, note that the set of imputation distribution procedures (1.21)-(1.22) can be empty. In the next section, consider the approach to check nonemptiness and construct IDPs from the IDP-core.

Suppose that the characteristic function $V(S; x^*(t), T - t)$, $t \in [t_0, T]$, is defined in some relevant way (for example, like in [146]). Suppose that it is a strictly monotonically decreasing function for any $t \in [t_0, T]$, decreasing faster than the linear law. We construct the Core $C(x_0, T - t_0)$ for the initial instant $t = t_0$. Afterwards, we choose an imputation from the Core $\xi(x_0, T - t_0) \in C(x_0, T - t_0)$. According to Definition 1.1.2, we choose the corresponding imputation distribution procedure $\beta(t) = \beta(t, \xi(x_0, T - t_0))$ (IDP) for $\xi(x_0, T - t_0)$ in some relevant way. As it follows from Proposition 1.2.5, for the imputations $\xi(x_0, T - t_0) \in \text{IDP-core}(x^*(t_0), T - t_0) = C(x^*(t_0), T - t_0)$ we can always find an IDP $\beta'(t)$ satisfying conditions (1.21)-(1.22) such that:

$$\int_{t_0}^T \beta'(t) dt = \xi(x_0, T - t_0).$$

On the other hand, $\beta(t) = \beta(t, \xi(x_0, T - t_0))$ could not satisfy conditions (1.21)-(1.22). It could be the case if $\beta(t) = \beta(t, \xi(x_0, T - t_0))$ is chosen so that

$$\begin{aligned} \beta(t) &= c = \text{const}, \quad t \in [t_0, t'), \quad (t' \in (t_0, T)), \\ \beta(t) &= \frac{\xi(x_0, T - t_0) - c(t' - t_0)}{T - t'}, \quad t \in [t', T], \end{aligned}$$

and the derivative of $V(S; x^*(t), T - t)$ is positive:

$$U(S; x^*(t), T - t) > 0, \quad t \in [t_0, T].$$

This case will be demonstrated on the model example in section 5.

1.2.3 Properties of IDP-core

Here several properties related to time-consistency of the IDP-core are presented.

Proposition 1.2.6 *If $C(x^*(t), T - t)$ is not empty for any $t \in [t_0, T]$, then all imputations of $C(x^*(t_0), T - t_0)$ are time-consistent.*

Proof Consider an imputation $\xi(x_0, T - t_0)$ from the Core $C(x^*(t_0), T - t_0)$. According to Proposition 1.2.5, $IDP - core(x^*(t_0), T - t_0) = C(x^*(t_0), T - t_0)$ and therefore, $\xi(x_0, T - t_0) \in IDP - core(x^*(t_0), T - t_0)$. From the definition of the IDP-core it follows that $\xi(x_0, T - t_0)$ is time-consistent. The proposition is proved. \square

As shown below, the IDP-core as a cooperative solution is strong time-consistent and not only build using the Axiom of time-consistency.

In the paper [101] it was shown that the IDP-core possesses strong time consistency.

Definition 1.2.7 *The cooperative solution $IDP - core(x_0, T - t_0)$ is a strong time-consistent solution in the game $\Gamma(x_0, T - t_0)$ if:*

- 1 $IDP - core(x^*(t), T - t) \neq \emptyset, \forall t \in [t_0, T]$,
- 2 $\forall \xi(x_0, T - t_0) \in IDP - core(x_0, T - t_0)$ there exists IDP $\beta(\tau), \tau \in [t_0, T]$, such that

$$\xi(x_0, T - t_0) = \int_{t_0}^T \beta(\tau) d\tau, \text{ and the following conditions are satisfied:}$$

$$IDP - core(x_0, T - t_0) \supseteq \int_{t_0}^t \beta(\tau) d\tau \oplus IDP - core(x^*(t), T - t), \forall t \in [t_0, T].$$

Here, \oplus is defined in the following way: for $a \in R^n$ and $D \subseteq R^n$, $a \oplus D = \{a + d : d \in D\}$.

The strong time consistency property of the IDP-core shows that when the game evolves along the cooperative trajectory $x^*(t)$, for any imputation $\xi(x_0, T - t_0) \in$

$IDP - core(x_0, T - t_0)$ there exists an IDP such that a deviation from the chosen imputation at the instant t to another imputation $\hat{\xi}(x^*(t), T - t) \in IDP - core(x^*(t), T - t)$ results in the total payoff in the game $\Gamma(x_0, T - t_0)$ corresponding to some imputation $\bar{\xi}(x_0, T - t_0)$ from the initial $IDP - core(x_0, T - t_0)$. Therefore, if the cooperative solution is strong time-consistent, its reconsideration during the game leads to one of the solutions of the initial IDP-core, therefore being unreasonable.

$IDP - core(x_0, T - t_0)$ is a strong time-consistent cooperative solution.

Theorem 1.2.2 *Suppose that $C(x^*(t), T - t) \neq \emptyset$ and the IDP-core $\bar{C}(x^*(t), T - t) \neq \emptyset, \forall t \in [t_0, T]$. Then $IDP - core(x_0, T - t_0)$ is a strong time-consistent cooperative solution in the game $\Gamma(x_0, T - t_0)$.*

Proof Because the Core and Subcore are both non-empty, for any imputation $\xi(x_0, T - t_0) \in \bar{C}(x_0, T - t_0) \subseteq C(x_0, T - t_0)$ there exists an IDP $\beta_i(t)$ such that $\xi_i(x_0, T - t_0) = \int_{t_0}^T \beta_i(t) dt, i \in N$. The game evolves along the cooperative trajectory $x^*(t), t \in [t_0, T]$, on which at each time instant the players find themselves in the current subgames $\Gamma_v(x^*(t), T - t)$.

Suppose that at the beginning of the game $\Gamma_v(x_0, T - t_0)$, the players have chosen an imputation $\xi(x_0, T - t_0) \in \bar{C}(x_0, T - t_0)$. Assume that at a time instant $t_{br}, t_{br} \in [t_0, T]$, a deviation from $\xi(x^*(t), T - t)$ to another imputation $\tilde{\xi}(x^*(t_{br}), T - t_{br})$ from the IDP-core $\bar{C}(x^*(t_{br}), T - t_{br})$ takes place. Let us show that the resulting imputation $\hat{\xi}(x_0, T - t_0) = \int_{t_0}^{t_{br}} \beta(t) dt + \tilde{\xi}(x^*(t_{br}), T - t_{br})$ will also belong to the IDP-core $\bar{C}(x_0, T - t_0)$. Well we have to prove the existence of an IDP $\hat{\beta}(t)$ such that $\int_{t_0}^T \hat{\beta}(t) dt = \hat{\xi}(x_0, T - t_0)$ and $\hat{\beta}(t) \in B(t)$ satisfies condition (1.21)-(1.22). By the definition of $\bar{C}(x_0, T - t_0)$, IDPs $\beta(t) \in B(t), t \in [t_0, T]$, and $\tilde{\beta}(t) \in B(t), t \in [t_{br}, T]$, (1.21)-(1.22). correspond to the imputations $\xi(x_0, T - t_0)$ and $\tilde{\xi}(x^*(t_{br}), T - t_{br})$, respectively.

Then the imputation $\hat{\beta}(t)$ has the form

$$\hat{\beta}(t) = \begin{cases} \beta(t), & t \in [t_0, t_{br}], \\ \tilde{\beta}(t), & t \in (t_{br}, T]. \end{cases}$$

To show that $\hat{\beta}(t) \in B(t), \forall t \in [t_0, T]$, it is necessary to prove that $\hat{\beta}(t)$ is an

integrable function. The functions $\beta(t)$ and $\tilde{\beta}(t)$ are integrable by construction because $\beta(t)$ and $\tilde{\beta}(t) \in B(t)$ on the intervals $[t_0, T]$ and $[t_{br}, T]$, respectively. Due to the definition of the integrable functions $\beta(t), \tilde{\beta}(t)$, it follows that the limit of integrable sums on the intervals $[t_0, t_{br}]$ and $(t_{br}, T]$ will converge. Hence, the total sum of these two integrable sums will also converge. Since $\hat{\beta}(t) \in B(t), \forall t \in [t_0, T]$, according to the definition of $\overline{C}(x_0, T - t_0)$ we obtain $\hat{\xi}(x_0, T - t_0) \in \overline{C}(x_0, T - t_0)$. The theorem is proved. \square

1.3 Nonemptiness of IDP-core using Linear Programming Approach

This section is devoted to the linear programming approach, described in section 2.4, for the non-emptiness properties of Core. The IDP-core can be constructed using a system of linear constraints for imputation distribution procedures. These constraints are defined for each instant in the game. From the nonemptiness of the set described by these constraints, it follows that the IDP-core is not empty.

Consider the following linear programming problem for a fixed t :

$$\begin{aligned} \sum_{i \in N} \beta_i &\longrightarrow \min \\ \sum_{i \in S} \beta_i &\geq U(S; x^*(t), T - t), \quad \forall S \subseteq N, S \neq N, S \neq \emptyset. \end{aligned} \quad (1.31)$$

Suppose that $\beta_i^0 = (\beta_1^0, \dots, \beta_n^0)$ is an optimal solution of the linear programming problem (1.31) with fixed t . The set of optimal solutions of problem (1.31) we denote by Y^0 .

Then the following theorem is true:

Theorem 1.3.1 *The set of IDPs satisfying conditions (1.21), (1.22), $t \in [t_0, T)$, is not empty if and only if $\forall t \in [t_0, T)$ the following condition is satisfied:*

$$\sum_{i \in N} \beta_i^0 \leq U(N; x^*(t), T - t), \quad (1.32)$$

where $\beta^0 \in Y^0$ is any solution of the linear programming problem (1.31).

Proof We start the proof with the sufficiency part. Suppose that condition (1.32) is satisfied. Then according to (1.31), for any $t \in [t_0, T]$ there exists $\hat{\beta}^0$ such that for

$$\beta_i = \beta_i^0 + \frac{U(N; x^*(t), T - t)}{n} - \frac{\sum_{i \in N} \beta_i^0}{n}, \quad i \in N, \quad (1.33)$$

conditions (1.21) and (1.22) are satisfied for any fixed $t \in [t_0, T]$. If it is true, we can compose the integrable function $\hat{\beta}^0(t)$ as a time-varying function for which conditions (1.21) and (1.22) will be satisfied.

Proof of the necessity part. If the IDP-core is not empty, then there exists at least one integrable function $\beta(t)$ satisfying conditions (1.21) and (1.22). As a result, for the solution of (1.31) condition (1.32) should be satisfied. \square

1.4 Differential Game Model of Resource Extraction

Consider a game-theoretic model of non-renewable resource extraction with asymmetric players [57; 168]. The amount of resource depends on the rates of extraction chosen by the players. The game involves n asymmetric players, with utility functions depending on the current amount of resource and the rates of extraction. Another related model example can be found in the paper [102].

We denote by $x(t) \in R^1$ the amount of resource at an instant t and by $u_i(t, x)$ a resource extraction rate chosen by player i at the instant t . As a class of strategies we will consider the class of feedback strategies, where the strategies are the functions of the time t and state x . Assume that $\forall t, u_i(t, x) \geq 0$, and $x(t) = 0$ implies $u_i(t, x) = 0$. The amount of resource $x(t)$ as a function of t depends on $u_i(t, x)$ in the following way:

$$\begin{aligned} \dot{x} &= - \sum_{i=1}^n a_i u_i(t, x), \quad a_i > 0, \quad i = 1, \dots, n, \\ x(t_0) &= x_0. \end{aligned} \quad (1.34)$$

The payoff function representing the income of player i is

$$K_i(x_0, T - t_0) = \int_{t_0}^T \log(u_i(\tau, x)) d\tau, \quad i = 1, \dots, n. \quad (1.35)$$

1.4.1 Cooperative Strategies and Cooperative Trajectory

Consider the cooperative version of the non-renewable resource extraction game [168]. Here, the players unite in the grand coalition and maximize their total utility, acting as one player. The corresponding optimal control problem is formalized in the following way:

$$\sum_{i=1}^n K_i(x_0, T - t_0) = \sum_{i=1}^n \int_{t_0}^T \log(u_i(\tau, x)) d\tau \rightarrow \max_{u_i, i=\overline{1, n}} \quad (1.36)$$

subject to

$$\begin{aligned} \dot{x} &= - \sum_{i=1}^n a_i u_i(t, x), \\ x(t_0) &= x_0 > 0, \\ u(t, x) &\geq 0. \end{aligned} \quad (1.37)$$

To solve the optimization problem (1.36), (1.37), we use the dynamic programming principle proposed by Bellman. We define the Bellman function as the maximum value of the total payoff of players (1.35) in the subgame $\Gamma(x, T - t)$ starting at an instant t in a position x :

$$W(t, x) = \max_{u_i, i=\overline{1, n}} \left\{ \sum_{i=1}^n K_i(x, T - t) \right\} = \max_{u_i, i=\overline{1, n}} \left\{ \sum_{i=1}^n \int_t^T \log u(\tau, x) d\tau \right\} \quad (1.38)$$

subject to equation (1.37) with $x_0 = x$ and $t_0 = t$.

It is proved that if there exists a continuously differentiable function $W(t, x)$ satisfying the Hamilton-Jacobi-Bellman equation

$$\begin{aligned} -W_t(t, x) &= \max_{u_i, i=\overline{1, n}} \left\{ \sum_{i=1}^n \log u_i(t, x) - W_x(t, x) \left(\sum_{i=1}^n a_i u_i(t, x) \right) \right\}, \\ \lim_{t \rightarrow T-0} W(t, x) &= 0, \end{aligned} \quad (1.39)$$

then the strategies $u_i^*(t, x)$ defined by maximizing the right-hand side of (1.39) deliver maximum to the functional in the optimization problem (1.36), (1.37).

From the first-order optimality condition of (1.39) we obtain

$$u_i^* = \frac{1}{a_i W_x(t, x)}.$$

Substituting it into (1.39) yields

$$\begin{aligned} W_t(t, x) &= n \log W_x(t, x) + \log A^{[N]} + n, & A^{[N]} &= \prod_{i=1}^n a_i \\ \lim_{t \rightarrow T-0} W(t, x) &= 0. \end{aligned} \quad (1.40)$$

The Bellman function will be considered in the form

$$W(t, x) = A(t) \log x + B(t).$$

Substituting it into (1.39), we obtain:

$$\begin{aligned} \dot{A} \log x + \dot{B} &= n \log A - n \log x + \log A^{[N]} + n, \\ \lim_{t \rightarrow T-0} A(t) &= \lim_{t \rightarrow T-0} B(t) = 0. \end{aligned} \quad (1.41)$$

The solution of (1.41) are the functions

$$\begin{aligned} A(t) &= n(T - t), \\ B(t) &= -(T - t) \left(\log A^{[N]} + n \log n(T - t) \right). \end{aligned} \quad (1.42)$$

Substituting $A(t)$ and $B(t)$ into the Bellman function, we obtain

$$W(t, x) = n(T - t) \log \frac{x}{n(T - t)} - (T - t) \log A^{[N]}, \quad t \in [t_0, T]. \quad (1.43)$$

The corresponding optimal control or cooperative strategy has the form

$$u_i^*(t, x) = \frac{1}{a_i W_x(t, x)} = \frac{x}{a_i n(T - t)}, \quad t \in [t_0, T]. \quad (1.44)$$

Substituting the optimal control into the motion equation (1.37), we obtain the

differential equation for the trajectory corresponding to the optimal control:

$$\begin{aligned}\dot{x} &= -\frac{x}{T-t}, \\ x(t_0) &= x_0.\end{aligned}\tag{1.45}$$

The solution has the form

$$x^*(t) = x_0 \frac{T-t}{T-t_0}, \quad t \in [t_0, T].\tag{1.46}$$

The trajectory $x^*(t)$ and strategy (control) $u^*(t, x)$ will be called cooperative.

To determine the value of the players' maximum total payoff corresponding to the optimization problem (1.36), (1.37) in the subgame along the cooperative trajectory $x^*(t)$ (1.46), it is necessary to substitute the cooperative trajectory into the Bellman function (1.43):

$$W(t, x^*(t)) = n(T-t) \log \frac{x_0}{n(T-t_0)} - (T-t) \log A^{[N]}, \quad t \in [t_0, T].\tag{1.47}$$

1.4.2 Characteristic Function

To construct the rule for allocating the maximum joint payoff among the players, it is necessary to define the characteristic function for each coalition $S \subseteq N$:

$$V(S; x, T-t) = \begin{cases} \sum_{i=1}^n K_i(x, T-t), & S = N \\ W_S(t, x), & S \subset N, \\ 0, & S = \emptyset, \end{cases}\tag{1.48}$$

where $W_S(t, x)$ is defined as the maximum joint payoff of coalition S provided that the players from coalition $N \setminus S$ use strategies from a fixed admissible Nash equilibrium $u^{NE} = (u_1^{NE}, \dots, u_n^{NE})$ (section 1.1.1) in the original game.

It can be shown that in the non-cooperative game case, Nash equilibrium strategies are

$$u_i^{NE}(t, x) = \frac{x}{a_i(T-t)}, \quad i \in N.\tag{1.49}$$

Consider the case of coalition $S \subset N$. We introduce the Bellman function $W_S(t, x)$ as the maximum total payoff of the players from coalition S in the subgame $\Gamma(x, T - t)$ starting at an instant t in a position x :

$$W_S(t, x) = \max_{u_i, i \in S} \sum_{i \in S} \left\{ \int_t^T \log u_i d\tau \right\} \quad (1.50)$$

$$\text{subject to } \dot{x}(\tau) = - \sum_{i \in N} a_i u_i \quad (1.51)$$

$$u_i = u_i^{NE}, \quad i \in N \setminus S. \quad (1.52)$$

The Hamilton-Jacobi-Bellman equation for this problem has the form

$$-\frac{\partial W_S(t, x)}{\partial t} = \max_{u_i, i \in S} \left\{ \sum_{i \in S} \log u_i(t, x) - \frac{\partial W_S(t, x)}{\partial x} \left(\sum_{j=1}^n a_j u_j(t, x) \right) \right\},$$

$$\lim_{t \rightarrow T-0} W_S(t, x) = 0. \quad (1.53)$$

From the first-order optimality condition for (1.53) we obtain

$$u_i^* = \frac{1}{a_i \frac{\partial W_S(t, x)}{\partial x}}. \quad (1.54)$$

Substituting it into (1.53) yields

$$\frac{\partial W_S}{\partial t} = k \log \frac{\partial W_S}{\partial x} + \log A^{[S]} + k + \frac{\partial W_S}{\partial x} \sum_{j \in N \setminus S} \frac{x}{T-t}, \quad A^{[S]} = \prod_{i=1}^n a_i$$

$$\lim_{t \rightarrow T-0} W_S(t, x) = 0, \quad (1.55)$$

where $k = |S|$, $n = |N|$. Consider the following Bellman function:

$$W_S(t, x) = A(t) \log x + B(t).$$

Substituting it into (1.55), we obtain:

$$\dot{A} \log x + \dot{B} = k \log A - k \log x + \log A^{[S]} + k + (n - k) \frac{A}{T-t}, \quad (1.56)$$

$$\lim_{t \rightarrow T-0} A(t) = \lim_{t \rightarrow T-0} B(t) = 0.$$

The solution of (1.56) are the functions

$$\begin{aligned} A(t) &= k(T - t), \\ B(t) &= -k(T - t) \left(\frac{\log A^{[S]}}{k} + \log k(T - t) + n - k \right). \end{aligned} \quad (1.57)$$

The solution of the optimization problem (1.50) is

$$W_S(t, x) = k(T - t) \left[\log \frac{x}{T - t} - \log k - \frac{\log A^{[S]}}{k} - n + k \right]. \quad (1.58)$$

According to the definition, we obtain the characteristic function for coalition $S \neq N$:

$$V(S, x, T - t) = W_S(t, x).$$

To determine the way to allocate the maximum joint payoff (1.47) among the players along the cooperative trajectory $x^*(t)$ (1.46), namely, for the subgame starting at the instant t on the cooperative trajectory $x^*(t)$ (1.46), it is necessary to define the characteristic function along the cooperative trajectory. Let us substitute the expression for $x^*(t)$ (1.46) into the expression for the characteristic function $V(S, T - t, x)$, $S \subset N$, (4.188):

$$W_S(t, x^*(t)) = k(T - t) \left[\log \frac{x_0}{T - t_0} - \log k - \frac{\log A^{[S]}}{k} - n + k \right]. \quad (1.59)$$

In the case $S = N$, the characteristic function is calculated using (1.47).

1.4.3 IDP-core

Suppose that all players unite in the grand coalition N . Then they can guarantee their joint payoff equal to $V(N; x^*(t), T - t)$. To determine how to allocate the maximum joint payoff among the players, we use the notion of imputations $\xi(x, T - t)$. In particular, we will use the IDP-core as a cooperative solution in the game. According to Theorem 1.2.1, the IDP-core can be constructed using the conditions

for IDPs $\beta_i(t)$, $i \in N$:

$$\begin{aligned} \sum_{i \in S} \beta_i(t) &\geq -kt \left[\log \frac{x_0}{T-t_0} - \log k - \frac{\log A^{[S]}}{k} - n + k \right], \quad \forall S \subset N, \\ \sum_{i \in N} \beta_i(t) &= -nt \log \frac{x_0}{n(T-t_0)} - (T-t) \log A^{[N]}, \quad \forall t \in [t_0, T]. \end{aligned} \quad (1.60)$$

1.4.4 Non-emptiness of IDP-core

To study non-emptiness conditions, we solve the linear programming problem, as presented in the paper [102], for $t \in [t_0, T]$ with a fixed step Δt :

$$\begin{aligned} \sum_{i \in N} \beta_i &\longrightarrow \min \\ \sum_{i \in S} \beta_i &\geq -kt \left[\log \frac{x_0}{T-t_0} - \log k - \frac{\log A^{[S]}}{k} - n + k \right], \quad \forall S \subset N. \end{aligned} \quad (1.61)$$

As a result, the vector function $\beta^0 = (\beta_1^0, \dots, \beta_n^0)$ is obtained using the numerical methods, and corresponding conditions are to be verified in order for the IDP-core to be non-empty:

$$\sum_{i \in N} \beta_i^0 \leq U(N; x^*(t), T-t). \quad (1.62)$$

We construct the IDP $\hat{\beta}^0(t)$ using $\beta^0(t)$ and show that it satisfies conditions (1.21),(1.22):

$$\hat{\beta}_i^0(t) = \beta_i^0(t) + \frac{U(N; x^*(t), T-t) - \sum_{i \in N} \beta_i^0(t)}{n}. \quad (1.63)$$

1.4.5 Core and IDP-core

According to Theorem 1.2.1, the imputation corresponding to the IDP $\hat{\beta}^0(t)$,

$$\xi(x_0, T-t_0) = \int_{t_0}^T \hat{\beta}^0(t) dt, \quad (1.64)$$

belongs to the Core $C(x_0, T - t_0)$ because for the given parameters, the IDP $\hat{\beta}^0(t)$, $t \in [t_0, T]$, satisfies conditions (1.21)-(1.22) or $\xi(x_0, T - t_0)$ belongs to $IDP - core(x_0, T - t_0)$. But if we use the Core $C(x_0, T - t_0)$ instead of $IDP - core(x_0, T - t_0)$ as a cooperative solution in the game, we can use any IDP for the imputation (1.64) such as

$$\begin{aligned} \beta(t) &= c = const, \quad t \in [t_0, t'], \quad (t' \in (t_0, T)), \\ \beta(t) &= \frac{\xi(x_0, T - t_0) - c(t' - t_0)}{T - t'}, \quad t \in [t', T], \end{aligned} \quad (1.65)$$

but this will not necessarily satisfy conditions (1.21)-(1.22) at some instant. Therefore, it appears to be dominated, and the imputation $\xi(x_0, T - t_0)$ corresponding to this IDP is time-inconsistent.

Fig. 1 shows the set defined by the system of constraints (1.21),(1.22). The solid line is the solution $\beta^0(t)$ of the corresponding linear programming problem (1.31) as a function of time, and the dashed line is the IDP $\hat{\beta}^0(t)$ and IDP $\beta(t)$ (1.65) corresponding to the imputation (1.64).

The function $\hat{\beta}_i^0(t)$ satisfies the constraints (1.21),(1.22). It can be seen that the IDP-core in this game model is not empty, and conditions (1.32) of Theorem 1.3.1 are satisfied.

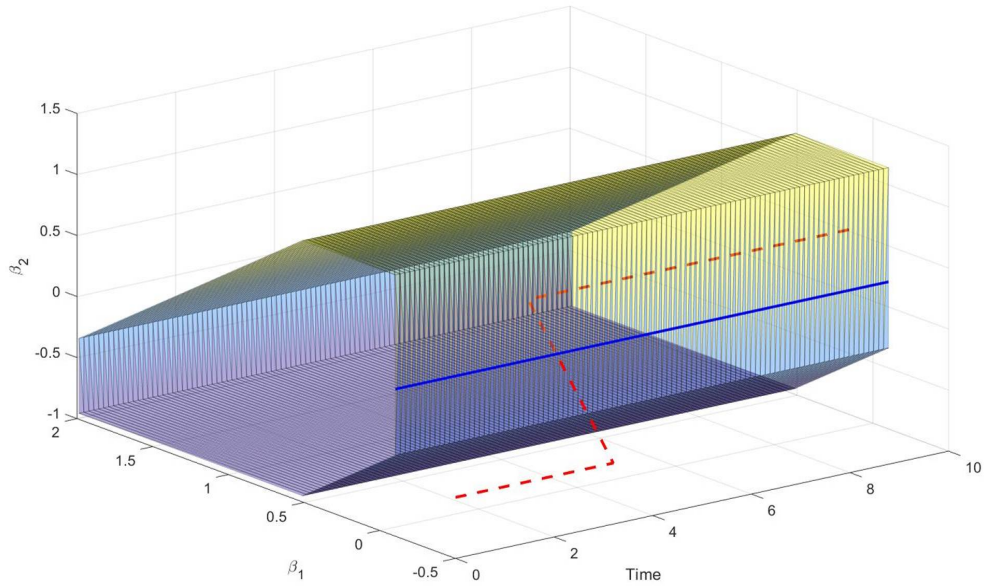


Figure 1.1: Axes: β_1, β_3, t . β_2 can be found using equality in (1.22).

Using Fig. 2 we can verify the non-emptiness conditions (1.32) of Theorem 1.3.1. The solid line in this figure shows the sum of the values $\beta_i^0(t)$, $i = \overline{1, 3}$:

$$S_{\beta^0}(t) = \beta_1^0(t) + \beta_2^0(t) + \beta_3^0(t). \quad (1.66)$$

The dashed line in Fig. 2 shows the characteristic function value for the grand coalition:

$$U(N; x^*(t), T - t) = U(\{1, 2, 3\}; x^*(t), T - t),$$

where $U(\{1, 2, 3\}; x^*(t), T - t)$ is defined in (1.17). In Fig. 2 it can be seen that

$$S_{\beta^0}(t) \leq U(\{1, 2, 3\}; x^*(t), T - t) \quad \forall t \in [t_0, T].$$

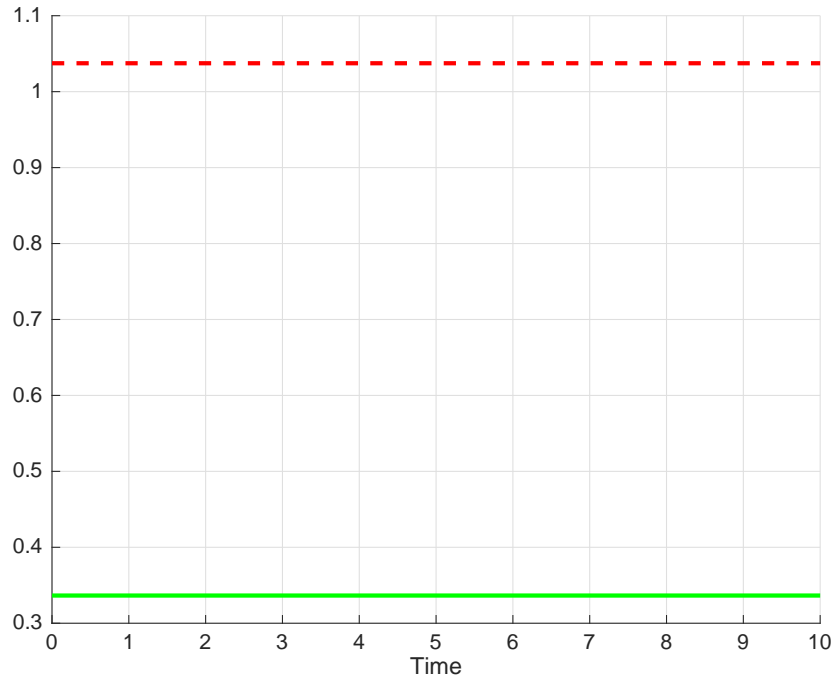


Figure 1.2: $U(\{1, 2, 3\}; x^*(t), T - t)$ (1.17) is dashed line, $S_{\beta^0}(t)$ (1.66) is solid line.

Chapter 2

Class of Discrete Games with Dynamic Updating

Information plays one of the main roles in game theory. Generally, classical differential game models consider only complete information: the game structure's information is known and will not change during the whole game. However, this section focuses on the case when information is updated dynamically, and players use only the truncated information, updating it at each stage. To model the behavior of players when information is updated dynamically, a novel approach of dynamic updating is introduced.

2.1 Discrete Game Model with Dynamic Updating

2.1.1 Original Discrete Game Model

Consider a general N -stage n -person nonzero-sum discrete-time dynamic game $\Gamma(x^0, N)$ [169] with an initial state $x^0 \in R^m$. The state dynamics of the game are defined by the following difference equation:

$$\begin{aligned} x_{k+1} &= f_k(x_k, u_1^k, u_2^k, \dots, u_n^k), \\ x_1 &= x^0, \end{aligned} \tag{2.1}$$

where $u_i^k \in U^i \subset R^{m_i}$ is the control vector of player i at stage k , $x_k \in X$ is the state of the game, and $k \in \{1, 2, \dots, N\} \equiv T$.

The payoff function of player i is

$$K_i(x^0, N, U) = \sum_{j=1}^N g_i^j[x_j, u_1^j, u_2^j, \dots, u_n^j, x_{j+1}] \left(\frac{1}{1+r} \right)^{j-1}, \text{ for } i \in \{1, 2, \dots, n\} \equiv I, \quad (2.2)$$

where r is a discount rate.

This game is a complete information game. However, in real-life-like games, players do not have complete information about the game structure on the interval on which the game is defined. We apply the dynamic updating approach to model the behavior of players when information about the game is updated dynamically.

2.1.2 Truncated Subgame

Assume that at each stage k , players have complete information about the motion equations and payoff functions within $k + \bar{T}$ stages, where \bar{T} is a fixed value, namely the information horizon. At stage k , information about the game is updated. At stage $k + 1$, players have complete information about the game structure on the interval until stage $k + 1 + \bar{T}$.

Definition 2.1.1 *Let $k = 1, \dots, N - \bar{T}$. A truncated subgame $\Gamma_k(x_{k,0}, k, k + \bar{T})$ is defined on an interval $[k, k + \bar{T}]$. The motion equation and payoff function of the truncated subgame on the stage interval $[k, k + \bar{T}]$ coincide with those of the original game on the same stage interval. The motion equation and initial condition of the truncated subgame have the following form:*

$$\begin{aligned} x_{k+1} &= f_k(x_k, u_1^k, u_2^k, \dots, u_n^k), \\ x_{1,0} &= x_1 = x^0. \end{aligned} \quad (2.3)$$

(We denote $x_{k,l} = x_{k+l}$, $u_i^{k,l} = u_i^{k+l}$, $l = \overline{0, \bar{T}}$ for the truncated subgame $\Gamma_k(x_{k,0}, k, k + \bar{T})$.)

The payoff function of player i in the truncated subgame $\Gamma_k(x_{k,0}, k, k + \bar{T})$ has

the form

$$K_i(x_{k,0}, k, k + \bar{T}, U) = \sum_{j=0}^{\bar{T}} g_i^{k+j} [x_{k,j}, u_1^{k,j}, u_2^{k,j}, \dots, u_n^{k,j}, x_{k,j+1}] \left(\frac{1}{1+r} \right)^{k+j-1}, \quad (2.4)$$

for $i \in \{1, 2, \dots, n\} \equiv I$.

2.2 Noncooperative Discrete Game with Dynamic Updating

Recently, a theory of noncooperative games has been developed to describe strategic interaction between several decision-makers with payoff functions depending on the choice of all players. By assumption, the players do not cooperate, and it is important to model their behavior using some principle of optimality. Many results have been obtained in noncooperative game theory; a general description can be found in the fundamental book [170]. In the papers [17; 171], the famous solution concept called Nash equilibrium was proposed.

Considering each truncated subgame, let $u_k^{NE,l}(x) = (u_1^{NE,k,l}(x), \dots, u_n^{NE,k,l}(x))$, $l \in [0, \bar{T}]$, be a vector of strategies that provides a feedback Nash equilibrium solution for the truncated subgame $\Gamma_k(x_{k,0}, k, k + \bar{T})$. We denote by $\bar{V}_i^k(l, x)$ the payoff of player $i \in I$ in feedback Nash equilibrium in the subgame defined on the interval $[l, k + \bar{T}]$, $l = k, \dots, k + \bar{T}$:

$$\bar{V}_i^k(l, x) = \sum_{j=l}^{k+\bar{T}} g_i^{k+j} [x_{k,j}, u_1^{NE,k,j}, u_2^{NE,k,j}, \dots, u_n^{NE,k,j}, x_{k,j+1}] \left(\frac{1}{1+r} \right)^{k+j-1}.$$

Theorem 2.2.1 ([172]) *Suppose that there exist functions $\bar{V}_i^k(l, x)$ satisfying (2.5). Then the strategies that maximize the right-hand side of (2.5) constitute the feedback*

Nash equilibrium:

$$\begin{aligned}
\bar{V}_i^k(l, x) &= \max_{u_i^{k,l}} \left\{ g_i^k[x, u_1^{NE,k,l}(x), \dots, u_{i-1}^{NE,k,l}, u_i^{k,l}, u_{i+1}^{NE,k,l}, \dots, u_n^{NE,k,l}(x), \right. \\
&\quad \left. \bar{f}_k(x, u_i^{k,l}) \right] \left(\frac{1}{1+r} \right)^{k+l-1} + \bar{V}_i^k(l+1, \bar{f}_k(x, u_i^{k,l})) \left. \right\} = \\
&= g_i^k[x, u_1^{NE,k,l}(x), \dots, u_n^{NE,k,l}(x), f_k(x_k, u_1^{NE,k,l}(x), \dots, u_n^{NE,k,l}(x))] \times \\
&\quad \times \left(\frac{1}{1+r} \right)^{k+l-1} + \bar{V}_i^k(l+1, f_k(x_k, u_1^{NE,k,l}(x), \dots, u_n^{NE,k,l}(x))),
\end{aligned} \tag{2.5}$$

where $\bar{V}_i^k(\bar{T} + 1, x) = 0$, $i \in I$.

We denote by $u_k^{NE,l}(x)$, $l \in [0, \bar{T}]$, the strategies obtained by maximization the right-hand side of (2.5). We denote by $x_k^{NE} = (x_{k,0}^{NE}, \dots, x_{k,\bar{T}}^{NE})$ the trajectory obtained with the strategies $u_k^{NE,l}(x)$. The presented solution is valid only for a truncated subgame, not for the whole game defined on the interval $[0, N]$. It is also impossible to construct the Nash equilibrium using the classical approaches since information about the game is updated dynamically, and it is different on different intervals. To construct the resulting equilibrium strategies and corresponding trajectory for the class of games with dynamic updating, we introduce the notion of the resulting strategies and corresponding resulting trajectory:

Definition 2.2.1 *The resulting feedback Nash equilibrium strategies $\{\hat{u}_i^{NE,j}\}_{j=1}^N$, $\forall i \in I$ are constructed in the following way:*

$$\hat{u}_i^{NE,j} = \begin{cases} u_i^{NE,j,0}, & j \in \{1, \dots, N - \bar{T}\}, \\ u_i^{NE,N-\bar{T},j-N-\bar{T}}, & j \in \{N - \bar{T} + 1, \dots, N\}. \end{cases}$$

Definition 2.2.2 *The resulting noncooperative trajectory $\{\hat{x}_j^{NE}\}_{j=1}^N$ is a combination of x_k^{NE} for each truncated subgame $\Gamma_k(x_{k,0}^{NE}, k, k + \bar{T})$:*

$$\{\hat{x}_j^{NE}\}_{j=0}^N = \begin{cases} x_{j,0}^{NE}, & j = \overline{1, N - \bar{T}}, \\ x_{N-\bar{T},j-(N-\bar{T})}^{NE}, & j = \overline{N - \bar{T} + 1, N}. \end{cases} \tag{2.6}$$

Along the resulting noncooperative trajectory, the players obtain a payoff according to the following formula:

$$\begin{aligned} \hat{V}(\hat{x}_k^{NE}, N - k) &= \\ &= \begin{cases} \sum_{j=k}^{N-1-\bar{T}} [\bar{V}_i^j(0, x_{j,0}^{NE}) - \bar{V}_i^j(1, x_{j,1}^{NE})] + \bar{V}_i^{N-\bar{T}}(0, x_{N-\bar{T},0}^{NE}), & k = \overline{1, N - \bar{T}}, \\ \bar{V}_i^{N-\bar{T}}(k - N + \bar{T}, x_{N-\bar{T}, k-N+\bar{T}}^{NE}), & k = \overline{N - \bar{T} + 1, N}. \end{cases} \end{aligned} \quad (2.7)$$

2.3 Cooperative Discrete Game with Dynamic Updating

The theory of cooperative dynamic games offers socially convenient and group efficient solutions for different decision problems involving strategic actions. One of the fundamental problems in the theory of cooperative dynamic games is the formulation of optimal or cooperative behavior for players and the definition of an allocation rule or cooperative solution, which is the set of imputations chosen by the players in the game. Suppose that the players decide to cooperate in each truncated subgame. We denote by $\Gamma_k^c(x_{k,0}, k, k + \bar{T})$ a truncated cooperative subgame defined on an interval $[k, k + \bar{T}]$ with an initial condition $x_{k,0}$. The maximum joint payoff of all players in this game can be defined by maximizing the functional

$$\sum_{i \in I} K_i(x_{k,0}, k, k + \bar{T}, U) = \sum_{i \in I} \sum_{j=0}^{\bar{T}} g_i^{k+j} [x_{k,j}, u_1^{k,j}, u_2^{k,j}, \dots, u_n^{k,j}, x_{k,j+1}] \left(\frac{1}{1+r} \right)^{k+j-1}, \quad (2.8)$$

subject to

$$x_{k+1} = f_k(x_k, u_1^k, u_2^k, \dots, u_n^k), \quad (2.9)$$

$$x_{1,0} = x_1 = x^0. \quad (2.10)$$

Assume that the Bellman function for each truncated subgame $\Gamma_k^c(x_{k,0}, k, k + \bar{T})$

has the following form [173]:

$$W^k(0, x_{k,0}) = \max_{u_i(k), i \in I} \left\{ \sum_{i \in I} \sum_{j=0}^{\bar{T}} g_i^{k+j} [x_{k,j}, u_1^{k,j}, \dots, u_n^{k,j}, f_k(x_{k,j}, u_1^{k,j}, \dots, u_n^{k,j})] \left(\frac{1}{1+r} \right)^{k+j-1} \right\}, \quad (2.11)$$

where $u_i(k) = \{u_i^{k,0}, \dots, u_i^{k,\bar{T}}\}$.

Theorem 2.3.1 *Suppose that there exists a function $W^k(l, x)$ in the truncated subgame $\Gamma_k^c(x_{k,0}, k, k + \bar{T})$, $l \in [0, \bar{T}]$, such that the following recursive relations are satisfied:*

$$W^k(l, x) = \max_{u_i(k), i \in I} \left\{ \sum_{i \in I} g_i^k [x, u_1^{k,l}, \dots, u_n^{k,l}, f_k(x, u_1^{k,l}, \dots, u_n^{k,l})] \left(\frac{1}{1+r} \right)^{k+l-1} + W^k(l+1, f_k(x, u_1^k(x), \dots, u_n^k(x))) \right\}, \quad (2.12)$$

where $W^k(\bar{T} + 1, x) = 0$. Assume that the maximum in (2.12) is achieved under a control $u_i^*(k)$. Then $u_i^*(k)$ is optimal in the control problem (2.8), (2.9).

Definition 2.3.1 *The resulting cooperative strategies $\{\hat{u}_i^{*,j}\}_{j=1}^N$, $\forall i \in I$, are constructed in the following way:*

$$\hat{u}_i^{*,j} = \begin{cases} u_i^{*,j,0}, & j \in \{1, \dots, N - \bar{T}\}, \\ u_i^{*,N-\bar{T},j-N-\bar{T}}, & j \in \{N - \bar{T} + 1, \dots, N\}. \end{cases}$$

In a similar way, we define the resulting cooperative trajectory. We denote by u^* the optimal strategies obtained through maximization of the right-hand side of (2.12). The corresponding trajectory is defined by $x_k^* = (x_{k,0}^*, \dots, x_{k,\bar{T}}^*)$.

Definition 2.3.2 *The resulting cooperative trajectory $\{\tilde{x}_l^*\}_{l=1}^N$ is a combination of the cooperative trajectories x_k^* for each truncated cooperative subgame $\Gamma_k^c(x_k^*, k, k +$*

\bar{T}) defined on the successive intervals $[k, k + \bar{T}]$:

$$\{\tilde{x}_j^*\}_{j=1}^N = \begin{cases} x_{j,0}^*, & j = \overline{1, N - \bar{T}}, \\ x_{N-\bar{T}, j-(N-\bar{T})}^*, & j = \overline{N - \bar{T} + 1, N}. \end{cases} \quad (2.13)$$

On the interval $[k, k + 1]$, the resulting cooperative trajectory coincides with the cooperative trajectory x_k^* in the truncated cooperative subgame $\Gamma_k^c(x_k^*, k, k + \bar{T})$. At stage $k + 1$, information about the game structure is updated. On the interval $[k + 1, k + 2]$, the trajectory \tilde{x}_k^* coincides with the cooperative trajectory x_{k+1}^* in the truncated cooperative subgame $\Gamma_{k+1}^c(x_{k+1}^*, k + 1, k + 1 + \bar{T})$, etc.

The characteristic function of a coalition is an essential concept in the theory of cooperative games. In this section, it is defined like in [174] as the total payoff of players from coalition $S \subseteq I$ in Nash equilibrium in the game with the following set of players: coalition S (acting as one player) and the players from the set $I \setminus S$ (acting as individuals). For each coalition $S \subseteq I$, we define the values of the characteristic function for each truncated subgame as was done in [112]:

$$V^k(S; x_{k,0}^*) = \begin{cases} \sum_{i \in I} K_i^k(x_{k,0}^*; u), & S = I, \\ \tilde{V}_k(S, x_{k,0}^*), & S \subseteq I, \\ 0, & S = \emptyset, \end{cases} \quad (2.14)$$

where $\tilde{V}_k(S, x_{k,0}^*)$ is defined as the total payoff of all players from coalition $S \setminus I$ in the feedback Nash equilibrium $u^{NE} = (\bar{u}_1^{NE}, \dots, \bar{u}_n^{NE})$ in the game with the following set of players: coalition S (acting as one player) and the players from the set $I \setminus S$, i.e., in the game with $|I \setminus S| + 1$ players. For $|S| = 1$, the cooperative setting turns into the noncooperative one. In other cases, the players from coalition S act as one player maximizing their total payoff.

2.3.1 Resulting Cooperative Solution and Related Theorems

An imputation $\xi^k(x_k^*, k, k + \bar{T})$ for each truncated cooperative subgame $\Gamma_{k+1}^c(x_k^*, k, k + \bar{T})$ is defined as an arbitrary vector satisfying the following conditions:

individual rationality: $\xi_i^k(x_k^*, k, k + \bar{T}) \geq V(\{i\}; x_k^*, k, k + \bar{T})$,

group rationality: $\sum_{i \in I} \xi_i^k(x_k^*, k, k + \bar{T}) = V(I; x_k^*, k, k + \bar{T})$.

We denote by $E_k(x^*, k, k + \bar{T})$ the set of all possible imputations for each truncated subgame. Suppose that for each truncated subgame, a non-empty cooperative solution is defined by

$$M_k(x^*, k, k + \bar{T}) \subseteq E_k(x^*, k, k + \bar{T}). \quad (2.15)$$

It can be the core, nucleus, or Shapley value.

To guarantee the time consistency theorem [169] in each cooperative truncated subgame, it is necessary to use the notion of the imputation distribution procedure. The theorem of time consistency was introduced in 1977 by L. Petrosyan [175].

Following the continuous-time analysis of Yeung and Petrosyan [176] for cooperative differential games, we formulate a discrete-time version of the imputation distribution procedure [177] so that the agreed upon imputations will be time consistent. We denote by $B_i^k(x_i^{*,k})$ the payment that player i obtains at stage k under the cooperative agreement along the cooperative trajectory $\left\{x_i^{*,k}\right\}_k^N$. The payment scheme involving $B_i^k(x_i^{*,k})$ constitutes an IDP in the sense that the imputation of player i over the stages from k to N can be expressed as

$$\xi_i^k(x_i^{*,k}) = B_i^k(x_i^{*,k}) + \sum_{j=k+1}^N B_i^j(x_i^{*,j}), \quad i \in I. \quad (2.16)$$

Definition 2.3.3 *A solution $M(x_k^*, k)$ is called time consistent if for any imputation*

$\xi^k(x^{,k}) \in M(x_k^*, k)$, there exists an IDP $B^k(x_i^{*,k}) = (B_1^k(x_i^{*,k}), \dots, B_n^k(x_i^{*,k}))$ such*

that, for $k = \overline{1, N}$,

$$\sum_k^N B^k(x_i^{*,k}) \in M(x_k^*, k) \quad \left(\sum_k^N B^k(x_i^{*,k}) = \xi^k(x_i^{*,k}) \right).$$

The agreed-upon optimality principle or cooperative solution is time consistent if condition (2.3.3) is maintained at any time instant throughout the game along the cooperative trajectory.

To construct the resulting cooperative solution or cooperative solution in the game with dynamic updating $\Gamma(x^0, N)$, it is necessary to use a special approach because the standard approach can only be applied to each truncated subgame with dynamic updating defined on the interval $[k, k + \bar{T}]$. The IDP also provides the time consistency theorem of the new solution concept and the ability to determine solutions within the interval $[k, k + \bar{T}]$.

Suppose that in each truncated subgame $\Gamma_k^c(x_k^*, k, k + \bar{T})$, a cooperative solution $M_k(x^*, k, k + \bar{T}) \neq \emptyset$ along the cooperative trajectory x_k^* is selected. Suppose that an imputation

$$\xi^k(x_k^*, k, k + \bar{T}) \in M_k(x^*, k, k + \bar{T})$$

and the corresponding IDP for $k \in \overline{1, N - \bar{T}}$ are selected in a way to guarantee the time consistency theorem:

$$\beta^{k,l}(x_{k,l}^*) = [\beta_1^{k,l}(x_{k,l}^*), \dots, \beta_n^{k,l}(x_{k,l}^*)], \text{ for } l \in \overline{0, \bar{T}}.$$

The corresponding IDP $\beta^{k,l}(x_{k,l}^*)$ can be obtained in the following form [169]:

$$\beta^{k,l}(x_{k,l}^*) = \xi^k(x_{k,l}^*, k + l, k + \bar{T}) - \xi^k(x_{k,l+1}^*, k + l + 1, k + \bar{T}), \text{ } l \in \overline{0, \bar{T}}, \quad (2.17)$$

where $\xi^k(x_{k,\bar{T}+1}^*, k + \bar{T} + 1, k + \bar{T}) = 0$.

However, time consistency only holds on the interval $[k, k + \bar{T}]$ for each truncated subgame $\Gamma_k^c(x_k^*, k, k + \bar{T})$, $k \in \{1, \dots, N\}$. To guarantee time consistency for the whole interval $[1, N]$, we introduce the resulting IDP.

Definition 2.3.4 *The resulting IDP $\hat{\beta}^k(\tilde{x}^*)$ is defined for the set of chosen imputations in each truncated subgame $\xi^k(x_{k,0}^*, k, k + \bar{T}) \in M_k(x^*, k, k + \bar{T})$ using*

the corresponding $\beta^k(x_k^*)$ as follows:

$$\left\{ \hat{\beta}^k(x^*) \right\}_{k=1}^N = \begin{cases} \beta^{k,0}(x_{k,0}^*), & k = \overline{1, N - \bar{T}}, \\ \beta^{N-\bar{T}, k-N+\bar{T}}(x_{N-\bar{T}, k-N+\bar{T}}^*), & k = \overline{N - \bar{T} + 1, N}. \end{cases} \quad (2.18)$$

Using the resulting IDP $\hat{\beta}^k(\tilde{x}^*)$, we define the following vector.

Definition 2.3.5 *The resulting imputation $\hat{\xi}(\tilde{x}^*, N - k)$ is a vector defined using the resulting IDP $\hat{\beta}^k(x^*)$ in the following way: for stage $k = \overline{1, N}$,*

$$\hat{\xi}(\tilde{x}^*, N - k) = \sum_{l=k}^N \hat{\beta}^l(\tilde{x}^*), \quad (2.19)$$

in particular, $\hat{\xi}(\tilde{x}^*, N) = \sum_{l=1}^N \hat{\beta}^l(\tilde{x}^*)$.

Using the notion of resulting imputation, it is possible to define the allocation rule for the joint payoff.

Definition 2.3.6 *The resulting solution $\hat{M}_k(\tilde{x}^*, N - k)$ is the set of resulting imputations $\hat{\xi}(\tilde{x}^*, N - k)$ for all possible resulting IDPs, $\hat{\beta}^k(\tilde{x}^*)$ (defined by different imputations in each truncated subgame).*

Any resulting imputation $\hat{\xi}(\tilde{x}^*, N) \in \hat{M}(\tilde{x}^*, N)$ and the corresponding resulting IDP $\hat{\beta}^k(\tilde{x}^*)$ allocate the total payoff of players along the resulting cooperative trajectory \tilde{x}_i^* in the original game with prescribed duration, i.e., the following condition holds:

$$\sum_{i \in I} \sum_{j=1}^k \hat{\beta}_i^j(\tilde{x}_j^*) = \sum_{i \in I} \sum_{j=1}^k g_i^j[\tilde{x}_j^*, \tilde{u}^*, \tilde{x}_{i+1}^*] \left(\frac{1}{1+r} \right)^{j-1}. \quad (2.20)$$

The resulting solution $\hat{M}_k(\tilde{x}^*, N - k)$ is time consistent by construction. However, it can also be proven that an arbitrary resulting cooperative solution is strong time consistent in the game model with dynamic updating. By an arbitrary solution, we understand a solution composed of, e.g., some vectors from the core at the first stage, the Shapley value at the second stage, etc.

Definition 2.3.7 A cooperative solution $M(x^0, N)$ is called strong time consistent if $\forall k = \overline{1, N}$ and every $\xi(x^0, N) \in M(x^0, N)$, the corresponding IDP $\beta^k(x_k^*)$ satisfies the condition

$$\sum_{j=1}^k \beta^j(x_j^*) \oplus M(x_{k,0}^*, N - k) \subseteq M(x^0, N), \quad (2.21)$$

where $a \oplus A = \{a + a' : a' \in A\}$.

Theorem 2.3.2 An arbitrary resulting cooperative solution $\hat{M}(\tilde{x}^*, N)$ is strong time consistent in the game model with dynamic updating.

2.3.2 Properties of Resulting Cooperative Solution

2.3.2.1 Construction of Characteristic Function in Game Model with Dynamic Updating

In the previous section, we have defined the notion of the resulting cooperative solution, i.e., a way for allocating the cooperative payoff along the chosen trajectory in the game with dynamic updating. However, we have not shown that the constructed set is a set of imputations or a set of vectors satisfying the individual and group rationality Theoremerties. To justify the resulting cooperative solution, we need to define the characteristic function for the game model with dynamic updating.

Definition 2.3.8 The resulting characteristic function $\bar{V}(S; \tilde{x}_k^*, N - k)$, $S \subseteq I$, in the game $\Gamma(\tilde{x}_k^*, N - k)$ with dynamic updating is a function calculated using the characteristic function values $V(S; x_k^*, k, k + \bar{T})$ in every truncated subgame $\Gamma_k^c(x_{k,0}, k, k + \bar{T})$ along the resulting cooperative trajectory \tilde{x}_k^* for $k \in \{1, \dots, N\}$:

$$\bar{V}(S; \tilde{x}_k^*, N - k) = \begin{cases} \sum_{l=k}^{N-\bar{T}-1} [V(S; x_l^*, l, l + \bar{T}) - V(S; x_{l+1}^*, l + 1, l + \bar{T})] + \\ \quad + V(S; x_{N-\bar{T}}^*, N - \bar{T}, N), \quad k = \overline{1, N - \bar{T}}, \\ V(S; x_k^*, k, N), \quad k = \overline{N - \bar{T} + 1, N}. \end{cases} \quad (2.22)$$

Using the resulting characteristic function, we can show that the resulting imputation $\hat{\xi}(\tilde{x}^*, N)$ is an imputation in the game with dynamic updating.

Theorem 2.3.3 *The resulting imputation $\hat{\xi}(\tilde{x}^*, N)$ is an imputation in the game $\Gamma(x^0, N)$ with dynamic updating if for $k \in \{1, \dots, N - \bar{T}\}$, the following condition is satisfied:*

$$\begin{aligned} & \xi_i^k(x_{k,l}^*, k+l, k+\bar{T}) - \xi_i^k(x_{k,l+1}^*, k+l+1, k+\bar{T}) \geq \\ & \geq V(\{i\}; x_{k,l}^*, k+l, k+\bar{T}) - V(\{i\}; x_{k,l+1}^*, k+l+1, k+\bar{T}), \quad i \in I, \end{aligned} \quad (2.23)$$

where $l = \overline{0, \bar{T}}$.

Proof It is necessary to prove that the following conditions are satisfied:

$$\sum_{i=1}^n \hat{\xi}_i(\tilde{x}_k^*, N-k) = \bar{V}(I; \tilde{x}_k^*, N-k), \quad (2.24)$$

$$\hat{\xi}_i(\tilde{x}_k^*, N-k) \geq \bar{V}(\{i\}; \tilde{x}_k^*, N-k), \quad i \in I. \quad (2.25)$$

We introduce the function

$$y(k) = \begin{cases} N - \bar{T} + 1, & k = \overline{1, N - \bar{T}}, \\ k, & k = \overline{N - \bar{T} + 1, N}. \end{cases}$$

According to Definition 2.3.5, the right-hand side of (2.24) can be rewritten as

$$\begin{aligned} & \sum_{i=1}^n \hat{\xi}_i(\tilde{x}_k^*, N-k) = \sum_{i=1}^n \sum_{j=k}^N \hat{\beta}_i^j(\tilde{x}_k^*) = \\ & = \sum_{i=1}^n \left[\sum_{j=k}^{N-\bar{T}} (\bar{\xi}_i^j(x_{j,0}^*, j, j+\bar{T}) - \bar{\xi}_i^k(x_{j,1}^*, j+1, j+\bar{T})) + \right. \\ & \left. + \sum_{j=y(k)}^N (\bar{\xi}_i^{N-\bar{T}}(x_{N-\bar{T}, j-N+\bar{T}}^*, j, N) - \bar{\xi}_i^{N-\bar{T}}(x_{N-\bar{T}, j+1-N+\bar{T}}^*, j+1, N)) \right] = \quad (2.26) \\ & = \sum_{j=k}^{N-\bar{T}} \left[V(I; x_{j,0}^*, j, j+\bar{T}) - V(I; x_{j,1}^*, j+1, j+\bar{T}) \right] + \\ & + \sum_{j=y(k)}^N \left[V(I; x_{N-\bar{T}, j-N+\bar{T}}^*, j, N) - V(I; x_{N-\bar{T}, j+1-N+\bar{T}}^*, j+1, N) \right]. \end{aligned}$$

The right-hand side of (2.26) is equal to $\bar{V}(I; \tilde{x}_k^*, N-k)$ according to Definition

2.3.8; therefore, (2.24) is correct. Now, consider (2.25), where the left-hand side can be rewritten as

$$\begin{aligned} \hat{\xi}_i(\tilde{x}_k^*, N - k) &= \sum_{j=k}^{N-\bar{T}} (\bar{\xi}_i^j(x_{j,0}^*, j, j + \bar{T}) - \bar{\xi}_i^k(x_{j,1}^*, j + 1, j + \bar{T})) + \\ &+ \sum_{j=y(k)}^N (\bar{\xi}_i^{N-\bar{T}}(x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - \bar{\xi}_i^{N-\bar{T}}(x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N)). \end{aligned} \quad (2.27)$$

(2.25) is fulfilled for each stage if for $k = \overline{1, N}$, the following condition is satisfied:

$$\begin{aligned} &\bar{\xi}_i^j(x_{j,0}^*, j, j + \bar{T}) - \bar{\xi}_i^k(x_{j,1}^*, j + 1, j + \bar{T}) \geq \\ &V(\{i\}; x_{j,0}^*, j, j + \bar{T}) - V(\{i\}; x_{j,1}^*, j + 1, j + \bar{T}), j = \overline{k, N - \bar{T}}, \\ &\bar{\xi}_i^{N-\bar{T}}(x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - \bar{\xi}_i^{N-\bar{T}}(x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N) \geq \\ &V(\{i\}; x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - V(\{i\}; x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N), j = \overline{N - \bar{T} + 1, N}. \end{aligned} \quad (2.28)$$

Fulfillment of condition (2.28) implies the fulfillment of the condition for each truncated subgame $k = \{1, \dots, N - \bar{T}\}$:

$$\begin{aligned} &\xi_i^k(x_{k,l}^*, k + l, k + \bar{T}) - \xi_i^k(x_{k,l+1}^*, k + l + 1, k + \bar{T}) \geq \\ &V(\{i\}; x_{k,l}^*, k + l, k + \bar{T}) - V(\{i\}; x_{k,l+1}^*, k + l + 1, k + \bar{T}), l = \overline{0, \bar{T}}, i \in I. \end{aligned} \quad (2.29)$$

The theorem is proven. □

2.3.2.2 Relationship of Solutions in Truncated Subgames and Resulting Solutions

In this section, it is shown that if the players choose imputations $\xi^k(x_k^*, k, k + \bar{T}) \in E_k(x^*, k, k + \bar{T})$ based on $V(S; x^*, k, k + \bar{T})$ for $k \in \{1, \dots, N - \bar{T}\}$ and $S \subseteq I$ in every truncated subgame by the same rule, then the resulting imputation $\hat{\xi}(\tilde{x}^*, N - k)$ corresponds to the imputation chosen by the same rule using the resulting characteristic function $\bar{V}(S; x^*, N - k)$, $S \subseteq I$. Further, we will prove it for several classical optimality principles.

Assume that if in every truncated subgame $\Gamma_k^c(x_k^*, k, k + \bar{T})$ the players choose the Shapley value $Sh^k(x_k^*, k, k + \bar{T})$ as the imputation, then the corresponding resulting

imputation $\hat{\xi}(\tilde{x}^*, N-k)$ will coincide with the Shapley value $\hat{Sh}(\tilde{x}_k^*, N-k)$ calculated using the resulting characteristic function $\bar{V}(S; \tilde{x}^*, N-k)$, $S \subseteq I$.

Theorem 2.3.4 *Suppose that in every truncated subgame $\Gamma_k^c(x_k^*, k, k + \bar{T})$:*

$$\xi^k(x_{k,0}^*, k, k + \bar{T}) = Sh^k(x_k^*, k, k + \bar{T}),$$

where $k \in \{1, \dots, N - \bar{T}\}$. Then the corresponding resulting imputation $\hat{\xi}(\tilde{x}_k^*, N-k)$ coincides with $\hat{Sh}(\tilde{x}_k^*, N-k)$:

$$\hat{\xi}(\tilde{x}_k^*, N-k) = \hat{Sh}(\tilde{x}_k^*, N-k), \quad k \in [1, N],$$

where $\hat{Sh}(\tilde{x}_k^*, N-k)$ is the Shapley value calculated using the resulting characteristic function $\bar{V}(S; \tilde{x}_k^*, N-k)$, $S \subseteq I$.

Proof Due to Definition 2.3.5 and (2.17), the resulting imputation is calculated

using the Shapley values $Sh^k(x_k^*, k, k + \bar{T})$ in each truncated subgame:

$$\begin{aligned}
\hat{\xi}(x_k^*, N - k) &= \sum_{j=k}^{N-\bar{T}} [Sh^j(x_{j,0}^*, j, j + \bar{T}) - Sh^j(x_{j,1}^*, j + 1, j + \bar{T})] + \\
&+ \sum_{j=y(k)}^N [Sh^{N-\bar{T}}(x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - Sh^{N-\bar{T}}(x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N)] \\
&= \sum_{S \subset I, i \in S} \frac{(|I| - |S|)! (|S| - 1)!}{|I|!} \left\{ \sum_{j=k}^{N-\bar{T}} [(V(S; x_{j,0}^*, j, j + \bar{T}) - \right. \\
&- V(S \setminus i; x_{j,0}^*, j, j + \bar{T})) - (V(S; x_{j,1}^*, j + 1, j + \bar{T}) - V(S \setminus i; x_{j,1}^*, j + 1, j + \bar{T}))] + \\
&+ \sum_{j=y(k)}^N [(V(S; x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - V(S \setminus i; x_{N-\bar{T},j-N+\bar{T}}^*, j, N)) - \\
&- (V(S; x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N) - V(S \setminus i; x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N))] \left. \right\} = \\
&= \sum_{S \subset I, i \in S} \frac{(|I| - |S|)! (|S| - 1)!}{|I|!} \left\{ \sum_{j=k}^{N-\bar{T}} [(V(S; x_{j,0}^*, j, j + \bar{T}) - \right. \\
&- V(S; x_{j,1}^*, j + 1, j + \bar{T})) - (V(S \setminus i; x_{j,0}^*, j, j + \bar{T}) - V(S \setminus i; x_{j,1}^*, j + 1, j + \bar{T}))] + \\
&+ \sum_{j=y(k)}^N [(V(S; x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - V(S; x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N)) - \\
&- (V(S \setminus i; x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - V(S \setminus i; x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N))] \left. \right\}, \tag{2.30}
\end{aligned}$$

$\hat{S}h(x_k^*, N - k)$ is the Shapley value calculated using the resulting characteristic

function $\bar{V}(S, \tilde{x}_k^*, N - k)$:

$$\begin{aligned}
& \hat{S}h(x_k^*, N - k) = \\
& = \sum_{S \subset I, i \in S} \frac{(|I| - |S|)! (|S| - 1)!}{|I|!} (\bar{V}(S; \tilde{x}_k^*, N - k) - \bar{V}(S \setminus i; \tilde{x}_k^*, N - k)) = \\
& = \sum_{S \subset I, i \in S} \frac{(|I| - |S|)! (|S| - 1)!}{|I|!} \left\{ \sum_{j=k}^{N-\bar{T}} [(V(S; x_{j,0}^*, j, j + \bar{T}) - \right. \\
& - V(S; x_{j,1}^*, j + 1, j + \bar{T})) - (V(S \setminus i; x_{j,0}^*, j, j + \bar{T}) - V(S \setminus i; x_{j,1}^*, j + 1, j + \bar{T}))] + \\
& + \sum_{j=y(k)}^N [(V(S; x_{N-\bar{T}, j-N+\bar{T}}^*, j, N) - V(S; x_{N-\bar{T}, j+1-N+\bar{T}}^*, j + 1, N)) - \\
& \left. - (V(S \setminus i; x_{N-\bar{T}, j-N+\bar{T}}^*, j, N) - V(S \setminus i; x_{N-\bar{T}, j+1-N+\bar{T}}^*, j + 1, N))] \right\}. \tag{2.31}
\end{aligned}$$

Comparing (2.30) with (2.31), we arrive in the desired result. The theorem is proven. \square

Suppose that in every truncated subgame $\Gamma_k^c(x_k^*, k, k + \bar{T})$, the players choose the core $C(x_k^*, k, k + \bar{T})$ as the cooperative solution. Then each resulting imputation $\hat{\xi}(x^*, N - k)$ belongs to the core defined using the resulting characteristic function $\bar{V}(S, \tilde{x}_k^*, N - k)$.

Theorem 2.3.5 *Suppose that in every truncated subgame $\Gamma_k^c(x_k^*, k, k + \bar{T})$:*

$$M_k(x^*, k, k + \bar{T}) = C(x_k^*, k, k + \bar{T}),$$

where $k \in \{1, \dots, N - \bar{T}\}$. Then for every $\xi^k(x_k^*, k, k + \bar{T}) \in C(x_k^*, k, k + \bar{T})$, the following relation is satisfied:

$$\hat{\xi}(\tilde{x}_k^*, N - k) \in \hat{C}(\tilde{x}_k^*, N - k), \quad k \in [k, N],$$

where $\hat{C}(\tilde{x}_k^*, N - k)$ is the core calculated using the resulting characteristic function

$\bar{V}(S, \tilde{x}_k^*, N - k)$, $S \subseteq I$, if

$$\begin{aligned} & \sum_{i \in S} \bar{\xi}_i^k(x_{k,0}^*, k, k + \bar{T}) - V(S; x^*, k, k + \bar{T}) \geq \\ & \geq \sum_{i \in S} \bar{\xi}_i^k(x_{k,1}^*, k + 1, k + \bar{T}) - V(S; x^*, k + 1, k + \bar{T}) \geq 0, \quad S \subseteq I. \end{aligned} \quad (2.32)$$

Proof The statement of the theorem is satisfied when the two statements below are true:

- If in every truncated subgame the players choose the imputation $\xi^k(x_k^*, k, k + \bar{T}) \in C(x_k^*, k, k + \bar{T})$ calculated using $V(S; x^*, k, k + \bar{T})$ for $k \in \{1, \dots, N - \bar{T}\}$, then the resulting imputation $\hat{\xi}(\tilde{x}_k^*, N - k)$ will belong to the core $\hat{C}(\tilde{x}_k^*, N - k)$ calculated using the resulting characteristic function $\bar{V}(S, \tilde{x}_k^*, N - k)$.
- The core $\hat{C}(\tilde{x}_k^*, N - k)$ should not contain an imputation $\hat{\xi}(\tilde{x}_k^*, N - k)$ for which it is impossible to find the set of imputations in the truncated subgame $\xi^k(x_k^*, k, k + \bar{T}) \in C(x_k^*, k, k + \bar{T})$.

Consider the first statement. If the set of imputations $\xi^k(x_k^*, k, k + \bar{T})$ satisfies the system of inequalities

$$\sum_{i \in S} \xi_i^k(x_{k,0}^*, k, k + \bar{T}) \geq V(S; x^*, k, k + \bar{T}), \quad S \subseteq I, \quad (2.33)$$

for $k \in \{1, \dots, N - \bar{T}\}$, $i \in I$, then the resulting imputation $\hat{\xi}(\tilde{x}_k^*, N - k)$ satisfies the system of inequalities

$$\sum_{i \in S} \hat{\xi}_i(\tilde{x}_k^*, N - k) \geq \bar{V}(S, \tilde{x}_k^*, N - k). \quad (2.34)$$

We substitute the expression for $\xi_i^k(x_{k,0}^*, k, k + \bar{T})$ into the left-hand side of (2.34) and (2.22) into the right-hand side of (2.34). Therefore, for every $S \subseteq I$, $k \in \{1, \dots, N - \bar{T}\}$, the fulfillment of (2.32) implies the fulfillment of (2.33); hence, (2.34) is satisfied. The first statement is proven.

Consider the second statement. For any $\hat{\xi}(\tilde{x}_k^*, N - k) \in \hat{C}(\tilde{x}_k^*, N - k)$, $k \in \{1, \dots, N - \bar{T}\}$, the following inequalities are satisfied in the set of imputations

$\xi_i^k(x_{k,0}^*, k, k + \bar{T})$, $i \in \{1, \dots, n\}$, $k \in \{1, \dots, N - \bar{T}\}$:

$$\begin{aligned}
& \sum_{i \in S} \left[\sum_{j=k}^{N-\bar{T}} (\xi_i^j(x_{j,0}^*, j, j + \bar{T}) - \xi_i^k(x_{j,1}^*, j + 1, j + \bar{T})) + \right. \\
& \left. + \sum_{j=y(k)}^N (\xi_i^{N-\bar{T}}(x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - \xi_i^{N-\bar{T}}(x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N)) \right] \geq \\
& \geq \sum_{j=k}^{N-\bar{T}} (V(S; x_{j,0}^*, j, j + \bar{T}) - V(S; x_{j,1}^*, j + 1, j + \bar{T})) + \\
& + \sum_{j=y(k)}^N (V(S; x_{N-\bar{T},j-N+\bar{T}}^*, j, N) - V(S; x_{N-\bar{T},j+1-N+\bar{T}}^*, j + 1, N)).
\end{aligned} \tag{2.35}$$

If there exists a set of imputations $\xi_i^k(x_{k,0}^*, k, k + \bar{T})$, where $i \in \{1, \dots, n\}$ and $k \in \{1, \dots, N - \bar{T}\}$ in (2.35), for which inequalities (2.33) are not satisfied, these unsatisfied inequalities will lead to the following condition for $k \in \{1, \dots, N - \bar{T}\}$ being satisfied:

$$\sum_{i \in S} \xi_i^k(x_{k,0}^*, k, k + \bar{T}) - V(S; x^*, k, k + \bar{T}) < 0. \tag{2.36}$$

This result contradicts (2.32), so the theorem is proven. □

2.4 Random Information Horizon Case

In this section, we consider a case when the length of the information horizon is not fixed; this means that the players may have certain information on truncated intervals $[k, \bar{T}_k]$, but the length of the intervals is a random variable. We denote by \bar{T}_k the length of the information horizon at stage k . \bar{T}_k can take values from the interval $[\hat{T}_{k-1} + 1, N]$, where \hat{T}_{k-1} is a realization of the random horizon \bar{T}_{k-1} in the previous truncated subgame, in particular, $\hat{T}_0 = 0$. It is worth mentioning that once $\bar{T}_k = N$, the updating procedure stops. \bar{T}_k is a discrete random variable. At the initial position x_0 , the following probabilities are defined for realizations on the interval $[1, N]$: $\{\gamma_1, \dots, \gamma_N\}$. In the position x_1 (the second position), the

probabilities are recalculated in the following way:

$$\frac{\gamma_{\bar{T}_{k-1}+1}}{\sum_{\zeta=\bar{T}_{k-1}+1}^N \gamma_{\zeta}}, \dots, \frac{\gamma_N}{\sum_{\zeta=\bar{T}_{k-1}+1}^N \gamma_{\zeta}}.$$

The motion equation on the interval $[k, \bar{T}_k]$ is

$$\begin{aligned} x_{k+1} &= f_k(x_k, u_1^k, u_2^k, \dots, u_n^k), \\ x_{1,0} &= x_1 = x^0. \end{aligned} \quad (2.37)$$

The payoff function of player $i \in \{1, \dots, n\}$ for the k^{th} truncated subgame is

$$E \left\{ \sum_{\zeta=k}^{\bar{T}_k} g_i^{\zeta}(x_{\zeta}, u_{\zeta}, x_{\zeta+1}) \right\} = \sum_{\bar{T}_k=\bar{T}_{k-1}+1}^N \frac{\gamma_{\bar{T}_k}}{\sum_{j=\bar{T}_{k-1}+1}^N \gamma_j} \sum_{\zeta=k}^{\bar{T}_k} g_i^{\zeta}(x_{\zeta}, u_{\zeta}, x_{\zeta+1}). \quad (2.38)$$

Consider the noncooperative game model. To define the feedback Nash equilibrium in this game, we introduce the Bellman function $V_i^k(x, k)$ for the k^{th} truncated subgame as follows:

$$V_i^k(x, k) = \max_{u_i^k, \dots, u_i^{\bar{T}_k}} \left\{ \sum_{\bar{T}_k=\bar{T}_{k-1}+1}^N \frac{\gamma_{\bar{T}_k}}{\sum_{j=\bar{T}_{k-1}+1}^N \gamma_j} \sum_{\zeta=k}^{\bar{T}_k} g_i^{\zeta}(x_{\zeta}, u_{\zeta}, x_{\zeta+1}) \mid x_k = x \right\}$$

with the set of strategies $\{u_i^j = \psi_i^j, j = \overline{k, \bar{T}_k}, i \in I\}$. Then

$$\begin{aligned} V_i^k(x, k) &= \max_{u_i^k, \dots, u_i^{\bar{T}_k}} \left\{ \sum_{\bar{T}_k=\bar{T}_{k-1}+1}^N \frac{\gamma_{\bar{T}_k}}{\sum_{j=\bar{T}_{k-1}+1}^N \gamma_j} \sum_{\zeta=k}^{\bar{T}_k} g_i^{\zeta}(x_{\zeta}, u_{\zeta}, x_{\zeta+1}) \right\} = \\ &= \sum_{\bar{T}_k=\bar{T}_{k-1}+1}^N \frac{\gamma_{\bar{T}_k}}{\sum_{j=\bar{T}_{k-1}+1}^N \gamma_j} \sum_{\zeta=k}^{\bar{T}_k} g_i^{\zeta}(x_{\zeta}^*, \psi_{\zeta}, x_{\zeta+1}^*) = \\ &= \sum_{l=k}^{\bar{T}_{k-1}} g_i^l(x_l^*, \psi_l, x_{l+1}^*) + \sum_{\bar{T}_k=\bar{T}_{k-1}+1}^N \frac{\gamma_{\bar{T}_k}}{\sum_{j=\bar{T}_{k-1}+1}^N \gamma_j} \sum_{\zeta=\bar{T}_{k-1}+1}^{\bar{T}_k} g_i^{\zeta}(x_{\zeta}^*, \psi_{\zeta}, x_{\zeta+1}^*), \end{aligned}$$

where $x_{k+1}^* = f_k(x_k^*, \psi_k(x_k^*))$.

Like in the theorem presented in [178], a feedback Nash equilibrium in the noncooperative case of each truncated subgame with random horizon can be

characterized as follows.

Theorem 2.4.1 *A set of strategies $\{\psi_i^j, j = \overline{k, N}, i \in I\}$ provides a feedback Nash equilibrium solution of the game (2.37), (2.38) if there exist functions $V_i^k(x, j)$, for $j \in \{k, \dots, N\}$ and $i \in I$, such that the following recursive relations are satisfied:*

$$\begin{aligned} \hat{V}_i^k(x, N) &= \max_{u_i^N} \left\{ g_i^N(x, \psi_1^N, \dots, \psi_{i-1}^N, u_i^N, \psi_{i+1}^N, \dots, \psi_n^N, \tilde{f}_N(x, u_i^N)) \right\}, \\ \hat{V}_i^k(x, \tau) &= \max_{u_i^\tau} \left\{ g_i^\tau(x, \psi_1^\tau, \dots, \psi_{i-1}^\tau, u_i^\tau, \psi_{i+1}^\tau, \dots, \psi_n^\tau, \tilde{f}_\tau(x, u_i^\tau)) + \right. \\ &\quad \left. + \frac{\sum_{j=\tau+1}^N \gamma_j}{\sum_{j=\tau}^N \gamma_j} \hat{V}_i^k(\tilde{f}_\tau(x, u_i^\tau), \tau + 1) \right\}, \text{ for } \tau \in \{\bar{T}_{k-1} + 1, \dots, N - 1\}, \\ V_i^k(x, \nu) &= \max_{u_i^\nu} \left\{ g_i^\nu(x, \psi_1^\nu, \dots, \psi_{i-1}^\nu, u_i^\nu, \psi_{i+1}^\nu, \dots, \psi_n^\nu, \tilde{f}_\nu(x, u_i^\nu)) + \right. \\ &\quad \left. + V_i^k(\tilde{f}_\nu(x, u_i^\nu), \nu + 1) \right\}, \text{ for } \nu \in \overline{k, \bar{T}_{k-1}}, \\ V_i^k(x, \bar{T}_{k-1} + 1) &= \hat{V}_i^k(x, \bar{T}_{k-1} + 1), \end{aligned} \tag{2.39}$$

where $\tilde{f}_k(x, u_i^k) = f_k(x, \psi_1^k, \dots, \psi_{i-1}^k, u_i^k, \psi_{i+1}^k, \dots, \psi_n^k)$.

Consider the case when all players agree to cooperate in each truncated subgame. To maximize their expected joint payoff in the k^{th} truncated subgame, we need to solve the following discrete-time dynamic programming problem:

$$E \left\{ \sum_{i \in I} \sum_{\zeta=k}^{\bar{T}_k} g_i^\zeta(x_\zeta, u_\zeta, x_{\zeta+1}) \right\} = \sum_{i \in I} \left[\sum_{\bar{T}_k = \bar{T}_{k-1} + 1}^N \frac{\gamma_{\bar{T}_k}}{\sum_{j=\bar{T}_{k-1} + 1}^N \gamma_j} \sum_{\zeta=k}^{\bar{T}_k} g_i^\zeta(x_\zeta, u_\zeta, x_{\zeta+1}) \right] \tag{2.40}$$

subject to (2.37).

Like in the noncooperative case, we can characterize an optimal solution of (2.37) and (2.40) as follows.

Corollary 2.4.1 *A set of strategies $\{\phi_i^j, j = \overline{k, N}, i \in I\}$ provides optimal cooperative strategies in the game (2.37) and (2.40) if there exists a function $W_i^k(x, j)$, for $j \in \{k, \dots, N\}$ and $i \in I$, such that the following recursive relations*

are satisfied:

$$\begin{aligned}
\hat{W}_i^k(x, N) &= \max_{u_1^N, \dots, u_n^N} \left\{ \sum_{i \in I} g_i^N(x, u_1^N, \dots, u_n^N, f_N(x, u_1^N, \dots, u_n^N)) \right\}, \\
\hat{W}_i^k(x, \tau) &= \max_{u_1^\tau, \dots, u_n^\tau} \left\{ \sum_{i \in I} [g_i^\tau(x, u_1^\tau, \dots, u_n^\tau, f_\tau(x, u_1^\tau, \dots, u_n^\tau))] + \right. \\
&\quad \left. + \frac{\sum_{j=\tau+1}^N \gamma_j}{\sum_{j=\tau}^N \gamma_j} \hat{W}_i^k(f_\tau(x, u_1^\tau, \dots, u_n^\tau), \tau + 1) \right\}, \text{ for } \tau \in \{\bar{T}_{k-1} + 1, \dots, N - 1\}, \\
W_i^k(x, \nu) &= \max_{u_1^\nu, \dots, u_n^\nu} \left\{ \sum_{i \in I} [g_i^\nu(x, u_1^\nu, \dots, u_n^\nu, f_\nu(x, u_1^\nu, \dots, u_n^\nu))] + \right. \\
&\quad \left. + W_i^k(f_\nu(x, u_1^\nu, \dots, u_n^\nu), \nu + 1) \right\}, \text{ for } \nu \in \overline{k, \bar{T}_{k-1}},
\end{aligned} \tag{2.41}$$

where $W_i^k(x, \bar{T}_{k-1} + 1) = \hat{W}_i^k(x, \bar{T}_{k-1} + 1)$.

Using the set of optimal strategies $\{\phi_i^j, j = \overline{k, N}, i \in I\}$, we can generate x_k^* from (2.37), obtaining the corresponding cooperative trajectory.

2.5 Dynamic Oligopoly Marketing Model of Advertising

2.5.1 Original Game Model

Consider an n -firm dynamic oligopoly marketing model of advertising defined on a closed time interval $[t_0, T]$. The solution for the noncooperative case of this model in differential form was presented in [179]. In this section, the model is transformed to the dynamic setting and studied for not only the noncooperative case, but mainly for the cooperative case [114]. Advertising efforts as strategies are used by firms to compete on the oligopoly market, and each firm tries to increase its market share while competitors try to minimize it using their advertising efforts. We denote by x_i^k the market share of firm $i \in I \equiv \{1, \dots, n\}$ at stage k . The full list of notations is presented in the table below. The dynamic oligopoly marketing model of advertising is defined as follows. The motion equation of firm $i \in I$ (its market share) has the

form

$$\begin{cases} x_i^{k+1} = h\left(\frac{n}{n-1}\rho_i u_i^k \sqrt{1-x_i^k} - \frac{1}{n-1} \sum_{j \in I} \rho_j u_j^k \sqrt{1-x_j^k}\right) - (h\delta - 1)x_i^k + h\delta \frac{1}{n}, \\ x_i^0 = z_i^0, \end{cases} \quad (2.42)$$

where z_i^0 is a positive constant.

The payoff function of firm $i \in I$ in the dynamic game model (the profit of firm $i \in I$) is given by

$$K_i^0(x^0, u^0) = \sum_{l=0}^N m_i x_i^l - (u_i^l)^2. \quad (2.43)$$

Unlike the motion equation (2.2) in the model example, here we use the discount rate $r = 0$. The main reason is that we do not want to complicate the result of the model and show the difference between the approach of dynamic updating and the original differential game first. The next thing is that we may model the advertising process on a short time interval, i.e., with a small value T .

In the table below, the list of parameters of the model is presented.

Table 2.1: List of variables and parameters

Notation	Explanation
$x_i^k \leq 1$	Market share of firm $i \in I \equiv \{1, \dots, n\}$ at stage k
$u_i^k \geq 0$	Advertising effort rate of firm $i \in I$ at stage k
$\rho_i > 0$	Advertising effectiveness parameter of firm $i \in I$
$\delta > 0$	Churn parameter
$m_i > 0$	Industry sales multiplied by the per unit profit margin of firm $i \in I$
$C(u_i^k)$	Cost of advertising of firm $i \in I$ at stage k , parameterized by $(u_i^k)^2$

The original marketing game model is defined on the interval $[0, N]$. Suppose that players have complete information about the motion equations and payoff functions on the interval $[k, k + \bar{T}]$. To construct a solution in the game model with dynamic updating, we use the notion of truncated subgame $\Gamma_k(x^{k,0}, k, k + \bar{T})$.

The payoff function of firm $i \in I$ in the truncated subgame $\Gamma_k(x^{k,0}, k, k + \bar{T})$ has the form

$$\bar{K}_i^k(x_i^{k,0}, u) = \sum_{j=k}^{k+\bar{T}} m_i x_i^j - (u_i^j)^2. \quad (2.44)$$

In (2.42), there is a logical consistency requirement that the market shares should satisfy:

$$\sum_{i \in I} x_i^\gamma = 1, \quad x_i^\gamma \geq 0, \quad i \in I, \quad \gamma = \overline{0, N}. \quad (2.45)$$

2.5.2 Noncooperative Outcomes in Truncated Subgame

First, we consider the noncooperative case for each truncated subgame $\Gamma_k(x_k^0, k, k + \bar{T})$, $k \in \{0, \dots, N - \bar{T}\}$ with the initial condition x_k^0 . According to Theorem 2.2.1, the set of strategies $\{\bar{u}_i^{k,l}, \text{ for } i \in \{1, \dots, n\}, l = \overline{0, \bar{T}} \text{ and } k \in \{0, \dots, N - \bar{T}\}\}$, provides a feedback Nash equilibrium solution of the game (2.42)–(2.44) if there exist functions $V_i^k(l, x_i^{k,l})$, for $i \in \{1, \dots, n\}$, $l = \overline{0, \bar{T}}$ and $k \in \{0, \dots, N - \bar{T}\}$, in each truncated subgame $\Gamma_k(x_k^0, k, k + \bar{T})$ such that the following recursive relations are satisfied for each $i \in I$:

$$\begin{aligned} V_i^k(l, x_i^{k,l}) &= \max_{u_i^{k,l}} \left\{ m_i x_i^{k,l} - (u_i^{k,l})^2 + V_i^k(l+1, x_i^{k,l+1}) \right\} = \\ &= m_i x_i^{k,l} - (\bar{u}_i^{k,l})^2 + V_i^k(l+1, x_i^{k,l+1}). \end{aligned}$$

Theorem 2.5.1 *The payoff of players in each truncated subgame $\Gamma_k(x_k^0, k, k + \bar{T})$ in feedback Nash equilibrium on the interval $[k, k + \bar{T}]$ has the following form:*

$$V_i^k(l, x_i^{k,l}) = \sum_{i \in I} \tilde{A}_i^{k,l} x_i^{k,l} + \tilde{B}^{k,l}(i), \quad i \in \{1, \dots, n\}, \quad k \in \{0, \dots, N - \bar{T}\}, \quad l = \overline{0, \bar{T}},$$

where $\tilde{A}_i^{k,l}$, $\tilde{B}^{k,l}(i)$ are determined from the relations

$$\begin{aligned} \tilde{A}_i^{k,l} &= m_i - (\hat{G}_i^{k,l+1} Z_i)^2 - \tilde{A}_i^{k,l+1}(\delta h - 1), \\ \tilde{A}_j^{k,l} &= -2(\hat{G}_j^{k,l+1} Z_j)^2 - \tilde{A}_j^{k,l+1}(\delta h - 1), \quad j \neq i, \quad j \in I \setminus i, \\ \tilde{B}^{k,l}(i) &= \sum_{i \in I} [2(\hat{G}_i^{k,l+1} Z_i)^2 + \frac{\tilde{A}_i^{k,l+1} h \delta}{n}] + \tilde{B}^{k,l+1}(i) - (\hat{G}_i^{k,l+1} Z_i)^2, \end{aligned}$$

with the initial conditions $\tilde{A}_i^{k, \bar{T}+1} = \tilde{B}^{k, \bar{T}+1}(i) = 0$, $i \in \{1, \dots, n\}$, where $\hat{G}_i^{k,l+1} = n \tilde{A}_i^{k,l+1} - \sum_{j \in I} \tilde{A}_j^{k,l+1}$, $Z_i = \frac{h \rho_i}{2(n-1)}$, $\hat{G}_i^{k,l+1} = \max \left\{ 0, \hat{G}_i^{k,l+1} \right\}$.

Proof The Bellman equation for firm $i \in I$ has the form

$$V_i^k(l, x_i^{k,l}) = \max_{u_i^k \geq 0} \left\{ m_i x_i^{k,l} - (u_i^{k,l})^2 + V_i^k(l+1, x_i^{k,l+1}) \right\}.$$

Using (2.42) and $V_i^k(l, x_i^{k,l}) = \sum_{i \in I} \tilde{A}_i^{k,l} x_i^{k,l} + \tilde{B}^{k,l}(i)$, for $i \in I$, $l = \overline{0, \bar{T}}$ and $k \in \{0, \dots, N\}$, we write the Bellman equation as

$$\begin{aligned} \sum_{i \in I} \tilde{A}_i^{k,l} x_i^{k,l} + \tilde{B}^{k,l}(i) &= \max_{u_i^k \geq 0} \left\{ m_i x_i^{k,l} - (u_i^{k,l})^2 + \sum_{i \in I} \tilde{A}_i^{k,l+1} x_i^{k,l+1} + \tilde{B}^{k,l+1}(i) \right\} = \\ &= \max_{u_i^k \geq 0} \left\{ \sum_{i \in I} \tilde{A}_i^{k,l+1} \left[h \left(\frac{n}{n-1} \rho_i u_i^{k,l} \sqrt{1-x_i^{k,l}} - \frac{1}{n-1} \sum_{j \in I} \rho_j u_j^{k,l} \sqrt{1-x_j^{k,l}} \right) - \right. \right. \\ &\quad \left. \left. - (h\delta - 1)x_i^{k,l} + h\delta \frac{1}{n} \right] + \tilde{B}^{k,l+1}(i) + m_i x_i^{k,l} - (u_i^{k,l})^2 \right\}. \end{aligned} \tag{2.46}$$

Performing the indicated maximization in this expression yields

$$-2u_i^{k,l} + \frac{h\rho_i}{n-1} n \tilde{A}_i^{k,l+1} \sqrt{1-x_i^{k,l}} - \frac{h\rho_i}{n-1} \sqrt{1-x_i^{k,l}} \sum_{j \in I} \tilde{A}_j^{k,l+1} = 0,$$

for $i \in I$, $l = \overline{0, \bar{T}}$ and $k \in \{0, \dots, N - \bar{T}\}$. Then the feedback Nash equilibrium strategy of firm $i \in I$ can be obtained in the form

$$\begin{aligned} \bar{u}_i^{k,l} &= \max \left\{ 0, \left[n \tilde{A}_i^{k,l+1} - \sum_{j \in I} \tilde{A}_j^{k,l+1} \right] \frac{h\rho_i}{2(n-1)} \sqrt{1-x_i^{k,l}} \right\} = \\ &= \frac{h\rho_i \sqrt{1-x_i^{k,l}}}{2(n-1)} \max \left\{ 0, n \tilde{A}_i^{k,l+1} - \sum_{j \in I} \tilde{A}_j^{k,l+1} \right\}. \end{aligned}$$

Anticipating that the controls will be shown to be positive, we substitute this into (2.46) and collect like terms together to express the parameters $\tilde{A}_i^{k,l}$ and $\tilde{B}^{k,l}(i)$ as

follows:

$$\begin{aligned}\tilde{A}_i^{k,l} &= m_i - (\hat{G}_i^{k,l+1} Z_i)^2 - \tilde{A}_i^{k,l+1}(\delta h - 1), \\ \tilde{A}_j^{k,l} &= -2(\hat{G}_j^{k,l+1} Z_j)^2 - \tilde{A}_j^{k,l+1}(\delta h - 1), \quad j \neq i, \quad j \in I \setminus i, \\ \tilde{B}^{k,l}(i) &= \sum_{i \in I} [2(\hat{G}_i^{k,l+1} Z_i)^2 + \frac{\tilde{A}_i^{k,l+1} h \delta}{n}] + \tilde{B}^{k,l+1}(i) - (\hat{G}_i^{k,l+1} Z_i)^2,\end{aligned}$$

where $\hat{G}_i^{k,l+1} = n\tilde{A}_i^{k,l+1} - \sum_{j \in I} \tilde{A}_j^{k,l+1}$, $\hat{G}_i^{k,l+1} = \max \{0, \hat{G}_i^{k,l+1}\}$, $Z_i = \frac{h\rho_i}{2(n-1)}$. \square

The corresponding feedback Nash equilibrium strategies are

$$\bar{u}_i^{k,l} = \hat{G}_i^{k,l+1} Z_i \sqrt{1 - x_i^{k,l}}, \quad k \in \{0, \dots, N - \bar{T}\}, \quad i \in \{1, \dots, n\}. \quad (2.47)$$

Consider now the case of a random horizon for a truncated subgame. Suppose that the players know some information about the game structure, but the duration of this information being right (or not changed) is unknown. Suppose that the probabilities are symmetric on the interval $[1, N]$, $\gamma_1 = \dots = \gamma_N = \frac{1}{N}$. In our model, the first truncated subgame starts at stage 0. Due to Theorem 2.4.1 for the k^{th} truncated subgame, if there exist continuously differentiable functions $V_i^k(x, f)$, $i \in \{1, \dots, n\}$, $f \in \{k, \dots, \bar{T}_{k-1}\}$, and $\hat{V}_i^k(x, \tau)$, for $\tau \in \{\bar{T}_{k-1} + 1, \dots, N\}$, such that the following recursive relations are satisfied:

$$\begin{aligned}\hat{V}_i^k(x, N) &= \max_{u_i^N \geq 0} \{m_i x_i^N - (u_i^N)^2\}, \\ \hat{V}_i^k(x, \tau) &= \max_{u_i^\tau \geq 0} \left\{ m_i x_i^\tau - (u_i^\tau)^2 + \frac{N - \tau}{N - \tau + 1} \hat{V}_i^k(x_i^{\tau+1}, \tau + 1) \right\}, \\ &\text{for } \tau \in \{\bar{T}_{k-1} + 1, \dots, N - 1\}, \\ V_i^k(x, f) &= \max_{u_i^f \geq 0} \left\{ m_i x_i^f - (u_i^f)^2 + V_i^k(x_i^{f+1}, f + 1) \right\}, \quad \text{for } f \in \overline{k, \bar{T}_{k-1}}. \\ V_i^k(x, \bar{T}_{k-1}) &= \hat{V}_i^k(x, \bar{T}_{k-1}),\end{aligned} \quad (2.48)$$

then the strategies u_i , $i \in I$, maximizing the right-hand side of the system (2.48) will constitute the feedback Nash equilibrium. The corresponding Bellman functions can be obtained as follows:

Theorem 2.5.2 *The Bellman functions for the feedback Nash equilibrium of the k^{th} truncated subgame in (2.48) are*

1

$$\hat{V}_i^k(x, \tau) = \sum_{i \in I} A_i^{k, \tau} x_i^\tau + B^{k, \tau}(i), \tau \in \{\bar{T}_{k-1} + 1, \dots, N - 1\}, \quad (2.49)$$

$$\begin{aligned} \text{where } A_i^{k, \tau} &= m_i - \left(\frac{N - \tau}{N - \tau + 1} \bar{Y}_i^{k, \tau+1} Z_i \right)^2 - \frac{N - \tau}{N - \tau + 1} A_i^{k, \tau+1} (\delta h - 1), \\ A_j^{k, \tau} &= -2 \left(\frac{N - \tau}{N - \tau + 1} \bar{Y}_j^{k, \tau+1} Z_j \right)^2 - \frac{N - \tau}{N - \tau + 1} A_j^{k, \tau+1} (\delta h - 1), \quad j \in I \setminus i, \\ B^{k, \tau}(i) &= \sum_{i \in I} \left[2 \left(\frac{N - \tau}{N - \tau + 1} \bar{Y}_i^{k, \tau+1} Z_i \right)^2 + \frac{N - \tau}{N - \tau + 1} \frac{A_i^{k, \tau+1} h \delta}{n} \right] + \\ &+ \frac{N - \tau}{N - \tau + 1} B^{k, \tau+1}(i) - \left(\frac{N - \tau}{N - \tau + 1} \bar{Y}_i^{k, \tau+1} Z_i \right)^2. \end{aligned}$$

2

$$V_i^k(x, f) = \sum_{i \in I} A_i^{k, f} x_i^f + B^{k, f}(i), f \in \{k, \dots, \bar{T}_{k-1}\}, \quad (2.50)$$

$$\begin{aligned} \text{where } A_i^{k, f} &= m_i - (\bar{Y}_i^{k, f+1} Z_i)^2 - A_i^{k, f+1} (\delta h - 1), \\ A_j^{k, f} &= -2 (\bar{Y}_j^{k, f+1} Z_j)^2 - A_j^{k, f+1} (\delta h - 1), \quad j \in I \setminus i, \\ B^{k, f}(i) &= \sum_{i \in I} \left[2 (\bar{Y}_i^{k, f+1} Z_i)^2 + \frac{A_i^{k, f+1} h \delta}{n} \right] + B^{k, f+1}(i) - (\bar{Y}_i^{k, f+1} Z_i)^2 \end{aligned}$$

with the initial conditions $A_i^{k, N} = m_i$, $B^{k, N+1}(i) = 0$, $i \in \{1, \dots, n\}$, $A_j^{k, N} = 0$, for $j \in I \setminus i$ and $Y_i^{k, t+1} = n A_i^{k, t+1} - \sum_{j \in I} A_j^{k, t+1}$, $\bar{Y}_i^{k, t+1} = \max \{0, Y_i^{k, t+1}\}$ $t \in \{k, \dots, N - 1\}$.

Proof Consider the Bellman function at the last stage (stage N). Performing the indicated maximization yields

$$\psi_i^N = 0, \quad \text{for } i \in I.$$

Using (2.49) and $\bar{V}_i^k(x_i^N, N) = m_i x_i^N$, we write the second equation in (2.48) as

$$\sum_{i \in I} A_i^{k, \tau} x_i^\tau + B^{k, \tau}(i) = \max_{u_i^\tau} \left\{ m_i x_i^\tau - (u_i^\tau)^2 + \frac{N - \tau}{N - \tau - 1} \left(\sum_{i \in I} A_i^{k, \tau+1} x_i^{\tau+1} + B^{k, \tau+1}(i) \right) \right\}.$$

Performing the indicated maximization in this expression yields

$$-2(u_i^\tau) + \frac{N - \tau}{N - \tau - 1} \left(\frac{h\rho_i}{n - 1} nA_i^{k,\tau+1} \sqrt{1 - x_i^{k,\tau}} - \frac{h\rho_i}{n - 1} \sqrt{1 - x_i^{k,\tau}} \sum_{j \in I} A_j^{k,\tau+1} \right) = 0$$

for $i \in I$ and $\tau \in \{\bar{T}_{k-1} + 1, \dots, N - 1\}$. Then the feedback Nash equilibrium strategy of firm $i \in I$ can be obtained in the form

$$\begin{aligned} \psi_i^\tau &= \max \left\{ 0, \frac{N - \tau}{N - \tau - 1} \left(nA_i^{k,\tau+1} - \sum_{j \in I} A_j^{k,\tau+1} \right) \frac{h\rho_i}{2(n - 1)} \sqrt{1 - x_i^{k,\tau}} \right\} = \\ &= \frac{N - \tau}{N - \tau - 1} \frac{h\rho_i \sqrt{1 - x_i^{k,\tau}}}{2(n - 1)} \max \left\{ 0, nA_i^{k,\tau+1} - \sum_{j \in I} A_j^{k,\tau+1} \right\}. \end{aligned}$$

Anticipating that the controls will be shown to be positive, we substitute this into the second equation in (2.48) and collect like terms together to express the parameters $A_i^{k,\tau}$ and $B^{k,\tau}(i)$ as follows:

$$\begin{aligned} A_i^{k,\tau} &= m_i - \left(\frac{N - \tau}{N - \tau + 1} \bar{Y}_i^{k,\tau+1} Z_i \right)^2 - \frac{N - \tau}{N - \tau + 1} A_i^{k,\tau+1} (\delta h - 1), \\ A_j^{k,\tau} &= -2 \left(\frac{N - \tau}{N - \tau + 1} \bar{Y}_j^{k,\tau+1} Z_j \right)^2 - \frac{N - \tau}{N - \tau + 1} A_j^{k,\tau+1} (\delta h - 1), \quad j \in I \setminus i, \\ B^{k,\tau}(i) &= \sum_{i \in I} \left[2 \left(\frac{N - \tau}{N - \tau + 1} \bar{Y}_i^{k,\tau+1} Z_i \right)^2 + \frac{N - \tau}{N - \tau + 1} \frac{A_i^{k,\tau+1} h \delta}{n} \right] + \\ &+ \frac{N - \tau}{N - \tau + 1} B^{k,\tau+1}(i) - \left(\frac{N - \tau}{N - \tau + 1} \bar{Y}_i^{k,\tau+1} Z_i \right)^2. \end{aligned}$$

At stage $f \in \{k, \dots, \bar{T}_{k-1}\}$, performing the indicated maximization in the third equation in (2.48) yields

$$\begin{aligned} \psi_i^f &= \max \left\{ 0, \left(nA_i^{k,f+1} - \sum_{j \in I} A_j^{k,f+1} \right) \frac{h\rho_i}{2(n - 1)} \sqrt{1 - x_i^{k,f}} \right\} = \\ &= \frac{h\rho_i \sqrt{1 - x_i^{k,f}}}{2(n - 1)} \max \left\{ 0, nA_i^{k,f+1} - \sum_{j \in I} A_j^{k,f+1} \right\}. \end{aligned}$$

Anticipating that the controls will be shown to be positive, we substitute this into the third equation in (2.48) and collect like terms together to express the parameters

$A_i^{k,f}$ and $B^{k,f}(i)$ as follows:

$$\begin{aligned} A_i^{k,f} &= m_i - (\bar{Y}_i^{k+1} Z_i)^2 - A_i^{k,f+1}(\delta h - 1), \\ A_j^{k,f} &= -2(\bar{Y}_j^{k,f+1} Z_j)^2 - A_j^{k,f+1}(\delta h - 1), \quad j \in I \setminus i, \\ B^{k,f}(i) &= \sum_{i \in I} [2(\bar{Y}_i^{k,f+1} Z_i)^2 + \frac{A_i^{k,f+1} h \delta}{n}] + B^{k,f+1}(i) - (\bar{Y}_i^{k,f+1} Z_i)^2, \end{aligned}$$

where $Y_i^{k,t+1} = nA_i^{k,t+1} - \sum_{j \in I} A_j^{k,t+1}$, $\bar{Y}_i^{k,t+1} = \max \left\{ 0, nA_i^{k,t+1} - \sum_{j \in I} A_j^{k,t+1} \right\}$, $t \in \{k, \dots, N-1\}$. \square

2.5.3 Cooperative Outcomes in Truncated Subgame

In this section, we focus on the cooperative setting, where all firms agree to cooperate in each truncated subgame. We denote by $\Gamma_k^c(x_k^0, k, k + \bar{T})$ the truncated cooperative subgame on the interval $[k, k + \bar{T}]$ with the initial condition x_k^0 . All firms aim to maximize their joint payoff:

$$\sum_{i \in I} \bar{K}_i^k(x_i^{k,0}, u) = \sum_{i \in I} \sum_{j=k}^{k+\bar{T}} m_i x_i^j - (u_i^j)^2. \quad (2.51)$$

Suppose that the set of strategies $\{u_i^{*,k,l}, \text{ for } i \in \{1, \dots, n\}, l = \overline{0, \bar{T}} \text{ and } k \in \{0, \dots, N - \bar{T}\}\}$, provides an optimal control in the game (2.42), (2.51). Due to Theorem 2.2, if there exist functions $\{W^k(l, x_k^l), x_k^l = (x_1^{k,l}, \dots, x_n^{k,l})\}$, for $i \in \{1, \dots, n\}, l = \overline{0, \bar{T}}$ and $k \in \{0, \dots, N - \bar{T}\}$ in each truncated subgame $\Gamma_k^c(x_k^0, k, k + \bar{T})$, then the following recursive relations are satisfied:

$$\begin{aligned} W^k(l, x_k^l) &= \max_{u_1^{k,l}, \dots, u_n^{k,l} \geq 0} \left\{ \sum_{i \in I} m_i x_i^{k,l} - (u_i^{k,l})^2 + W^k(l+1, x_k^{l+1}) \right\} = \\ &= \sum_{i \in I} m_i x_i^{k,l} - (u_i^{*,k,l})^2 + W^k(l+1, x_k^{l+1}). \end{aligned}$$

Theorem 2.5.3 *The maximum joint payoff of players in each truncated subgame $\Gamma_k^c(x_k^0, k, k + \bar{T})$ has the following form:*

$$W^k(l, x_k^l) = \sum_{i \in I} \hat{C}_i^{k,l} x_i^{k,l} + \hat{D}_i^{k,l}, \quad (2.52)$$

where $\hat{C}_i^{k,l}$, $\hat{D}_i^{k,l}$, $i \in \{1, \dots, n\}$, $k \in \{0, \dots, N - \bar{T}\}$, $l = \overline{0, \bar{T}}$, satisfy the relations

$$\begin{aligned}\hat{C}_i^{k,l} &= m_i - (\bar{G}_i^{k,l+1} Z_i)^2 - \hat{C}_i^{k,l+1}(\delta h - 1), \\ \hat{D}_i^{k,l} &= (\bar{G}_i^{k,l+1} Z_i)^2 + \frac{\hat{C}_i^{k,l+1} \delta h}{n} + D_i^{k,l+1}\end{aligned}$$

with the initial conditions $\hat{C}_i^{k, \bar{T}+1} = 0$, $\hat{D}_i^{k, \bar{T}+1} = 0$, $\bar{G}_i^{k,l+1} = n\hat{C}_i^{k,l+1} - \sum_{j \in I} \hat{C}_j^{k,l+1}$ and $\bar{G}_i^{k,l+1} = \max \left\{ 0, \bar{G}_i^{k,l+1} \right\}$.

Proof The Bellman equation in the cooperative setting is given by

$$W^k(l, x_i^{k,l}) = \max_{u_1^{k,l}, \dots, u_n^{k,l}} \sum_{i \in I} \left\{ m_i x_i^{k,l} - (u_i^{*,k,l})^2 + W^k(l+1, x_i^{k,l+1}) \right\}. \quad (2.53)$$

Substituting (2.52) into (2.53) yields

$$\begin{aligned}W^k(l, x_k^l) &= \max_{u_1^{k,l}, \dots, u_n^{k,l}} \sum_{i \in I} \left\{ m_i x_i^{k,l} - (u_i^{*,k,l})^2 + W^k(l+1, x_k^{l+1}) \right\} \\ \Rightarrow \sum_{i \in I} \hat{C}_i^{k,l} x_i^{k,l} + \hat{D}_i^{k,l} &= \max_{u_1^k, \dots, u_n^k} \left\{ \sum_{i \in I} [m_i x_i^{k,l} - (u_i^{k,l})^2] + \sum_{i \in I} \hat{C}_i^{k,l+1} x_i^{k,l+1} + \hat{D}_i^{k,l+1} \right\},\end{aligned} \quad (2.54)$$

$$W^k(\bar{T} + 1, x_k^{\bar{T}}) = 0.$$

We substitute the right-hand side of (2.42) into (2.54). Performing the indicated maximization in (2.54), we obtain

$$-2u_i^k + (n\hat{C}_i^{k,l+1} - \sum_{j \in I} \hat{C}_j^{k,l+1}) \frac{h\rho_i \sqrt{1 - x_i^{k,l}}}{n-1} = 0.$$

The optimal cooperative strategies can be calculated in the form

$$\begin{aligned}u_i^{*,k,l} &= \max \left\{ 0, (n\hat{C}_i^{k,l+1} - \sum_{j \in I} \hat{C}_j^{k,l+1}) \frac{h\rho_i \sqrt{1 - x_i^{k,l}}}{2(n-1)} \right\} = \\ &= \frac{h\rho_i \sqrt{1 - x_i^{k,l}}}{2(n-1)} \max \left\{ 0, n\hat{C}_i^{k,l+1} - \sum_{j \in I} \hat{C}_j^{k,l+1} \right\}\end{aligned}$$

for $i \in I$, $l = \overline{0, \bar{T}}$ and $k \in \{0, \dots, N\}$.

Anticipating that the cooperative strategies will be shown to be positive, we substitute it into (2.54). Hence, $\hat{C}_i^{k,l}$ and $\hat{D}_i^{k,l}$ are defined as

$$\begin{aligned}\hat{C}_i^{k,l} &= m_i - (\bar{G}_i^{k,l+1} Z_i)^2 - \hat{C}_i^{k,l+1}(\delta h - 1), \\ \hat{D}_i^{k,l} &= (\bar{G}_i^{k,l+1} Z_i)^2 + \frac{\hat{C}_i^{k,l+1} \delta h}{n} + D_i^{k,l+1}.\end{aligned}$$

□

The optimal cooperative strategies in each truncated subgame have the form

$$u_i^{*,k,l} = \bar{G}_i^{k,l+1} Z_i \sqrt{1 - x_i^{k,l}}, \forall i \in I.$$

We denote by $x_k^{*,l} = (x_1^{*,k,l}, \dots, x_n^{*,k,l})$, $l = \overline{0, \bar{T}}$, the solution of (2.52) with the optimal strategies u^* involved with the cooperative trajectory. The cooperative trajectory in each truncated subgame can be calculated using (2.42) as follows, where $k \in \{0, \dots, N - \bar{T}\}$:

$$x_i^{*,k,l+1} = 2n \bar{G}_i^{k,l+1} Z_i^2 (1 - x_i^{*,k,l}) - \sum_{j \in I} 2 \bar{G}_j^{k,l+1} Z_j^2 (1 - x_j^{*,k,l}) - (h\delta - 1)x_i^{*,k,l} + \frac{h\delta}{n}.$$

So far, we have formalized the case when all firms cooperate in the game model with dynamic updating and a fixed information horizon. Hence, we start to study the case when the information horizon is random, where \bar{T}_1 has symmetric probabilities on the interval $[1, N]$, $\gamma_1 = \dots = \gamma_N = \frac{1}{N}$. In the k^{th} cooperative truncated subgame, if there exist functions $W^k(x, f)$, for $i \in \{1, \dots, n\}$, $f \in \{k, \dots, \bar{T}_{k-1}\}$, and $\hat{W}^k(x, \tau)$, for $\tau \in \{\bar{T}_{k-1} + 1, \dots, N\}$, such that the following recursive relations

are satisfied:

$$\begin{aligned}
\hat{W}^k(x, N) &= \max_{u_1^N, \dots, u_n^N} \left\{ \sum_{i \in I} m_i x_i^N - (u_i^N)^2 \right\}, \\
\hat{W}^k(x, \tau) &= \max_{u_1^\tau, \dots, u_n^\tau} \left\{ \sum_{i \in I} m_i x_i^\tau - (u_i^\tau)^2 + \frac{N - \tau}{N - \tau + 1} \hat{W}^k(x_i^{\tau+1}, \tau + 1) \right\}, \\
&\text{for } \tau \in \{\bar{T}_{k-1} + 1, \dots, N - 1\}, \\
W^k(x, f) &= \max_{u_1^f, \dots, u_n^f} \left\{ \sum_{i \in I} m_i x_i^f - (u_i^f)^2 + W^k(x_i^{f+1}, f + 1) \right\}, \text{ for } f \in \overline{k, \bar{T}_{k-1}}, \\
W_i^k(x, \bar{T}_{k-1}) &= \hat{W}_i^k(x, \bar{T}_{k-1}),
\end{aligned} \tag{2.55}$$

then the strategies $\{u_i^{*,k,l}, \text{ for } i \in \{1, \dots, n\}, l = \overline{0, \bar{T}} \text{ and } k \in \{0, \dots, N - \bar{T}\}\}$, are optimal.

The value function can be obtained as follows.

Theorem 2.5.4 *The Bellman functions in the k^{th} cooperative truncated subgame in (2.55) are*

1

$$\hat{W}^k(x, \tau) = \sum_{i \in I} \tilde{C}_i^{k,\tau} x_i^\tau + \tilde{D}_i^{k,\tau}, \tau \in \{\bar{T}_{k-1} + 1, \dots, N - 1\}, \tag{2.56}$$

$$\begin{aligned}
\text{where } \tilde{C}_i^{k,\tau} &= m_i - \left(\frac{N - \tau}{N - \tau + 1} \hat{G}_i^{k,\tau+1} Z_i \right)^2 - \frac{N - \tau}{N - \tau + 1} \tilde{C}_i^{k,\tau+1} (\delta h - 1), \\
\tilde{D}_i^{k,\tau} &= \left(\frac{N - \tau}{N - \tau + 1} \hat{G}_i^{k,\tau+1} Z_i \right)^2 + \frac{N - \tau}{N - \tau + 1} \frac{\tilde{C}_i^{k,\tau+1} h \delta}{n} + \frac{N - \tau}{N - \tau + 1} \tilde{D}_i^{k,\tau+1};
\end{aligned}$$

2

$$W_i^k(x, f) = \sum_{i \in I} \tilde{C}_i^{k,f} x_i^f + \tilde{D}_i^{k,f}, f \in \{k, \dots, \bar{T}_{k-1}\}, \tag{2.57}$$

$$\begin{aligned}
\text{where } \tilde{C}_i^{k,l} &= m_i - \left(\hat{G}_i^{k,l+1} Z_i \right)^2 - \tilde{C}_i^{k,l+1} (\delta h - 1), \\
\tilde{D}_i^{k,l} &= \left(\hat{G}_i^{k,l+1} Z_i \right)^2 + \frac{\tilde{C}_i^{k,l+1} \delta h}{n} + \tilde{D}_i^{k,l+1},
\end{aligned}$$

with the initial conditions $\tilde{C}_i^{k,N} = m_i$, $\tilde{D}_i^{k,N} = 0$, $i \in \{1, \dots, n\}$, $\hat{G}_i^{k,t+1} = n \tilde{C}_i^{k,t+1} - \sum_{j \in I} \tilde{C}_j^{k,t+1}$, $\hat{G}_i^{k,t+1} = \max \{0, \hat{G}_i^{k,t+1}\}$, $t \in \{k, \dots, N - 1\}$.

Proof Consider the Bellman function at the last stage (stage N). Performing the indicated maximization in it yields

$$\phi_i^N = 0, \text{ for } i \in I.$$

Using (2.56) and $\bar{W}_i^k(x_i^N, N) = \sum_{i \in I} m_i x_i^N$, we write the second equation in (2.55) as

$$\sum_{i \in I} \tilde{C}_i^{k,\tau} x_i^\tau + \tilde{D}_i^{k,\tau} = \max_{u_1^\tau, \dots, u_n^\tau} \left\{ \sum_{i \in I} m_i x_i^\tau - (u_i^\tau)^2 + \frac{N - \tau}{N - \tau + 1} (\tilde{C}_i^{k,\tau+1} x_i^{\tau+1} + \tilde{D}_i^{k,\tau+1}) \right\}.$$

Performing the indicated maximization in this expression yields

$$-2(u_i^\tau) + \frac{N - \tau}{N - \tau - 1} \left(\frac{h\rho_i}{n - 1} n \tilde{C}_i^{k,\tau+1} \sqrt{1 - x_i^{k,\tau}} - \frac{h\rho_i}{n - 1} \sqrt{1 - x_i^{k,\tau}} \sum_{j \in I} \tilde{C}_j^{k,\tau+1} \right) = 0$$

for $i \in I$ and $\tau \in \{\bar{T}_{k-1} + 1, \dots, N - 1\}$. Then the feedback Nash equilibrium strategy of firm $i \in I$ can be obtained in the form

$$\begin{aligned} \phi_i^\tau &= \max \left\{ 0, \frac{N - \tau}{N - \tau - 1} \left(n \tilde{C}_i^{k,\tau+1} - \sum_{j \in I} \tilde{C}_j^{k,\tau+1} \right) \frac{h\rho_i}{2(n - 1)} \sqrt{1 - x_i^{k,\tau}} \right\} = \\ &= \frac{N - \tau}{N - \tau - 1} \frac{h\rho_i \sqrt{1 - x_i^{k,\tau}}}{2(n - 1)} \max \left\{ 0, n \tilde{C}_i^{k,\tau+1} - \sum_{j \in I} \tilde{C}_j^{k,\tau+1} \right\}. \end{aligned}$$

Anticipating that the controls will be shown to be positive, we substitute this into the second equation in (2.55) and collect like terms together to express the parameters $\tilde{C}_i^{k,\tau}$ and $\tilde{D}_i^{k,\tau}$ as follows:

$$\begin{aligned} \tilde{C}_i^{k,\tau} &= m_i - \left(\frac{N - \tau}{N - \tau + 1} \hat{G}_i^{k,\tau+1} Z_i \right)^2 - \frac{N - \tau}{N - \tau + 1} \tilde{C}_i^{k,\tau+1} (\delta h - 1), \\ \tilde{D}_i^{k,\tau} &= \left(\frac{N - \tau}{N - \tau + 1} \hat{G}_i^{k,\tau+1} Z_i \right)^2 + \frac{N - \tau}{N - \tau + 1} \frac{\tilde{C}_i^{k,\tau+1} h\delta}{n} + \frac{N - \tau}{N - \tau + 1} \tilde{D}_i^{k,\tau+1}. \end{aligned}$$

For stage $f \in \{k, \dots, \bar{T}_{k-1}\}$, performing the indicated maximization in third equation of (2.55), we obtain

$$\phi_i^f = \max \left\{ 0, \left(n \tilde{C}_i^{k,\tau+1} - \sum_{j \in I} \tilde{C}_j^{k,\tau+1} \right) \frac{h\rho_i}{2(n - 1)} \sqrt{1 - x_i^{k,f}} \right\} =$$

$$= \frac{h\rho_i\sqrt{1-x_i^{k,f}}}{2(n-1)} \max \left\{ 0, n\tilde{C}_i^{k,\tau+1} - \sum_{j \in I} \tilde{C}_j^{k,\tau+1} \right\}.$$

Anticipating that the controls will be shown to be positive, we substitute this into the third equation in (2.55) and collect like terms together to express the parameters $\tilde{C}_i^{k,f}$ and $\tilde{D}_i^{k,f}$ as follows:

$$\begin{aligned} \tilde{C}_i^{k,l} &= m_i - (\hat{G}_i^{k,l+1} Z_i)^2 - \tilde{C}_i^{k,l+1}(\delta h - 1), \\ \tilde{D}_i^{k,l} &= (\hat{G}_i^{k,l+1} Z_i)^2 + \frac{\tilde{C}_i^{k,l+1}\delta h}{n} + D_i^{k,l+1}, \end{aligned}$$

where $\hat{G}_i^{k,t+1} = n\tilde{C}_i^{k,t+1} - \sum_{j \in I} \tilde{C}_j^{k,t+1}$, $\hat{G}_i^{k,t+1} = \max \left\{ 0, \hat{G}_i^{k,t+1} \right\}$, $t \in \{k, \dots, N-1\}$.

□

2.5.4 Characteristic Function in Truncated Subgame

Using the characteristic function formula (2.14), we define the characteristic function for each truncated subgame:

$$V^k(S; x_k^{*,0}) = \begin{cases} \sum_{i \in I} K_i^k(x_k^{*,0}; u), & S = I, \\ \tilde{V}_k(S, x_k^{*,0}), & S \subset I, \\ 0, & S = \emptyset, \end{cases} \quad (2.58)$$

where $\tilde{V}_k(S, x_k^{*,0})$ is defined as the total payoff of players from coalition S in the feedback Nash equilibrium $u^{NE} = (\bar{u}_1, \dots, \bar{u}_n)$ in the game with the following set of players: coalition S (acting as one player) and players from the set $I \setminus S$, i.e., in the game with $|I \setminus S| + 1$ players.

Suppose that the Bellman function for coalition $S \subseteq I$ and players $i \in I \setminus S$ in the truncated subgame $\Gamma_k^c(x_k^0, k, k + \bar{T})$ has the following form:

$$\begin{aligned} V^k(S, x_k^0) &= \sum_{i \in I} \tilde{C}_i^{k,0} x_i^{k,0} + \bar{E}^{k,0}, \quad S \subseteq I \\ V^k(\{i\}, x_k^0) &= \sum_{i \in I} \hat{A}_i^{k,0} x_i^{k,0} + \hat{B}(i)^{k,0}, \quad i \in I \setminus S. \end{aligned}$$

Then the corresponding system of Bellman equations is given by

$$\begin{cases} V^k(S, x_k^0) = \max_{u_i^k, \forall i \in S} \left\{ \sum_{i \in S} [m_i x_i^{k,0} - (u_i^k)^2] + V^k(S, x_k^1) \right\}, \\ V^k(\{i\}, x_k^0) = \max_{u_i^k} \left\{ m_i x_i^{k,0} - (u_i^k)^2 + V^k(\{i\}, x_k^1) \right\}, \quad \forall i \in I \setminus S, \\ V^k(S, x_k^{\bar{T}+1}) = 0, \\ V^k(\{i\}, x_k^{\bar{T}+1}) = 0. \end{cases} \quad (2.59)$$

2.5.5 Numerical Simulation for Dynamic Advertising Game Model with Updating

We consider a specific three-firm oligopoly case on the stage interval with $N = 8$ and a fixed information horizon $\bar{T} = 2$. To illustrate our model, let the parameters for each firm be $\rho = [0.4, 0.5, 0.3]$, $h = 0.4$, $\delta = 0.09$, $m = [0.6, 1, 1.2]$, and the initial conditions be $z^0 = [0.3, 0.5, 0.2]$.

Figures 2.2 and 2.3 show the feedback Nash equilibrium strategies for each firm in the noncooperative case and cooperative strategies in the cooperative case, respectively, for both the original game model and the game model with dynamic updating.

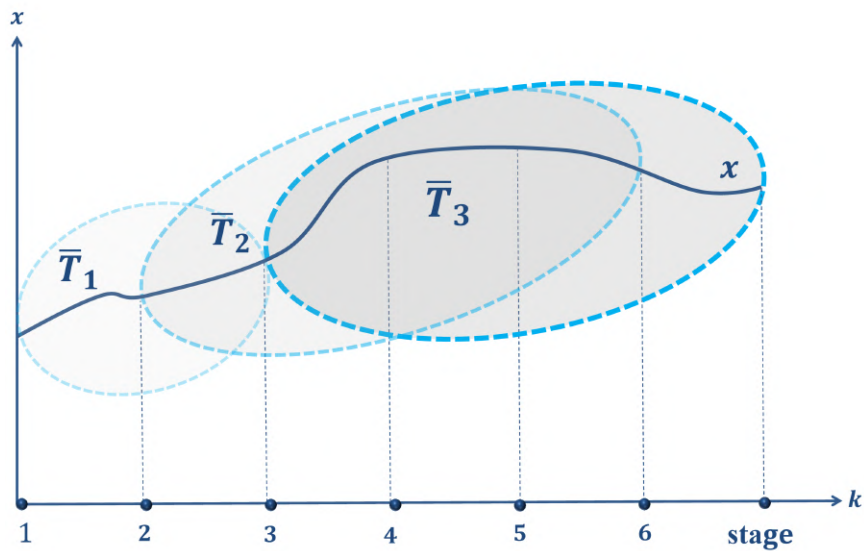


Figure 2.1: Length of each blue oval is a realization of infinite horizon.

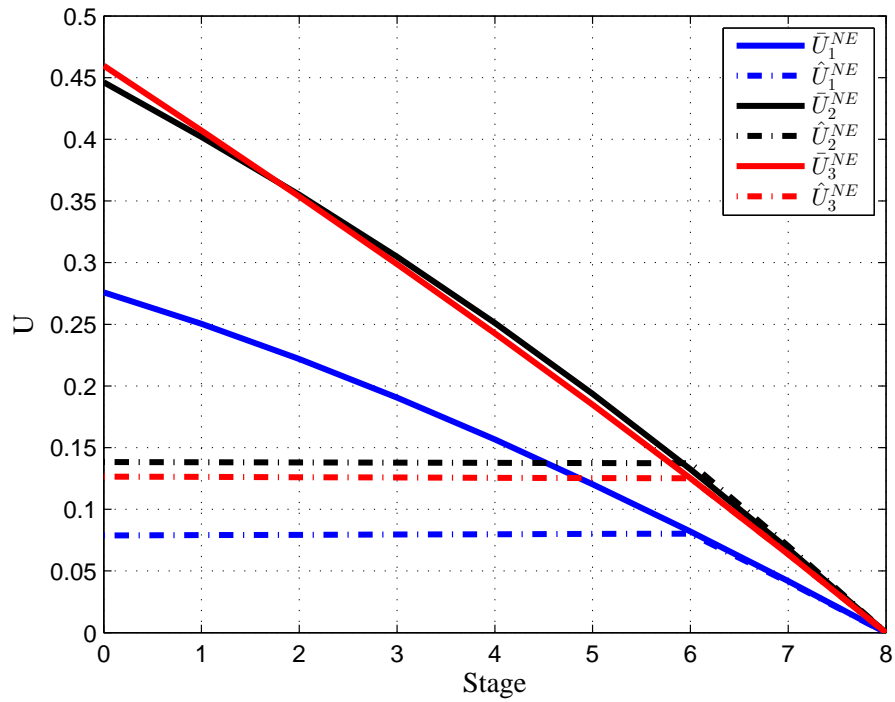


Figure 2.2: Feedback Nash equilibrium in noncooperative original game (solid line) and feedback Nash equilibrium in noncooperative game model with dynamic updating (dashed line).

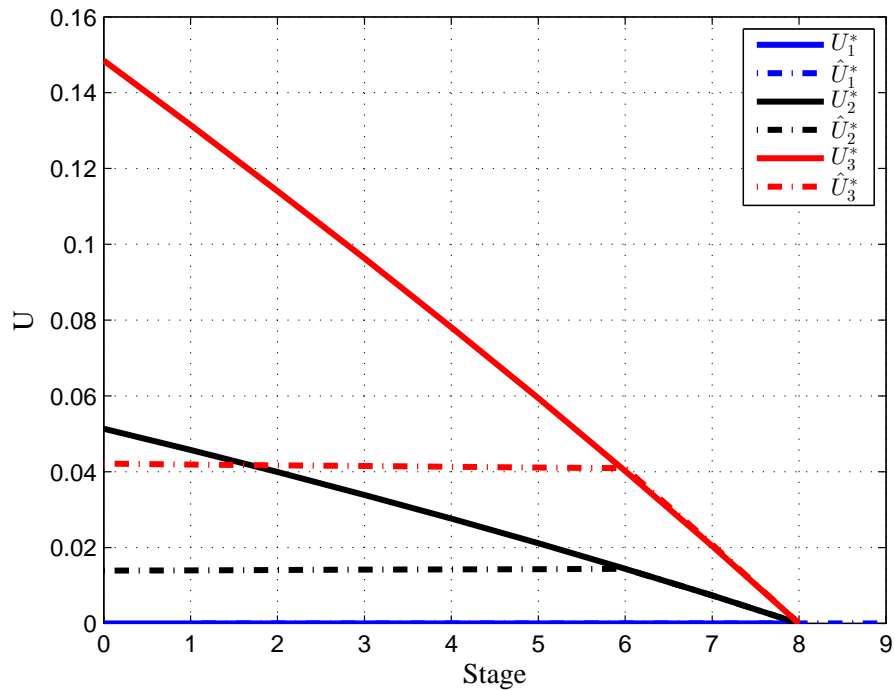


Figure 2.3: Optimal cooperative strategies in cooperative original game (solid line) and optimal cooperative strategies in cooperative game model with dynamic updating (dashed line).

Figures 2.4 and 2.5 show the dynamics of players' market shares corresponding to the feedback Nash equilibrium strategies for each firm in the noncooperative case and the dynamics of market shares corresponding to the optimal cooperative strategies in the cooperative case, respectively, for both the original game and the game model with dynamic updating.

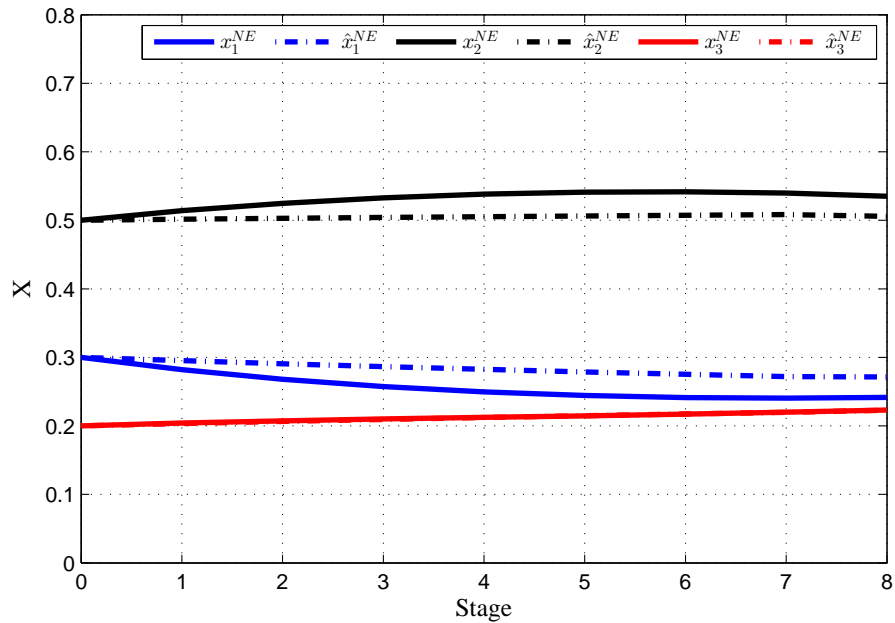


Figure 2.4: Noncooperative trajectory in original game (solid line) and resulting noncooperative trajectory (dashed line) in game model with dynamic updating.

Figures 2.6 and 2.7 show the optimal payoff of each firm in the noncooperative case and the characteristic function in the cooperative case for coalitions $\{1, 2\}$, $\{1, 3\}$, and $\{2, 3\}$, respectively, for both the original game and the game model with dynamic updating.

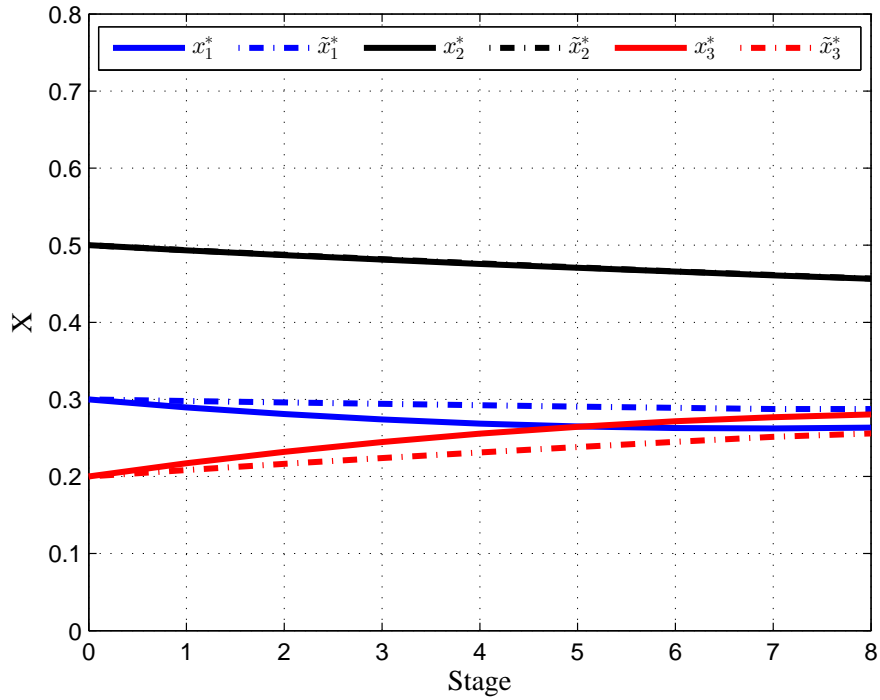


Figure 2.5: Cooperative trajectory in original game (solid line) and resulting cooperative trajectory (dashed line) in game model with dynamic updating.

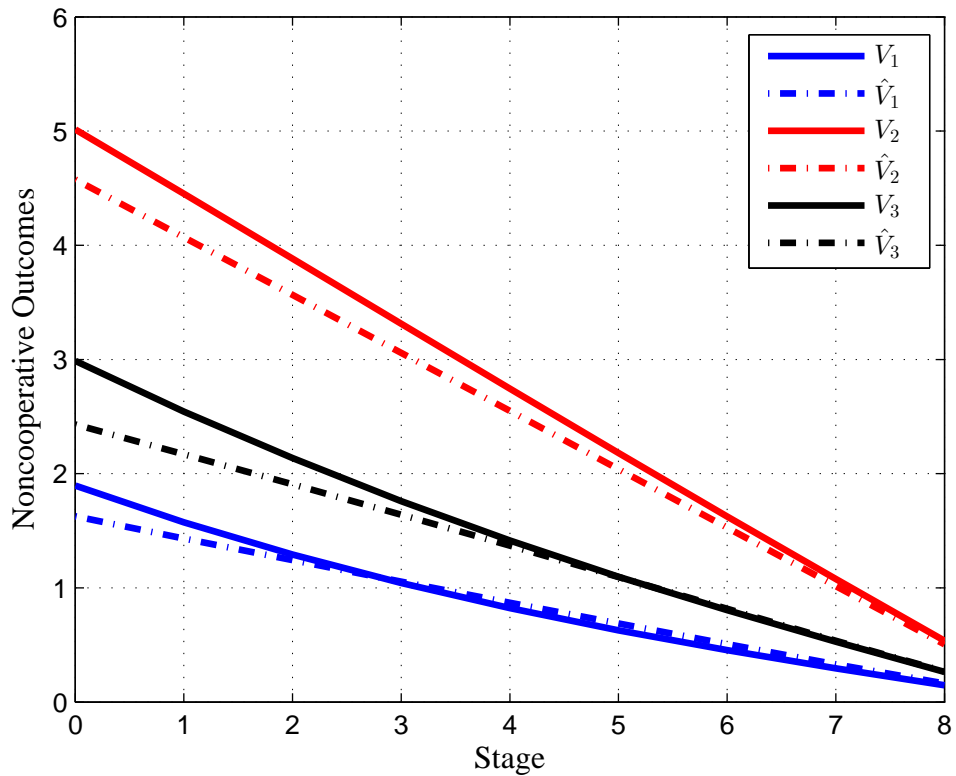


Figure 2.6: Noncooperative outcomes in original game (solid line) and resulting noncooperative outcomes (dashed line) in game model with dynamic updating.

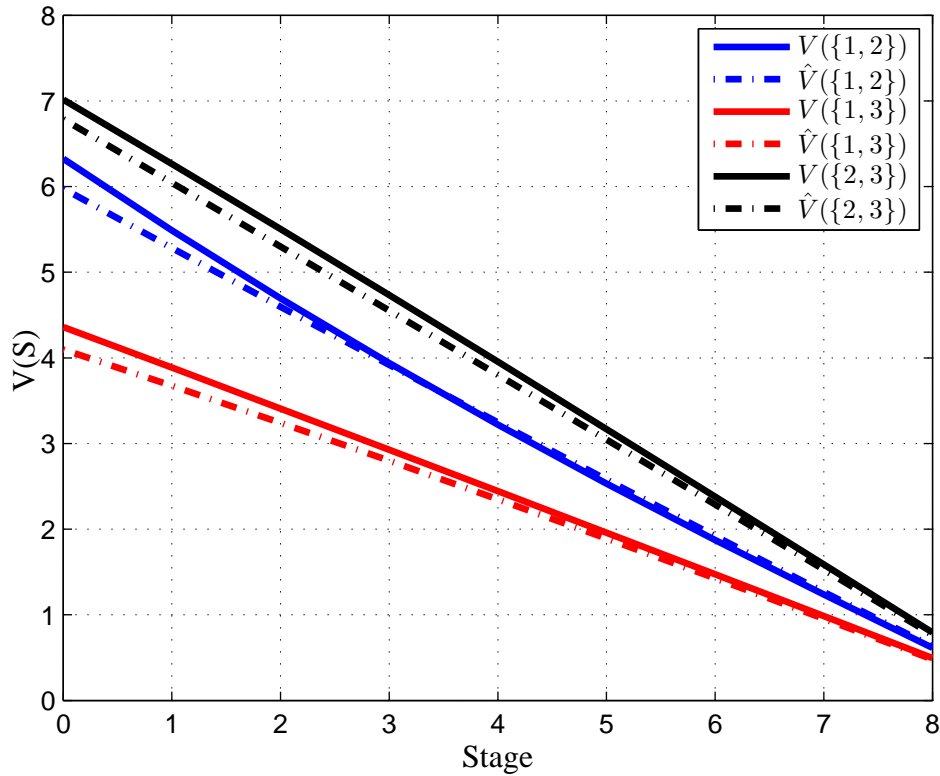


Figure 2.7: Characteristic function in original game model (solid line) and resulting characteristic function (dashed line) in game model with dynamic updating.

Figure 2.8 shows the difference between the resulting Shapley value and the Shapley value in the original game model. As we can see, the resulting Shapley value of the game with moving information horizon changes more steadily over stages rather than the Shapley value of the original game. It can be seen that the competition among the companies is more fierce in the original game rather than in the game with a moving information horizon, considering the market shares in Figure 2.5 and the advertising expenditure in Figure 2.6. The Shapley value of each company barely changes at the last stage, which means that the profit of each company does not change much; see Figure 2.8.

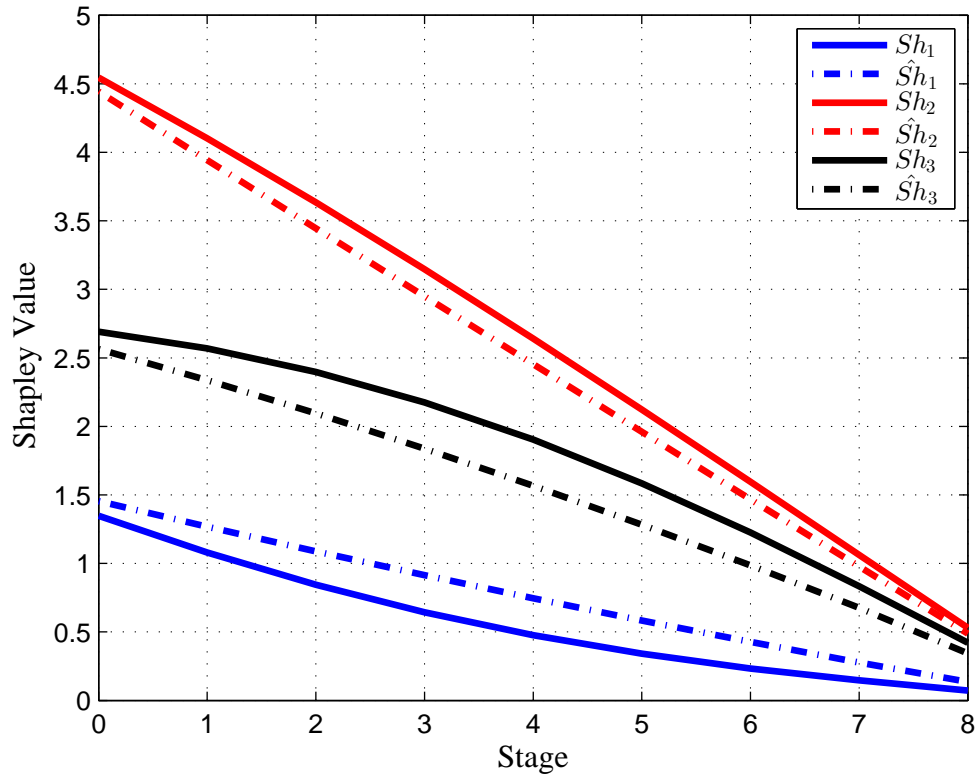


Figure 2.8: Shapley value in original game (solid line) and resulting Shapley value (dashed line) in game model with dynamic updating.

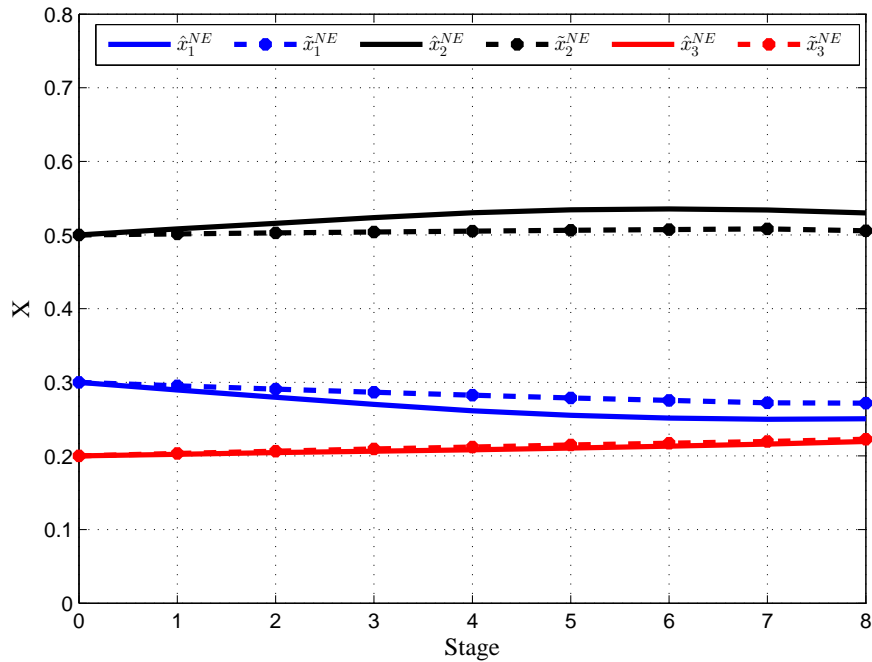


Figure 2.9: Resulting noncooperative trajectory with a fixed information horizon (dashed line) and resulting noncooperative trajectory with a random information horizon (solid line).

Suppose that the realizations of the game information horizon are $\bar{T}_1 = 2$, $\bar{T}_2 = 5$, and $\bar{T}_3 = 8$. Figure 2.9 shows the difference between the resulting noncooperative trajectories with the fixed and random information horizons.

Figure 2.10 shows the difference between the resulting noncooperative outcomes of each firm in the noncooperative case with the fixed and random information horizons.

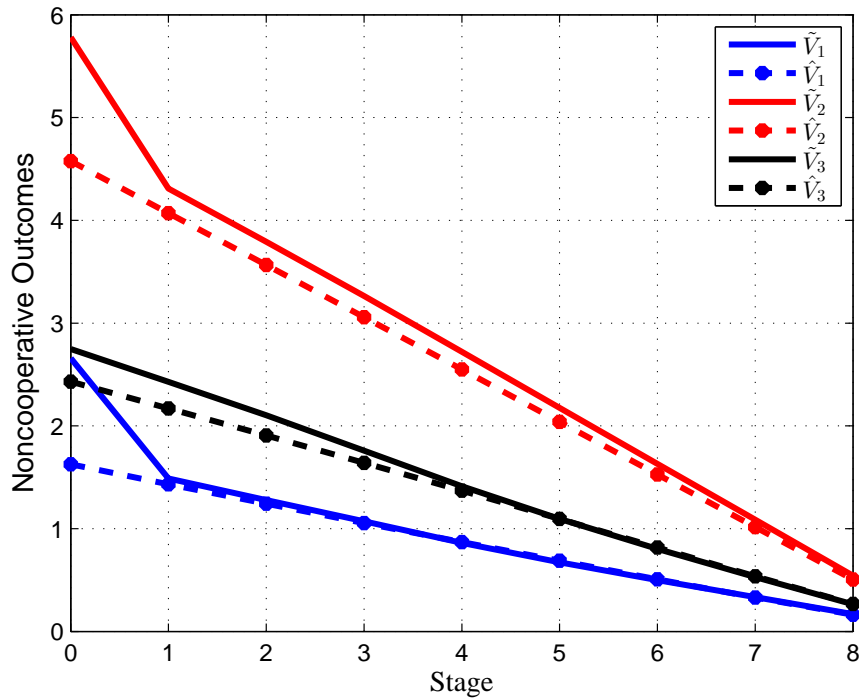


Figure 2.10: Resulting noncooperative outcomes with a fixed information horizon (dashed line) and resulting noncooperative outcomes with a random information horizon (solid line).

Figure 2.11 shows the difference between the resulting cooperative trajectories with the fixed and random information horizons.

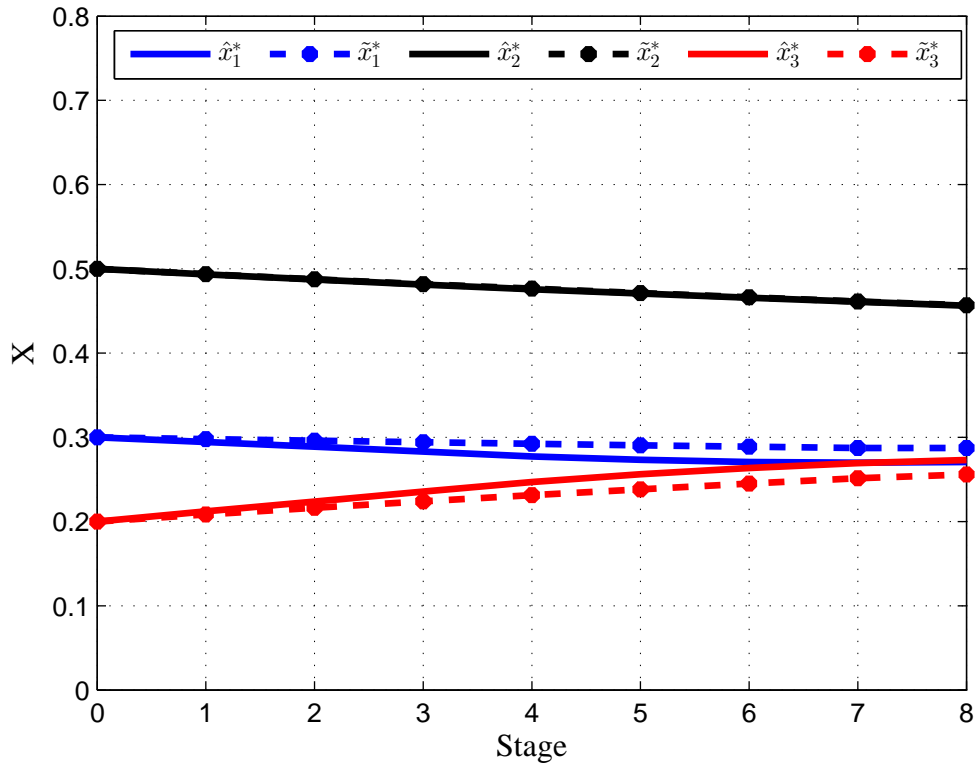


Figure 2.11: Resulting cooperative trajectory with a fixed information horizon (dashed line) and resulting cooperative trajectory with a random information horizon (solid line).

Chapter 3

Class of Differential Games with Dynamic Updating

This chapter presents a complete description of the dynamic updating approach for cooperative differential games with transferable utility. Here the difference between the previous chapter is that the dynamic updating approach is applied to the differential game model. The approach is used for constructing game-theoretic models and defining solutions for conflict-controlled processes where information about the process is updated dynamically. The time-consistency property of cooperative solutions with dynamic updating and the properties related to the transition from the solution used on truncated intervals to the cooperative solutions defined on the whole time interval are studied. Specific information updating models are considered, such as updating with an uncertain stochastic term and random horizon updating. At the end of this chapter, an applied oil market model is studied: the model of a real-life cooperative agreement is presented as a coalitional cooperative game with non-transferable utility, and historical data analysis is performed.

3.1 On General Class of Cooperative Differential Games with Dynamic Updating

3.1.1 Original Differential Game Model

Consider an n -player differential game $\Gamma(x_0, T - t_0)$ with a finite horizon $T - t_0$, an initial state $x_0 \in X \subset R^m$, where X is the state space, and an initial time instant t_0 (t_0 and T are fixed values). We denote by $N = \{1, \dots, n\}$ the set of players. At each time instant, player $i \in N$ chooses a control (strategy) $u_i \in U_i \subset \text{Comp}R^k$. The joint strategy (control) space is denoted by $U = U_1 \times \dots \times U_n$. The payoff function of player $i \in N$ has the form

$$K_i(x_0, T - t_0; u) = \int_{t_0}^T h_i(\tau, x(\tau), u(\tau, x)) e^{-r(\tau-t_0)} d\tau \quad (3.1)$$

subject to the dynamical system

$$\dot{x} = g(t, x(t), u(t, x)), \quad x(t_0) = x_0, \quad (3.2)$$

where $x(t) \in X \subset R^m$ is the trajectory (solution) of the system (3.2) with the control input $u = (u_1, \dots, u_n) \in U$; $r \geq 0$; the functions $h_i(t, x, u) : [t_0, T] \times X \times U \rightarrow R$ and $g(t, x, u) : [t_0, T] \times X \times U \rightarrow R$ are differentiable. The function $h_i(t, x, u)$ shows an instant payoff obtained by player $i \in N$ in a state $x(t) \in R^m$ with a control input $u = (u_1, \dots, u_n)$. The function $e^{-r(t-t_0)}$ defines the discount factor for the players' payoff. The solution of the system (3.2) determines the trajectory of the game.

When open-loop strategies are used, piece-wise continuity with a finite number of breaks is required. For feedback strategies, the approach from [180] is used, i.e. it is supposed that any set of strategies $u(t, x) = (u_1(t, x), \dots, u_n(t, x))$ is admissible as defined in section 1.1.1. For any n -tuple of strategies $u(t, x) = (u_1(t, x), \dots, u_n(t, x))$, it is required that the solution of the Cauchy problem (3.2) exists, unique and continuable solution on the time interval $[t_0, T]$. A significantly more sophisticated approach necessary for defining strategies in an zero-sum differential game is proposed and implemented in [181].

3.1.2 Truncated Subgame

As mentioned above, information about the functions $g(t, x, u)$ and $h_i(t, x, u)$ is updated at fixed time instants. Suppose that the interval $[t_0, T] \subset R$ is divided into $l = \frac{T-t_0}{\Delta t} - 1$ subintervals $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, l$ of an equal length $\Delta t \geq 0$. Then during the time interval $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, the players have complete information about the game dynamics $g(t, x, u)$ and the payoff function $h_i(t, x, u)$ on the time interval $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, where \bar{T} is a fixed value, namely the information horizon. More precisely, information about the game is described in the following way: during the interval $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, the game dynamics are $g(t, x, u) = g_k(t, x, u)$ and the payoff is $h_i(t, x, u) = h_i^k(t, x, u)$, where the functions $g_k(t, x, u), h_i^k(t, x, u) : [t_0 + k\Delta t, t_0 + (k+1)\Delta t] \times X \times R$, $k = j, \dots, j + \frac{\bar{T}}{\Delta t}$. At the time instant $t = t_0 + (j+1)\Delta t$, the information about the game dynamics $g(t, x, u)$ and payoff $h_i(t, x, u)$ is being updated. On the next time interval $(t_0 + (j+1)\Delta t, t_0 + (j+2)\Delta t]$, the players have complete information about $g(t, x, u)$ and $h_i(t, x, u)$ on the time interval $(t_0 + (j+1)\Delta t, t_0 + (j+1)\Delta t + \bar{T}]$. At the time instant $t = t_0 + (j+1)\Delta t$, the players receive additional information about the motion equation $g(t, x, u) = g_k(t, x, u)$ and payoff $h_i(t, x, u) = h_i^k(t, x, u)$ on the interval $(t_0 + k\Delta t, t_0 + (k+1)\Delta t + \bar{T}]$, where $k = j + 1 + \frac{\bar{T}}{\Delta t}$. This new information about the functions $g(t, x, u)$ and $h_i(t, x, u)$ has the meaning of dynamic updating: the new information is the updating itself.

To model this type of information structure, the following definition is introduced, where $x_{j,0} = x(t_0 + j\Delta t)$.

Definition 3.1.1 *Let $j = 0, \dots, l$. A truncated subgame $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ is defined on the time interval $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$. The dynamical system and the payoff function on the time interval $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ coincide with those of the game $\Gamma(x_0, T - t_0)$ on the same time interval. The payoff function of player $i \in N$ in the truncated subgame j is*

$$K_i^j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; u) = \int_{t_0 + j\Delta t}^{t_0 + j\Delta t + \bar{T}} h_i(\tau, x(\tau), u(\tau)) e^{-r(\tau - t_0)} d\tau \quad (3.3)$$

subject to the dynamical system

$$\dot{x} = g(t, x, u), \quad x(t_0 + j\Delta t) = x_{j,0} \quad (3.4)$$

with the initial condition $x_{j,0}$ of the truncated subgame $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$.

3.1.2.1 Solution of Cooperative Truncated Subgame

Consider a truncated cooperative subgame $\bar{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ defined on the time interval $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ with the initial condition $x(t_0 + j\Delta t) = x_{j,0}$. The total payoff of players to be maximized in this game is

$$\sum_{i \in N} K_i^j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; u) = \sum_{i \in N} \int_{t_0 + j\Delta t}^{t_0 + j\Delta t + \bar{T}} h_i(\tau, x(\tau), u(\tau)) e^{-r(\tau - t_0)} d\tau \quad (3.5)$$

subject to

$$\dot{x} = g(t, x, u), \quad x(t_0 + j\Delta t) = x_{j,0}. \quad (3.6)$$

This is an optimal control problem. Sufficient conditions for the solution and the optimal feedback are given by the following assertion [182]. We denote by $W^{(j\Delta t)}(t, x)$ the maximum value of the joint payoff (3.5):

$$W^{(j\Delta t)}(t, x) = \max_{u \in U} \left\{ \sum_{i \in N} K_i^j(x, t, t_0 + j\Delta t + \bar{T}; u) \right\}, \quad (3.7)$$

where x, t are the initial state and time of the subgame of the truncated game, respectively, and $U = U_1 \times \dots \times U_n$.

Theorem 3.1.1 *Assume that there exists a continuously differential function $W^{(j\Delta t)}(t, x) : [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}] \times R^m \rightarrow R$ satisfying the partial differential equation*

$$-W_t^{(j\Delta t)}(t, x) = \max_{u \in U} \left\{ \sum_{i=1}^n h_i(t, x, u) e^{-r(t-t_0)} + W_x^{(j\Delta t)}(t, x) g(t, x, u) \right\}, \quad (3.8)$$

where $\lim_{t \rightarrow -t_0 + j\Delta t + \bar{T}} W^{(j\Delta t)}(t, x) = 0$ and the maximum in (3.8) is achieved under controls $u_j^*(t, x)$. Then $u_j^*(t, x)$ is optimal in the control problem defined by (3.5), (3.6).

Theorem 3.1.1 requires that the function $W^{(j\Delta t)}$ be C^1 . However, it is possible to assume continuity only considering viscosity solutions using the Subbotin approach [183; 184]. But due to the shortage of space, it is not possible to properly introduce and define this solution here. In the model example, the solution $W^{(j\Delta t)}$ will be obtained from C^1 .

3.1.2.2 Conditionally Cooperative Trajectory

During the game $\Gamma(x_0, T - t_0)$ the players possess only truncated information about its structure. Obviously, this is not enough to construct optimal control and the corresponding trajectory for the game $\Gamma(x_0, T - t_0)$. As a cooperative trajectory in the game $\Gamma(x_0, T - t_0)$, a conditionally cooperative trajectory will be used:

Definition 3.1.2 *A conditionally cooperative trajectory $\{\hat{x}^*(t)\}_{t=t_0}^T$ is defined as a composition of the cooperative trajectories $x_j^*(t)$ in the truncated cooperative subgames $\bar{\Gamma}_j^c(x_{j-1}^*(t_0 + j\Delta t), t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ defined on the successive time intervals $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$:*

$$\{\hat{x}^*(t)\}_{t_0}^T = \begin{cases} x_0^*(t), & t \in [t_0, t_0 + \Delta t), \\ \vdots \\ x_j^*(t), & t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t), \\ \vdots \\ x_l^*(t), & t \in [t_0 + l\Delta t, t_0 + (l + 1)\Delta t), \end{cases} \quad (3.9)$$

On the time interval $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, a conditionally cooperative trajectory coincides with the cooperative trajectory $x_j^*(t)$ in the truncated cooperative subgame $\bar{\Gamma}_j^c(x_{j-1}^*(t_0 + j\Delta t), t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, where $x_{j-1}^*(t_0 + j\Delta t)$ is the position of the game at the instant $t = t_0 + j\Delta t$ on the cooperative trajectory of the previous subgame $\bar{\Gamma}_j^c(x_{j-2}^*(t_0 + (j - 1)\Delta t), t_0 + (j - 1)\Delta t, t_0 + (j - 1)\Delta t + \bar{T})$. In the position $x_j^*(t_0 + (j + 1)\Delta t)$ at the time instant $t = t_0 + (j + 1)\Delta t$, information about

the game structure is updated. On the time interval $(t_0 + (j+1)\Delta t, t_0 + (j+2)\Delta t]$, the trajectory $\hat{x}^*(t)$ coincides with the cooperative trajectory $x_{j+1}^*(t)$ in the truncated cooperative subgame $\bar{\Gamma}_{j+1}^c(x_j^*(t_0 + (j+1)\Delta t), t_0 + (j+1)\Delta t, t_0 + (j+1)\Delta t + \bar{T})$ starting at the time instant $t = t_0 + (j+1)\Delta t$ in the position $x_j^*(t_0 + (j+1)\Delta t)$. For $j = 0$, $x_{j-1}^*(t_0 + j\Delta t) = x_0$.

3.1.2.3 Characteristic Function in Truncated Subgame

For each coalition $S \subseteq N$ and $j = 0, \dots, l$ we define the values of the characteristic function as in [185] or [186]:

$$V_j(S; x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) = \begin{cases} \sum_{i \in N} K_i^j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; u_j^*), & S = N, \\ \tilde{V}_j(S, x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}), & S \subset N, \\ 0, & S = \emptyset, \end{cases} \quad (3.10)$$

where $\tilde{V}_j(S, x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ is defined as the total payoff for players from coalition S in the Nash equilibrium $u_j^{NE} = (u_1^{NE,j}, \dots, u_n^{NE,j})$ in the game with the following set of players: coalition S (acting as one player) and the players from the set $N \setminus S$, i.e., in the game with $|N \setminus S| + 1$ players.

An imputation $\xi_j(x_{j,0}^*, t_0 + j\Delta t)$ in each truncated cooperative subgame $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ is defined as an arbitrary vector that satisfies the conditions

$$\begin{aligned} \xi_i^j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) &\geq V_j(\{i\}, x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}), \quad i \in N, \\ \sum_{i \in N} \xi_i^j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) &= V_j(N, x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}). \end{aligned}$$

We denote by $E_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ the set of all possible imputations in each truncated subgame. Suppose that for each truncated subgame a non-empty solution is defined:

$$W_j(x_{j,0}^*, t, t_0 + j\Delta t + \bar{T}) \subset E_j(x_{j,0}^*, t, t_0 + j\Delta t + \bar{T}) \quad (3.11)$$

This can be the Core, NM solution, Nucleus, or Shapley value.

3.1.3 Concept of Cooperative Solution with Dynamic Updating

It is logical to assume that the distribution of the total payoff between players in the game $\Gamma(x_0, T - t_0)$ along the conditionally cooperative trajectory $\{\hat{x}^*(t)\}_{t=t_0}^T$ is defined as a combination of the imputations on time intervals $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, $j = 0, \dots, l$. This construction is a new solution concept called the resulting solution.

The combination of the family of sets $W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ does not allow obtaining a solution in the game $\Gamma(x_0, T - t_0)$ directly. For each $j = 0, \dots, l$ the solution in a truncated subgame $\hat{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ is defined on the time interval $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$. But information about the game is updated with the step Δt , and the use of such a solution on the time interval $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$ is not possible. A necessary part of the solution can be obtained by using the imputation distribution procedure for each truncated subgame. The IDP also provides the time-consistency property of the new solution concept and the ability to determine solutions within the time interval $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$.

Definition 3.1.1 *A solution $W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ ($\xi_j(x_{j,0}^*, t_0 + j\Delta t)$) is called time-consistent if for any imputation $\xi_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) \in W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ there exists an IDP $\beta_j(t, x_j^*)$ that satisfies the following conditions $\forall t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$:*

$$\left\{ \int_t^{t_0 + j\Delta t + \bar{T}} \beta_i^j(\tau, x_j^*) d\tau \right\} \in W(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \quad (3.12)$$

$$\left(\left\{ \int_t^{t_0 + j\Delta t + \bar{T}} \beta_i^j(\tau, x_j^*) d\tau \right\} = \xi_j(x_{j,0}^*, t_0 + j\Delta t) \right).$$

To construct a solution in the game $\Gamma(x_0, T - t_0)$, we need to define the IDP for all truncated subgames $\hat{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, $j = 0, \dots, l$. The family of subgames along the cooperative trajectory $x_j^*(t)$ is denoted by $\hat{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, where $t \in (t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ is the initial instant of the subgame. The characteristic function along $x_j^*(t)$ in the family of subgames $\hat{\Gamma}_j^c(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ is also defined as in (3.10). We denote by $E_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ the set of imputations in the subgame $\hat{\Gamma}_j^c(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$.

Suppose that in each truncated subgame $\hat{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, a solution $W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) \neq \emptyset$ along the cooperative trajectory $x_j^*(t)$ is selected. Also, suppose that for any truncated subgame $\hat{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, in the starting position $x_{j,0}^*$ an imputation

$$\xi_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) \in W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$$

and the corresponding IDP

$$\beta_j(t, x_j^*) = [\beta_1^j(t, x_j^*), \dots, \beta_n^j(t, x_j^*)], \quad t \in (t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$$

are selected so that the IDP guarantees the time-consistency of this imputation: [187]:

$$\xi_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) = \int_{t_0 + j\Delta t}^{t_0 + j\Delta t + \bar{T}} \beta_j(t, x_j^*) e^{-r(\tau - t_0)} dt. \quad (3.13)$$

The IDP $\beta_j(t, x_j^*)$ can be obtained by differentiating the imputation $\xi_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$; see the corresponding theorem below, which was presented in [188].

Theorem 3.1.2 *If the function $\xi_j(x_j^*, t, t_0 + j\Delta t + \bar{T})$ is continuously differentiable with respect to t and x_j^* , then*

$$\begin{aligned} \beta_j(t, x_j^*) = & -\xi_t^j(x_j^*, t, t_0 + j\Delta t + \bar{T}) - \\ & - \xi_{x_j^*}^j(x_j^*, t, t_0 + j\Delta t + \bar{T}) g \left(\tau, x_j^*(\tau), u_1^{*j}(\tau, x_j^*), \dots, u_n^{*j}(\tau, x_j^*) \right). \end{aligned} \quad (3.14)$$

The new solution concept in the game $\Gamma(x_0, T - t_0)$ consists of a combination of solutions $W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ (corresponding to the IDP) in truncated subgames $\hat{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, $j = 0, \dots, l$. Suppose that for each imputation $\xi_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) \in W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, there exists $\beta_j(t, x_j^*)$. We define the resulting IDP for the whole game $\Gamma(x_0, T - t_0)$ in the following way.

Definition 3.1.3 *The resulting IDP $\hat{\beta}(t, \hat{x}^*)$ is defined for each set $\xi_j(x_{j,0}^*, t_0 +$*

$j\Delta t, t_0 + j\Delta t + \bar{T}) \in W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ using the corresponding $\beta_j(t, x_j^*)$ as follows:

$$\hat{\beta}(t, \hat{x}^*) = \begin{cases} \beta_0(t, x_0^*), & t \in [t_0, t_0\Delta t], \\ \vdots \\ \beta_j(t, x_j^*), & t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t], \\ \vdots \\ \beta_l(t, x_l^*), & t \in [t_0 + l\Delta t, t_0 + (l+1)\Delta t]. \end{cases} \quad (3.15)$$

Using the resulting IDP $\hat{\beta}(t, \hat{x}^*)$, we construct the following vector.

Definition 3.1.4 *The resulting imputation $\hat{\xi}(\hat{x}^*(t), T - t)$ is the vector defined by the resulting IDP $\hat{\beta}(t, \hat{x}^*)$ in the following way, where $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$:*

$$\begin{aligned} \hat{\xi}(\hat{x}^*(t), T - t) &= \int_t^T \hat{\beta}(\tau, \hat{x}^*(\tau)) e^{-r(\tau-t_0)} d\tau = \int_t^{t_0+j\Delta t} \beta_j(\tau, x_j^*(\tau)) e^{-r(\tau-t_0)} d\tau \\ &+ \sum_{m=j+1}^l \left[\int_{t_0+m\Delta t}^{t_0+(m+1)\Delta t} \beta_m(\tau, x_m^*(\tau)) e^{-r(\tau-t_0)} d\tau \right]; \end{aligned} \quad (3.16)$$

in particular,

$$\hat{\xi}(x_0, T - t_0) = \int_{t_0}^T \hat{\beta}(\tau, \hat{x}^*(\tau)) e^{-r(\tau-t_0)} d\tau.$$

We introduce the concept of the resulting solution in the game $\Gamma(x_0, T - t_0)$ with dynamic updating.

Definition 3.1.5 *The resulting solution $\hat{W}(\hat{x}^*(t), T - t)$ is the set of the resulting imputations $\hat{\xi}(\hat{x}^*(t), T - t)$, (3.16) for all possible resulting IDPs $\hat{\beta}(t, \hat{x}^*)$ (3.15).*

As proved in [77], the resulting imputation $\hat{\xi}(\hat{x}^*(t), T - t)$ and the corresponding resulting solution $\hat{W}(\hat{x}^*(t), T - t)$ can be used to allocate the actual total payoffs among the players:

Assertion 3.1.1.1 Any resulting imputation $\hat{\xi}(x_0, T - t_0) \in \hat{W}(x_0, T - t_0)$ and the corresponding resulting IDP $\hat{\beta}(t, \hat{x}^*(t))$ allocates the total player payoffs (3.5) along the conditionally cooperative trajectory $\hat{x}^*(t)$ in the game with prescribed duration $\Gamma(x_0, T - t_0)$:

$$\sum_{i=1}^n \int_{t_0}^t \hat{\beta}_i(\tau, \hat{x}^*(\tau)) e^{-r(\tau-t_0)} d\tau = \sum_{i=1}^n \int_{t_0}^t h_i(\tau, \hat{x}^*(\tau), \hat{u}^*(\tau)) e^{-r(\tau-t_0)} d\tau. \quad (3.17)$$

The resulting solution $\hat{W}(x_0, T - t_0)$ is time-consistent by construction. As proved in [76], it also has the property of strong time-consistency.

Definition 3.1.6 A solution $W(x_0, T - t_0)$ is called strongly Δt -time-consistent if for any $j = 0, \dots, l$ and every $\xi(x_0, T - t_0) \in W(x_0, T - t_0)$, the corresponding IDP $\beta(t, x^*)$ satisfies the condition

$$\int_{t_0}^{t_0+j\Delta t} \beta(\tau, x^*(\tau)) e^{-r(\tau-t_0)} d\tau \oplus W(x_{j,0}^*, T - t_0 + j\Delta t) \subset W(x_0, T - t_0), \quad (3.18)$$

where $a \oplus A = \{a + a' : a' \in A\}$.

Theorem 3.1.3 An arbitrary resulting solution $\hat{W}(x_0, T - t_0)$ is strongly Δt -time-consistent in the game $\Gamma(x_0, T - t_0)$ with prescribed duration.

Under an arbitrary resulting solution, any resulting solution is constructed using solutions $W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, $j = 0, \dots, l$ as an arbitrary subset of the set of imputations $E_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$. The solutions in each truncated subgame can differ, i.e., in the first truncated subgame the players can choose the Core, in the second truncated subgame they can choose the Shapley value, etc.

3.1.4 Construction of Characteristic Function in Game with Dynamic Updating

As the characteristic function in the differential game $\Gamma(x_0, T - t_0)$ with dynamic updating, the resulting characteristic function is used.

Definition 3.1.7 *The resulting characteristic function $\bar{V}(S; \hat{x}^*(t), T - t)$ in the game $\Gamma(\hat{x}^*(t), T - t)$ with dynamic updating is the function calculated using the characteristic function values $V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ in every truncated subgame $\hat{\Gamma}_j^c(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ along the conditionally cooperative trajectory $\hat{x}^*(t)$ for $j = 0, \dots, l, \forall t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$.*

$$\begin{aligned} \bar{V}(S; \hat{x}^*(t), T - t) &= \sum_{m=j+1}^l \left[V_m(S; x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \right. \\ &\quad \left. - V_m(S; x_{m,1}^*, t_0 + (m + 1)\Delta t, t_0 + m\Delta t + \bar{T}) \right] + \\ &\quad + \left[V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(S; x_{j,1}^*, t_0 + (j + 1)\Delta t, t_0 + j\Delta t + \bar{T}) \right], \end{aligned} \quad (3.19)$$

where $x_{j,0}^* = \hat{x}^*(t_0 + j\Delta t)$, $x_{j,1}^* = \hat{x}^*(t_0 + (j + 1)\Delta t)$.

The following theorem shows that the resulting imputation $\hat{\xi}(x_0, T - t_0)$ can be used as an imputation in the game $\Gamma(x_0, T - t_0)$ with the characteristic function $\bar{V}(S; x_0, T - t_0)$.

Theorem 3.1.4 *The resulting imputation $\hat{\xi}(x_0, T - t_0)$ is an imputation in the game $\Gamma(x_0, T - t_0)$ with dynamic updating if $\forall t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, $j = 0, \dots, l$ the following condition is satisfied:*

$$\begin{aligned} &\xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \geq \\ &\geq \xi_i^j(x_{j,1}^*, t_0 + (j + 1)\Delta t, t_0 + j\Delta t + \bar{T}) - V_j(\{i\}; x_{j,1}^*, t_0 + (j + 1)\Delta t, t_0 + j\Delta t + \bar{T}). \end{aligned} \quad (3.20)$$

Proof First, let us show that $\forall t \in [t_0, T]$ the following conditions are fulfilled:

$$\sum_{i=1}^n \hat{\xi}_i(\hat{x}^*(t), T - t) = \bar{V}(N; \hat{x}^*(t), T - t), \quad (3.21)$$

$$\hat{\xi}_i(\hat{x}^*(t), t) \geq \bar{V}(\{i\}; \hat{x}^*(t), T - t). \quad (3.22)$$

According to the definition of $\hat{\xi}(\hat{x}^*(t), T - t)$ and $\bar{V}(S; \hat{x}^*(t), T - t)$, the left-hand

side of (3.21) can be rewritten as

$$\begin{aligned}
\sum_{i=1}^n \hat{\xi}_i(\hat{x}^*(t), T-t) &= \sum_{i=1}^n \int_t^T \hat{\beta}(\tau, \hat{x}^*(\tau)) d\tau = \\
&= \sum_{i=1}^n \left[\sum_{m=j+1}^l \left[\int_{m\Delta t}^{(m+1)\Delta t} \beta_m(\tau, x_m^*(\tau)) d\tau \right] + \left[\int_t^{j\Delta t} \beta_j(\tau, x_m^*(\tau)) d\tau \right] \right] = \\
&= \sum_{i=1}^n \left[\sum_{m=j+1}^l [\xi_i^m(x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \right. \\
&\quad \left. - \xi_i^m(x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T})] + \right. \\
&\quad \left. + [\xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - \xi_i^j(x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T})] \right]. \quad (3.23)
\end{aligned}$$

Since condition (3.24) in (3.23) is satisfied, (3.21) is correct.

$$\sum_{i=1}^n \xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = V_j(N; x_j^*(t), t, t_0 + j\Delta t + \bar{T}), \quad j = 0, \dots, l, \quad (3.24)$$

Now let us prove (3.22) by substituting the expression for $\hat{\xi}_i(\hat{x}^*(t), T-t)$ and $\bar{V}(\{i\}; \hat{x}^*(t), T-t)$ into the left-hand side of (3.22). Also, we substitute the

characteristic function $\bar{V}(\{i\}; x_0, T - t_0)$ (3.19) into the right-hand side of (3.22):

$$\begin{aligned}
& \sum_{m=j+1}^l \left[\xi_i^m(x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \right. \\
& \quad \left. - \xi_i^m(x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \right] + \\
& + \left[\xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - \xi_i^j(x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) \right] \geq \\
& \geq \sum_{m=j+1}^l \left[V_m(\{i\}; x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \right. \\
& \quad \left. - V_m(\{i\}; x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \right] + \\
& + \left[V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(\{i\}; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) \right] \\
& \tag{3.25}
\end{aligned}$$

(3.25) is fulfilled $\forall t \in [t_0, T]$ if for $\forall m = 0, \dots, l$ we have

$$\begin{aligned}
& \xi_i^m(x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \xi_i^m(x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \geq \\
& V_m(\{i\}; x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - V_m(\{i\}; x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \\
& \tag{3.26}
\end{aligned}$$

and $\forall t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $m = 0, \dots, l$ we have

$$\begin{aligned}
& \xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - \xi_i^j(x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) \geq \\
& V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(\{i\}; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}). \tag{3.27}
\end{aligned}$$

The fulfillment of condition (3.27) $\forall t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $m = 0, \dots, l$ implies the fulfillment of condition (3.26). We write (3.27) as follows:

$$\begin{aligned}
& \xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \geq \\
& \geq \xi_i^j(x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) - V_j(\{i\}; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}). \\
& \tag{3.28}
\end{aligned}$$

Condition (3.28) means that in every truncated subgame, changing the values of the

characteristic function and imputation according to the time happens evenly with respect to each other. The theorem is proved. \square

In this subsection, the notion of the characteristic function $\bar{V}(S; x_0, T - t_0)$ in the game $\Gamma(x_0, T - t_0)$ with dynamic updating has been introduced. It has been shown that the resulting imputation $\hat{\xi}(x_0, T - t_0)$ is an imputation in the classical meaning that satisfies individually rationality. Nevertheless, condition (3.20) is not fulfilled $\forall t \in [t_0, T]$.

3.1.5 Relationship of Solutions in Truncated Subgames and Resulting Solutions

In this subsection, it has been shown that if the players choose the imputation $\xi_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \in E_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ based on $V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, $j = 0, \dots, l$, in every truncated subgame by the same rule, then the resulting imputation $\hat{\xi}(\hat{x}^*(t), T - t)$ corresponds to the imputation chosen by the same rule using the resulting characteristic function $\bar{V}(S; \hat{x}^*(t), T - t)$. Below, this fact will be established for several optimality principles.

First of all, we show that if in every truncated subgame $\hat{\Gamma}_j(x_j^*, t, t_0 + j\Delta t + \bar{T})$, the players choose the Shapley value $Sh_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ as an imputation, then the resulting imputation $\hat{\xi}(\hat{x}^*(t), T - t)$ (3.16) will coincide with the Shapley value $\hat{S}h(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ calculated using the resulting characteristic function $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19).

Theorem 3.1.5 *Suppose that in every truncated subgame $\hat{\Gamma}_j(x_j^*, t, t_0 + j\Delta t + \bar{T})$,*

$$\xi_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = Sh_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}),$$

where $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$. Then the resulting imputation $\hat{\xi}(\hat{x}^(t), T - t)$ coincides with $\hat{S}h(\hat{x}^*(t), T - t)$:*

$$\hat{\xi}(\hat{x}^*(t), T - t) = \hat{S}h(\hat{x}^*(t), T - t), \quad \forall t \in [t, T],$$

where $\hat{S}h(\hat{x}^(t), T - t)$ is the Shapley value calculated using the resulting characteristic function $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19).*

Proof In this case, the resulting imputation $\hat{\xi}(\hat{x}^*(t), T-t)$ is calculated by formulas (3.15), (3.16) using the Shapley values $Sh_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ in every truncated subgame $\hat{\Gamma}_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$:

$$\begin{aligned} \hat{\xi}(\hat{x}^*(t), T-t) = & \sum_{m=j+1}^l \left[Sh_m(x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \right. \\ & \left. - Sh_m(x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \right] + \\ & + [Sh_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - Sh_j(x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T})], \quad (3.29) \end{aligned}$$

where $Sh_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$, is calculated by the formula

$$\begin{aligned} Sh_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = & \sum_{\substack{S \subset N \\ i \in S}} \frac{(|N| - |S|)! (|S| - 1)!}{|N|!} \\ & \cdot \left(V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(S \setminus \{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \right). \quad (3.30) \end{aligned}$$

$\hat{Sh}(\hat{x}^*(t), T-t)$ is calculated using the resulting characteristic function $\bar{V}(S; \hat{x}^*(t), T-t)$. We substitute the expression for $\bar{V}(S; \hat{x}^*(t), T-t)$ into the formula for $\hat{Sh}(\hat{x}^*(t),$

$T-t)$. Letting $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, we obtain:

$$\begin{aligned}
\hat{S}h(\hat{x}^*(t), T-t) &= \sum_{\substack{S \subset N \\ i \in S}} \frac{(|N| - |S|)! (|S| - 1)!}{|N|!} \\
&\cdot \left(\sum_{m=j+1}^l \left[\left[V_m(S; x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \right. \right. \right. \\
&\quad \left. \left. \left. - V_m(S; x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \right] - \right. \right. \\
&\quad \left. \left. \left. - \left[V_m(S \setminus \{i\}; x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \right. \right. \right. \\
&\quad \left. \left. \left. - V_m(S \setminus \{i\}; x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \right] \right] + \right. \\
&\quad \left. \left[\left[V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - \right. \right. \right. \\
&\quad \left. \left. \left. - V_j(S; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) \right] - \left[V_j(S \setminus \{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - \right. \right. \right. \\
&\quad \left. \left. \left. - V_j(S \setminus \{i\}; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) \right] \right] \right). \quad (3.31)
\end{aligned}$$

Substituting (3.30) into (3.29) yields formula (3.31). The theorem is proved. \square

The same result can be obtained for the proportional solution. Suppose that the characteristic function $V(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ is differentiable along the cooperative trajectory $x_j^*(t)$. We define the proportional solution using its IDP $\beta_j^{Prop}(t)$ in the following way:

$$\begin{aligned}
\beta_i^{Prop,j}(t) &= \frac{-\frac{d}{dt} V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T})}{\sum_{i \in N} -\frac{d}{dt} V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T})} \\
&\quad \left(-\frac{d}{dt} V_j(N; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \right), \quad i \in N. \quad (3.32)
\end{aligned}$$

The corresponding imputation is obtained by the direct integration of $\beta_j^{Prop}(t)$ using the formula (3.13) and denoted by $Prop_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$.

We prove that if in every truncated subgame $\hat{\Gamma}_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, the players choose the proportional solution $Prop_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ (3.32), then the resulting imputation $\hat{\xi}(\hat{x}^*(t), T-t)$ (3.16) will coincide with the proportional solution

$\hat{P}rop(\hat{x}^*(t), T-t)$ (3.32) calculated using the characteristic function $\bar{V}(S; \hat{x}^*(t), T-t)$ (3.19).

Theorem 3.1.6 *Suppose that in every truncated subgame $\hat{\Gamma}_j(x_j^*, t, t_0 + j\Delta t + \bar{T})$,*

$$\xi_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = Prop_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}),$$

where $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$. Then the resulting vector $\hat{\xi}(\hat{x}^*(t), T-t)$ coincides with $\hat{P}rop(\hat{x}^*(t), T-t)$ (3.32):

$$\hat{\xi}(\hat{x}^*(t), T-t) = \hat{P}rop(\hat{x}^*(t), T-t), \quad \forall t \in [t, T],$$

where $\hat{P}rop(\hat{x}^*(t), T-t)$ is the proportional solution calculated using the resulting characteristic function $\bar{V}(S; \hat{x}^*(t), T-t)$ (3.19).

Proof In this case, the resulting imputation $\hat{\xi}(\hat{x}^*(t), T-t)$ is calculated by formula (3.15) for the IDP using the combination of the IDP values for the proportional solution (3.32) in every truncated subgame on the interval $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$.

We show that the formula for $\hat{P}rop(\hat{x}^*(t), T-t)$ or for the IDP $\hat{\beta}^{Prop}(t)$ (3.32), where $\bar{V}(S; \hat{x}^*(t), T-t)$ (3.19) is used as a characteristic function, leads to the right-hand side of (3.32). We substitute the expression for $\bar{V}(S; \hat{x}^*(t), T-t)$ (3.19) into (3.32). Consider one of the terms. Letting $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, we have

$$\begin{aligned} -\frac{d}{dt}(\bar{V}(\{i\}; \hat{x}^*(t), T-t)) &= -\frac{d}{dt} \left(\sum_{k=j+1}^l \left[V_k(\{i\}; x_{k,0}^*, t_0 + k\Delta t, t_0 + k\Delta t + \bar{T}) - \right. \right. \\ &\quad \left. \left. - V_k(\{i\}; x_{k,0}^*, t_0 + (k+1)\Delta t, t_0 + k\Delta t + \bar{T}) \right] + \right. \\ &\quad \left. + \left[V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(\{i\}; x_{j,0}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) \right] \right). \end{aligned} \quad (3.33)$$

For $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, l$, there is only one term depending on t under the derivative sign (see (3.33)). Therefore,

$$-\frac{d}{dt}(\bar{V}(\{i\}; \hat{x}^*(t), T-t)) = -\frac{d}{dt}(V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T})). \quad (3.34)$$

We substitute (3.34) and the formula for $\bar{V}(N; \hat{x}^*(t), T - t)$ (3.32) into this expression. In this case, the right-hand sides of the IDP $\hat{\beta}^{Prop}(t)$ and (3.32) are equal to each other. The theorem is proved. \square

We show that if in every truncated subgame $\hat{\Gamma}_j(x_j^*, t, t_0 + j\Delta t + \bar{T})$, the players choose the Core $C_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ as an optimality principle, then every element of the resulting solution $\hat{\xi}(\hat{x}^*(t), T - t)$ (3.16) is the core calculated using the resulting characteristic function $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19).

Theorem 3.1.7 *Suppose that in every truncated subgame $\hat{\Gamma}_j(x_j^*, t, t_0 + j\Delta t + \bar{T})$,*

$$W_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = C_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}),$$

where $\forall t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$. Then for every $\xi_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \in C_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ that satisfies the condition

$$\begin{aligned} & \sum_{i \in S} \xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \geq \\ & \geq \sum_{i \in S} \xi_i^j(x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) - V_j(S; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}), \end{aligned} \quad (3.35)$$

we have

$$\hat{\xi}(\hat{x}^*(t), T - t) \in \hat{C}(\hat{x}^*(t), T - t), \quad \forall t \in [t, T],$$

where $\hat{C}(\hat{x}^*(t), T - t)$ is the Core calculated using the resulting characteristic function $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19).

Proof The following statements should be proven:

- 1 If the players in every truncated subgame choose the imputation $\xi_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \in C_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ calculated using $V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, $j = 0, \dots, l$, then the resulting imputation $\hat{\xi}(\hat{x}^*(t), T - t)$ will belong to the Core $\hat{C}(\hat{x}^*(t), T - t)$ calculated using the resulting characteristic function $\bar{V}(S; \hat{x}^*(t), T - t)$.

2 The Core $\hat{C}(\hat{x}^*(t), T - t)$ shouldn't contain any imputation $\hat{\xi}(\hat{x}^*(t), T - t)$ for which it is impossible to find a set of imputations in truncated subgames $\xi_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \in C_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$.

We prove the first statement: if the set of imputations $\xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ satisfies the system of inequalities

$$\sum_{i \in S} \xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \geq V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}), \quad S \subset N,$$

for each $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$, $i = 1, \dots, n$, then the resulting imputation $\hat{\xi}(\hat{x}^*(t), T - t)$ will satisfy the system of inequalities

$$\sum_{i \in S} \hat{\xi}_i(\hat{x}^*(t), T - t) \geq \bar{V}(S; \hat{x}^*(t), T - t), \quad \forall t \in [t_0, T], \quad S \subset N. \quad (3.36)$$

We substitute the expression for $\hat{\xi}(\hat{x}^*(t), T - t)$ into the left-hand side of (3.36) and (3.19) into the right-hand side of (3.36). Like in Theorem 3.1.6, we show that for every $S \subset N$, $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, the fulfilment of (3.36) leads to the fulfilment of (3.35). The first statement is proved.

Now we prove the second statement: in the set of imputations $\xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, $i = 1, \dots, n$, $j = 0, \dots, l$, satisfying the following system of inequalities $\forall j = 0, \dots, l$, $S \subset N$:

$$\begin{aligned}
& \sum_{m=j+1}^l \left[\xi_i^m(x_m^*(t), t, t_0 + m\Delta t + \bar{T}) - \right. \\
& \quad \left. - \xi_i^m(x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \right] + \\
& + \left[\xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - \xi_i^j(x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) \right] \geq \\
& \geq \sum_{m=j+1}^l \left[V_m(S; x_{m,0}^*, t_0 + m\Delta t, t_0 + m\Delta t + \bar{T}) - \right. \\
& \quad \left. - V_m(S; x_{m,1}^*, t_0 + (m+1)\Delta t, t_0 + m\Delta t + \bar{T}) \right] + \\
& + \left[V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(S; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) \right], \\
& \tag{3.37}
\end{aligned}$$

there exists at least one set of $\xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, $i = 1, \dots, n$, $j = 0, \dots, l$ satisfying

$$\begin{aligned}
& \sum_{i \in S} \xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \geq \\
& \geq V_j(S; x_j^*, t, t_0 + j\Delta t + \bar{T}), \quad \forall j = 0, \dots, l, \quad S \subset N. \tag{3.38}
\end{aligned}$$

This fact will be proved by contradiction. Suppose that for the imputations satisfying (3.35) and (3.38), inequality (3.37) is not satisfied. We show that for $\forall j = 0, \dots, l$ the following condition is satisfied:

$$\begin{aligned}
& \sum_{i \in S} \xi_i^j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \geq \\
& \geq \sum_{i \in S} \xi_i^j(x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}) - V_j(S; x_{j,1}^*, t_0 + (j+1)\Delta t, t_0 + j\Delta t + \bar{T}). \\
& \tag{3.39}
\end{aligned}$$

Due to (3.38), the right- and left-hand sides are always positive, and by (3.35) it follows that (3.39) is satisfied. The theorem is proved. \square

The same results can be obtained for the IDP-core, a new cooperative solution

presented in [100]. We introduce the following notation:

$$U_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = -\frac{d}{dt}V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}), \quad (3.40)$$

where $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ and $S \subseteq N$.

We define $B_j(t)$ as a set of integrable vector functions $\beta_j(t)$ satisfying the following system of inequalities:

$$B_j(t, x_j^*) = \left\{ \beta_j(t) = (\beta_1^j(t), \dots, \beta_n^j(t)) : \right. \\ \left. \sum_{i \in S} \beta_i^j(t) \geq U_j(S, x_j^*(t), t, t_0 + j\Delta t + \bar{T}), \right. \\ \left. \sum_{i \in N} \beta_i^j = U_j(N, x_j^*(t), t, t_0 + j\Delta t + \bar{T}), \forall S \subset N \right\}. \quad (3.41)$$

Suppose that $B_j(t, x_j^*) \neq \emptyset$, $\forall t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$. Then the set $B_j(t, x_j^*)$ can be used to define the following set of vectors.

Definition 3.1.8 *The set of all $\xi_j(x_j^*(t), t)$ for some integrable selectors $\beta_j(t, x_j^*) \in B_j(t, x_j^*)$ will be called the IDP-core and denoted by $\bar{C}_j(x_j^*(t), t_0 + j\Delta t + \bar{T})$, where*

$$\bar{C}_j(x_j^*(t), t) = \{ \xi_j(x_j^*(t), t), t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}] \} \quad (3.42)$$

and for $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$,

$$\xi_j(x_j^*(t), t) = \int_t^T \beta_j(\tau, x_j^*) d\tau. \quad (3.43)$$

As was proved in [100], the IDP-core is strongly time-consistent.

We show that if in every truncated subgame $\hat{\Gamma}_j(x_j^*, t, t_0 + j\Delta t + \bar{T})$, the players choose the IDP-core $\bar{C}_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ (3.42) as an optimality principle, then every element of the resulting solution (3.16) is the IDP-core calculated using the resulting characteristic function $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19).

Theorem 3.1.8 *Suppose that in every truncated subgame $\hat{\Gamma}_j(x_j^*, t, t_0 + j\Delta t + \bar{T})$,*

$$W_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = \bar{C}_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \neq \emptyset,$$

where $\forall t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$. Then

$$\hat{W}(\hat{x}^*(t), T - t) = \hat{C}(\hat{x}^*(t), T - t), \quad \forall t \in [t, T],$$

where $\hat{C}(\hat{x}^*(t), T - t)$ is the IDP-core calculated using the resulting characteristic function $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19).

Proof The resulting solution $\hat{W}(\hat{x}^*(t), T - t)$ consists of $\hat{\xi}(\hat{x}^*(t), T - t)$, each of which is defined by the IDP set of imputations $\xi_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \in \bar{C}_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, $j = 0, \dots, l$ by formula (3.15). By definition, each imputation from the IDP-core satisfies the following system of inequalities (3.41):

$$\begin{aligned} \sum_{i \in S} \beta_i^j(t, \hat{x}^*(t)) &\geq -\frac{d}{dt} \left(V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \right), \\ \sum_{i \in N} \beta_i^j(t, \hat{x}^*(t)) &= -\frac{d}{dt} \left(V_j(N; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) \right), \quad \forall S \subset N. \end{aligned}$$

Thus, the resulting solution $\hat{W}(\hat{x}^*(t), T - t)$ is defined by (3.41) for $t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, $j = 0, \dots, l$.

We write the expression for $\hat{C}(\hat{x}^*(t), T - t)$ with the resulting characteristic function $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19) for showing that it yields (3.41). We consider separately one of the constraints in it and substitute the expression for $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19). Letting $t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, we obtain:

$$\begin{aligned} -\frac{d}{dt} \left(\bar{V}(S; \hat{x}^*(t), T - t) \right) &= -\frac{d}{dt} \left(\sum_{k=j+1}^l \left[V_k(S; x_{k,0}^*, t_0 + k\Delta t, t_0 + k\Delta t + \bar{T}) - \right. \right. \\ &\quad \left. \left. - V_k(S; x_{k,0}^*, t_0 + (k + 1)\Delta t, t_0 + k\Delta t + \bar{T}) \right] + \right. \\ &\quad \left. + \left[V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) - V_j(S; x_{j,0}^*, t_0 + (j + 1)\Delta t, t_0 + j\Delta t + \bar{T}) \right] \right). \end{aligned} \tag{3.44}$$

For $t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, $j = 0, \dots, l$, there is only one term depending on t under the desired sign (see (3.44)). Hence,

$$-\frac{d}{dt}\left(\bar{V}(S; \hat{x}^*(t), T - t)\right) = -\frac{d}{dt}\left(V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T})\right). \quad (3.45)$$

We substitute (3.45) into the formula for $\hat{C}(\hat{x}^*(t), T - t)$:

$$\begin{aligned} \sum_{i \in S} \hat{\beta}_i(t, \hat{x}^*(t)) &\geq -\frac{d}{dt}\left(V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T})\right), \\ \sum_{i \in N} \hat{\beta}_i(t, \hat{x}^*(t)) &= -\frac{d}{dt}\left(V_j(N; \hat{x}^*(t), t, t_0 + j\Delta t + \bar{T})\right), \forall S \subset N. \end{aligned}$$

Thus, the IDP-core $\hat{C}(\hat{x}^*(t), T - t)$ with the resulting characteristic function $\bar{V}(S; \hat{x}^*(t), T - t)$ coincides with the resulting solution $\hat{W}(\hat{x}^*(t), T - t)$ calculated using the combination of solutions $\bar{C}_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ in the truncated subgames. The theorem is proved. \square

In this subsection, it has been shown that if in every truncated subgame $\hat{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, the players choose the proportional solution, Shapley value, an imputation from the Core, or an imputation from the IDP-core as an optimality principle, then the resulting imputation is also the proportional solution, Shapley value, an imputation from the Core, or an imputation from the IDP-core in the game $\Gamma(x_0, T - t_0)$ with dynamic updating.

The theorems proved below give a direct approach for calculating the resulting solution.

3.1.6 Cooperative Limited Resource Extraction Game with Dynamic Updating

Consider the resource extraction game defined on a closed time interval. The solution of the two-person game in the classical form was presented in [69]. The problem of time-consistency was studied by Yeung [188]. In this example, a limited resource extraction game with dynamic updating for three persons is presented. The Core is used as an optimality principle. The characteristic function is calculated like in

[185]. In the last part of the example, the property of strong time-consistency is discussed.

3.1.6.1 Original Game

The following dynamical system describes the change in the stock of resources $x(t) \in X \subset R$:

$$\dot{x} = a\sqrt{x(t)} - bx(t) - \sum_{i=1}^3 u_i, \quad x(t_0) = x_0,$$

where u_i is the production level of player $i = \overline{1, 3}$. The payoff of player i has the form

$$K_i(x_0, t_0; u) = \int_{t_0}^T h_i(\tau, x(\tau), u(\tau)) d\tau,$$

where

$$h_i(\tau, x(\tau), u(\tau)) = \sqrt{u_i(\tau)} - \frac{c_i}{\sqrt{x(\tau)}} u_i(\tau), \quad i = \overline{1, 3},$$

c_i is a constant, $c_i \neq c_k, \forall i \neq k = \overline{1, 3}$.

3.1.6.2 Truncated Subgame

The original game $\Gamma(x_0, T - t_0)$ is defined on the time interval $[t_0, T]$. Suppose that for any $t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, $j = 0, \dots, l$, the players have truncated information about the game. It includes information about the dynamical system and the payoff function on the time interval $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$. This model is constructed using the truncated subgame $\hat{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$. The dynamical system and the initial data have the following form:

$$\dot{x} = a\sqrt{x(t)} - bx(t) - \sum_{i=1}^3 u_i, \quad x(t_0 + j\Delta t) = x_{j,0}. \quad (3.46)$$

The payoff function of player i is

$$K_i^j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; u) = \int_{t_0 + j\Delta t}^{t_0 + j\Delta t + \bar{T}} h_i(\tau, x(\tau), u(\tau)) d\tau.$$

Consider the case when the players agree to cooperate in the truncated subgame $\hat{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$. Then the players act to maximize the total payoffs.

3.1.6.3 Cooperative Trajectory

Suppose that the maximum joint payoff of the players in each truncated subgame $\hat{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ has the following form [69]:

$$W^j(t, x) = A^j(t)\sqrt{x} + C^j(t), \quad (3.47)$$

where the functions $A^j(t)$ and $C^j(t)$ satisfy the system of differential equations

$$\begin{aligned} \dot{A}^j(t) &= \frac{b}{2}A^j(t) - \sum_{i=1}^3 \left[\frac{1}{4 \left[c_i + \frac{A^j(t)}{2} \right]} \right], \\ \dot{C}^j(t) &= -\frac{a}{2}A^j(t), \\ A^j(t_0 + j\Delta t + \bar{T}) &= 0, \quad C^j(t_0 + j\Delta t + \bar{T}) = 0. \end{aligned}$$

The cooperative trajectory $x_j^*(t)$ in each truncated subgame can be calculated as follows [69]:

$$x_j^*(t) = \varpi_j^2(t_0 + j\Delta t, t) \left[\sqrt{x_{j,0}^*} + \frac{1}{2}a \cdot \int_{t_0 + j\Delta t}^t \varpi_j(t_0 + j\Delta t, \tau)^{-1} d\tau \right]^2,$$

where $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ and

$$\varpi_j(t_0 + j\Delta t, t) = \exp \int_{t_0 + j\Delta t}^t - \left[\frac{1}{2}b + \sum_{i=1}^3 \left[\frac{1}{4 \left[c_i + \frac{A^j(\tau)}{2} \right]^2} \right] \right] d\tau.$$

The initial position for the cooperative trajectory in each truncated subgame is defined from the previous truncated subgame: $x_{0,0}^* = x_0$ and $x_{j,0}^* = x_{j-1}^*(t_0 + j\Delta t)$ for $1 \leq j \leq l$. The conditionally cooperative trajectory $\hat{x}^*(t)$ is defined as follows:

$$\hat{x}_j^*(t) = x_j^*(t), \quad t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t], \quad j = 0, \dots, l.$$

3.1.6.4 Characteristic Function

To allocate the cooperative payoff between the players in each truncated subgame, it is necessary to determine the characteristic function values $V_j(S; x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ for each coalition $S \subset N$. According to formula (3.10), the maximum total payoff for the players, $W_j(t_0 + j\Delta t, x_{j,0})$ (3.47), corresponds to the characteristic function value $V_j(N; x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ of coalition $S = N$ in the truncated subgame $\hat{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$:

$$V_j(N; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = W_j(t, x_j^*(t)), \quad (3.48)$$

where $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$. Next, the characteristic function values for the following coalitions are determined:

$$\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}.$$

For each coalition $\{i\}$, $i = \overline{1, 3}$, the Nash equilibrium is determined in the truncated subgame $\hat{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ and consequently, $V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T})$.

3.1.6.5 One Player Coalitions

The Nash equilibrium in the truncated subgame $\hat{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ is determined by the following strategies of the players:

$$u_i^j(t, x) = \frac{x}{4[c_i + A_i^j(t)/2]^2}, \quad i = \overline{1, 3},$$

where the functions $A_i^j(t)$ for $i = \overline{1, 3}$ are defined by the system of differential equations

$$\begin{aligned} \dot{A}_i^j(t) &= A_i^j(t) \left[\frac{b}{2} + \sum_{k \neq i} \frac{1}{8(c_k + A_k^j(t)/2)^2} \right] - \frac{1}{4(c_i + A_i^j(t)/2)}, \\ \dot{C}_i^j(t) &= -\frac{a}{2} A_i^j(t), \\ A_i^j(t_0 + j\Delta t + \bar{T}) &= 0, \quad C_i^j(t_0 + j\Delta t + \bar{T}) = 0. \end{aligned}$$

The corresponding payoff of players $i = \overline{1, 3}$ in the Nash equilibrium is determined by the function

$$V_i^j(t, x) = A_i^j(t)\sqrt{x} + C_i^j(t), \quad i = \overline{1, 3}.$$

The characteristic function value for the coalitions consisting of one player ($S = \{i\}$, $i \in N$) is calculated as follows:

$$V_j(\{i\}; x_j^*(t), t, t_0 + j\Delta t + \overline{T}) = V_i^j(t, x_j^*(t)), \quad (3.49)$$

where $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \overline{T}]$, $j = 0, \dots, l$.

3.1.6.6 Two Player Coalitions

According to formula (3.10), the characteristic function $V_j(S; x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \overline{T})$ for the coalitions consisting of two players ($S = \{1, 2\}, \{1, 3\}, \{2, 3\}$) is defined as the best reply of coalition S to the strategies from the Nash equilibrium $u_j^{NE} = (u_1^{NE,j}, u_2^{NE,j}, u_3^{NE,j})$ in the truncated subgame $\hat{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \overline{T})$, used by the players from N/S . In our case, this means that the players from coalition S act as one player and maximize their total payoff. Using this approach, we determine the equilibrium between the two players: the combined player (coalition S), and a player not included in coalition S (coalition N/S).

Consider the formula for $V_j(S; x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \overline{T})$ in the case $S = \{1, 2\}$. The formula for the remaining coalitions can be obtained by the same principle:

$$\begin{aligned} V_{\{1,2\}}^j(t, x) &= A_{\{1,2\}}^j(t)\sqrt{x} + C_{\{1,2\}}^j(t), \\ V_3^j(t, x) &= A_3^j(t)\sqrt{x} + C_3^j(t), \end{aligned}$$

where the functions $A_{\{1,2\}}^j(t)$, $A_3^j(t)$, $C_{\{1,2\}}^j(t)$, and $C_3^j(t)$ satisfy the system of

differential equations

$$\begin{aligned}\dot{A}_{\{1,2\}}^j(t) &= A_{\{1,2\}}^j(t) \left[\frac{b}{2} + \frac{1}{8(c_3 + A_3^j(t)/2)^2} \right] - \sum_{k \in S} \frac{1}{4(c_k + A_{\{1,2\}}^j(t)/2)}, \\ \dot{A}_3^j(t) &= A_3^j(t) \left[\frac{b}{2} + \sum_{k \in S} \frac{1}{8(c_k + A_{\{1,2\}}^j(t)/2)^2} \right] - \frac{1}{4(c_3 + A_3^j(t)/2)}, \\ \dot{C}_{\{1,2\}}^j(t) &= -\frac{a}{2}A_{\{1,2\}}^j(t), \quad \dot{C}_3^j(t) = -\frac{a}{2}A_3^j(t)\end{aligned}$$

with the initial condition $A_{\{1,2\}}^j(t_0 + j\Delta t + \bar{T}) = 0$, $A_3^j(t_0 + j\Delta t + \bar{T}) = 0$, $C_{\{1,2\}}^j(t_0 + j\Delta t + \bar{T}) = 0$, $C_3^j(t_0 + j\Delta t + \bar{T}) = 0$.

The characteristic function value for coalition $S = \{1, 2\}$ is calculated as follows:

$$V_j(\{1, 2\}; x_j^*(t), t, t_0 + j\Delta t + \bar{T}) = V_{\{1,2\}}^j(t, x_j^*(t)), \quad (3.50)$$

where $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$.

3.1.6.7 Concept of Solution

Suppose that in each cooperative truncated subgame $\hat{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, the players use the Core as an optimality principle. This means that the players in each truncated subgame choose an imputation $\xi_j(x_j^*, t, t_0 + j\Delta t + \bar{T}) \in C_j(x_j^*(t), T - t)$ by the following rule:

$$\sum_{i \in S} \xi_i^j(x_j^*, t, t_0 + j\Delta t + \bar{T}) \geq V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}), \quad S \subset N,$$

for any $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$. The resulting imputation $\hat{\xi}(\hat{x}^*(t), T - t)$ for any set of distributions in the truncated subgames $\xi_j(x_j^*, t, t_0 + j\Delta t + \bar{T}) \in C_j(x_j^*(t), T - t)$, $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, l$, can be calculated by formula (3.16). We denote by $\hat{C}(\hat{x}^*(t), T - t)$ the set of imputations $\hat{\xi}(\hat{x}^*(t), T - t)$ constructed using (3.15) and (3.16).

Using the results obtained in Sections 4 and 5, the solution $\hat{C}(\hat{x}^*(t), T - t)$ can

be constructed by the following rule:

$$\sum_{i \in S} \hat{\xi}(\hat{x}^*(t), T - t) \geq \bar{V}(S; \hat{x}^*(t), T - t), \quad S \subset N, \quad (3.51)$$

where $\bar{V}(S; \hat{x}^*(t), T - t)$ is calculated using formula (3.19).

Further, using an example of a particular deviation from $\hat{C}(\hat{x}^*(t), T - t)$, we show that the solution constructed is strongly Δt -time-consistent in the game $\Gamma(x_0, T - t_0)$.

3.1.6.8 Numerical Simulation

Consider a numerical example of the resource extraction game defined on the time interval $T - t_0 = 4$, in which information about the game is known on the time interval with the duration $\bar{T} = 2$ and is updated with the step $\Delta t = 1$. Let us choose the following parameters: for the motion equation, $a = 5$ and $b = 0.3$; for the payoff function, $c_1 = 0.15$, $c_2 = 0.65$, and $c_3 = 0.45$; for the initial conditions, $t_0 = 0$ and $x_0 = 250$.

Figure 3.1 shows the optimal strategies for the first player in the game with dynamic updating (solid line) and the optimal strategies in the original game [69] (dotted line) with prescribed duration.

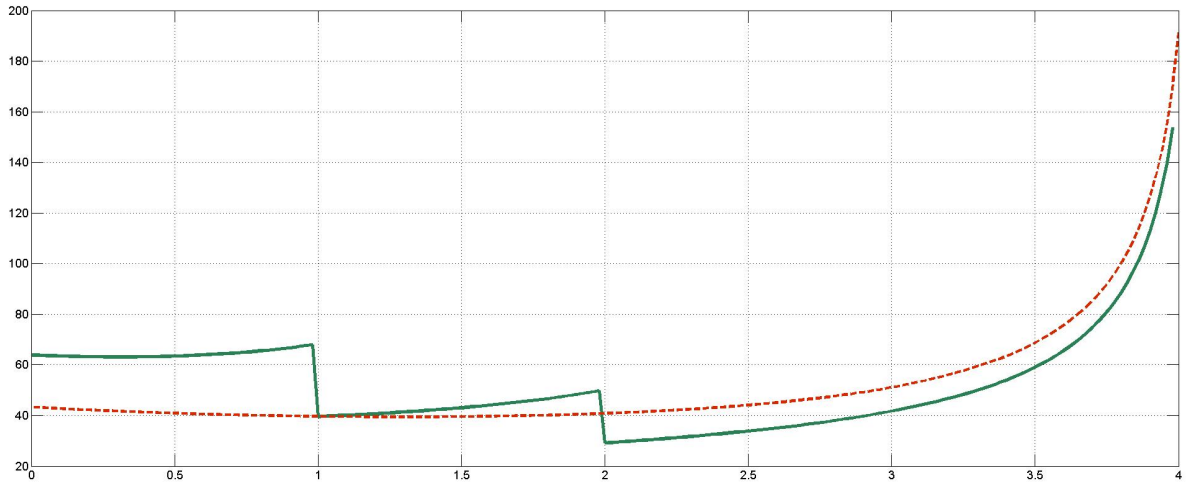


Figure 3.1: Optimal strategy for first player in game with dynamic updating (solid line) and optimal strategies in original game [69] (dotted line) with prescribed duration.

The conditionally cooperative trajectory $\hat{x}^*(t)$ is constructed using the cooper-

ative trajectories in the truncated subgames $\hat{\Gamma}_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ with the dynamical system (3.46). In Figure 3.2 the comparison of the conditionally cooperative trajectory $\hat{x}^*(t)$ (solid line) in the game with dynamic updating and the cooperative trajectory $x^*(t)$ (dashed line) in the original game $\Gamma(x_0, T - t_0)$ [69] is displayed. Under limited information, resource extraction occurs faster because the players are guided by a reduced time interval. The abscissa axis in Figure 3.2 corresponds to the time t , whereas the ordinate axis to the stock of resource x .

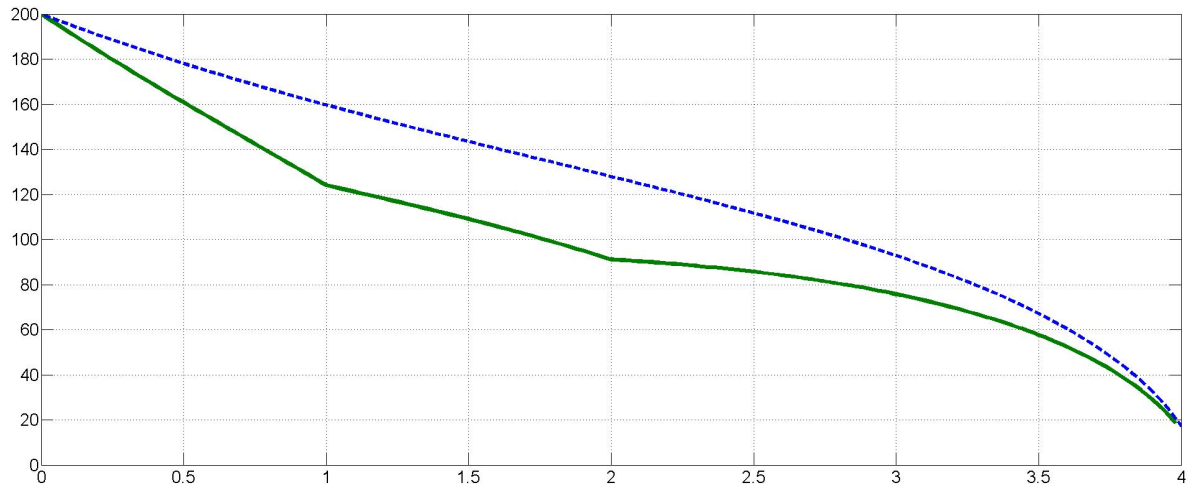


Figure 3.2: Conditionally cooperative trajectory $\hat{x}^*(t)$ (solid line) in game with dynamic updating and cooperative trajectory $x^*(t)$ (dashed line) in original game [69].

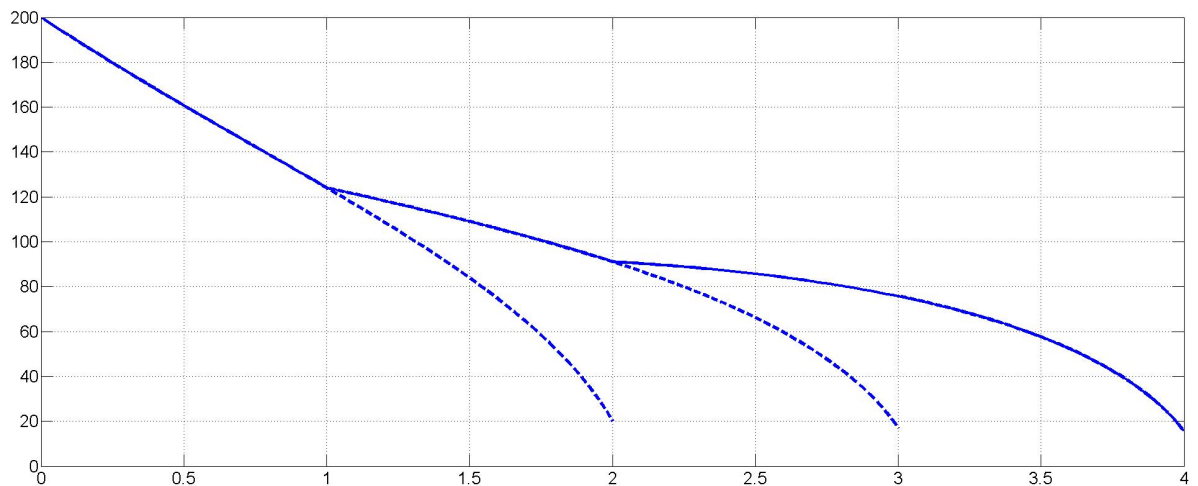


Figure 3.3: Conditionally cooperative trajectory $\hat{x}^*(t)$ (solid line) in game with dynamic updating and corresponding cooperative trajectories in truncated subgames (dashed lines).

Based on the characteristic function values

$$V_j(S; x_j^*(t), t, t_0 + j\Delta t + \bar{T}), \quad t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t], \quad S \subset N, \quad i = 0, \dots, l,$$

calculated using (3.48), (3.49), (3.50), we obtain the expression for the resulting characteristic function $\bar{V}(S; \hat{x}^*(t), T - t)$ (3.19), $t \in [t_0, T]$. Using (3.51), we construct $\hat{C}(\hat{x}^*(t), T - t)$ in the game with dynamic updating $\Gamma(x_0, T - t_0)$ (see Fig. 3.5).

Now we demonstrate the strong Δt -time-consistency of the solution $\hat{C}(\hat{x}^*(t), T - t)$. Suppose that at the beginning of the game $\Gamma(x_0, T - t_0)$, the players agreed to use the proportional solution $\hat{P}rop(\hat{x}^*(t), T - t)$ (3.32). (Below, it will be shown that $\hat{P}rop(\hat{x}^*(t), T - t) \in \hat{C}(\hat{x}^*(t), T - t)$ for the given parameters.) Suppose that at some instant $t_{br} = t_0 + m\Delta t \in [t_0, T]$, the players decided to choose another imputation from $\hat{C}(\hat{x}^*(t_{br}), T - t_{br})$ instead of the proportional solution, e.g., the Shapley value $\hat{S}h(\hat{x}^*(t), T - t)$, $t \in [t_{br}, T]$ (3.30). The IDP for the proportional solution and the Shapley value are calculated using formula (3.14).

Let $m = 2$. Then the resulting IDP (3.15) for the combined solution has the following form:

$$\hat{\beta}(t, \hat{x}^*) = \begin{cases} \hat{\beta}^{Prop}(t, \hat{x}^*), & t \in [t_0, t_{br}], \\ \hat{\beta}^{Sh}(t, \hat{x}^*), & t \in (t_{br}, T]. \end{cases} \quad (3.52)$$

Figure 3.4 shows the resulting IDP for the proportional solution $\hat{\beta}^{Prop}(t, \hat{x}^*)$ (solid line) and $\hat{\beta}(t, \hat{x}^*)$ for the combined solution (3.52) (dashed line).

To obtain the imputation (3.16) corresponding to the combined solution $\hat{\beta}(t, \hat{x}^*)$ (3.52), we integrate it with respect to t . We denote the result of the integration by $\hat{\xi}(\hat{x}^*(t), T - t)$. According to the resulting imputation $\hat{\xi}(\hat{x}^*(t), T - t)$, the players allocate their joint payoff in the game $\Gamma(x_0, T - t_0)$ with dynamic updating in the following way:

$$\hat{\xi}(\hat{x}^*(t), T - t) = (12.3, 30.2, 16.8).$$

In Figure 3.5, the imputation corresponding to the combined solution $\hat{\xi}(\hat{x}^*(t), T - t)$ (dashed line) belongs to $\hat{C}(\hat{x}^*(t), T - t)$ (the selected domain) for all $t \in [t_0, T]$. This illustrates the strong Δt -time-consistency of $\hat{C}(\hat{x}^*(t), T - t)$ since the imputation $\hat{\xi}(\hat{x}^*(t), T - t)$ has been constructed by the deviation of players from the proportional

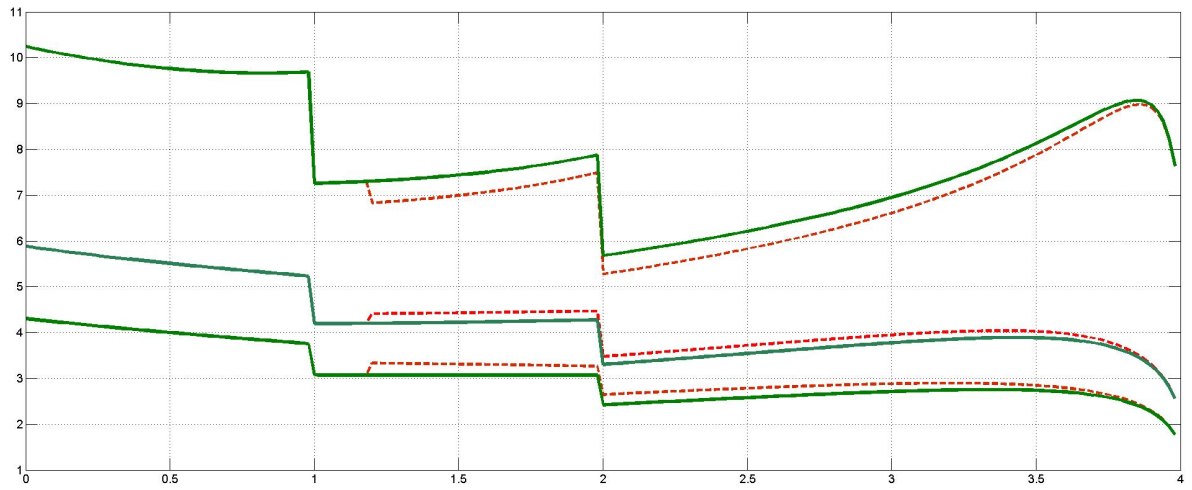


Figure 3.4: IDP $\hat{\beta}^{Prop}(t, \hat{x}^*)$ for proportional solution (solid line) and IDP $\hat{\beta}(t, \hat{x}^*)$ for combined solution (3.52) (dashed line).

solution $\hat{P}rop(\hat{x}^*(t), T - t)$ (solid line) at the instant $t_{br} = t_0 + j\Delta t$ in favor of the Shapley value $\hat{S}h(\hat{x}^*(t_{br}), T - t_{br})$.

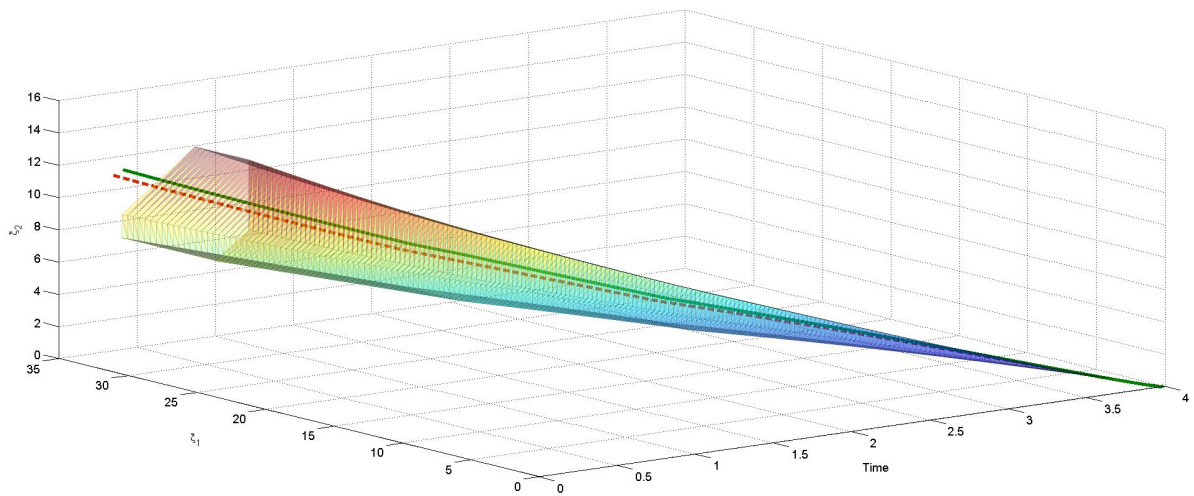


Figure 3.5: Axis: ξ_1, ξ_3, t . ξ_2 can be calculated using normalization condition.

3.2 Cooperative Differential Games with Dynamic Updating and Uncertain Stochastic Dynamics

3.2.1 Original Game Model and Assumptions

The n -person differential game $\Gamma(x_0, 0, T)$ on a time interval $[0, T]$ starting from an initial state $x_0 \in R^n$ is defined by the motion equations

$$\dot{x} = g(t, x, u), \quad x(0) = x_0, \quad (3.53)$$

where $x \in R^m$, $u = (u_1 \dots u_n)$. We denote by $N = \{1, 2, \dots, n\}$ the set of players. Player i chooses a control $u_i \in U_i \subset \text{Comp}R^k$ for $i = 0, \dots, n$.

The payoff function of player i is defined as

$$K_i(x_0, 0, T; u) := \int_0^T h_i(x(\tau), u(\tau)) d\tau + q_i(x(T)), \quad (3.54)$$

where $x(\tau)$ is the trajectory (solution) of the system (3.53) on the time interval $[0, T]$ with the control input u , and $q_i(x(T))$ is the terminal payoff.

Suppose that the function $g(x, u)$ has the following form:

$$g(t, x, u) = \begin{cases} g_0(t, x, u), & t \in [0, \Delta t] \\ g_1(t, x, u), & t \in]\Delta t, 2\Delta t] \\ \dots \\ g_j(t, x, u), & t \in]j\Delta t, (j+1)\Delta t] \\ \dots \\ g_{l-1}(t, x, u), & t \in](l-1)\Delta t, l\Delta t]. \end{cases}$$

We need to emphasize the part of the game dynamics $g_j(t, x, u)$ on the time interval $]j\Delta t, (j+1)\Delta t]$, $j = 0, \dots, l-1$, to formulate the behavior of the players on this time interval.

Suppose that the information for players is being updated at fixed time instants

$t = j\Delta t, j = 0, \dots, l-1$. During the time interval $]j\Delta t, (j+1)\Delta t]$, the players have certain information about the game dynamics described by the function $g_j(t, x, u)$. However, they do not know the game dynamics on the time interval $](j+1)\Delta t, T]$. Any forecast on this interval is based on the present information about the process, which is the game dynamics function $g_j(t, x, u)$ in the current model. The only option for the players is to act as if the game structure will not change and use the function $g_j(t, x, u)$ to forecast the game dynamics on the time interval $[(j+1)\Delta t, T]$.

To apply the dynamic updating approach, the players must have an opportunity to use combined truncated information about the game structure. The combined truncated information at any time instant $t \in [j\Delta t, (j+1)\Delta t]$ includes the motion equation on this interval,

$$\dot{x} = g_j(t, x, u), \quad (3.55)$$

and a forecast of the motion on the future time interval $[(j+1)\Delta t, T]$. The forecast is defined by the stochastic equation with the same deterministic part:

$$dx := g_j(t, x, u)dt + \sigma(t, x)dz(t), \quad (3.56)$$

where $\sigma(t, x)$ is an $m \times \theta$ matrix, and $z(t)$ is a θ -dimensional Wiener process. When information about the game dynamics is updated at the time instants $j\Delta t$, the players recalculate their decisions using the new updated information.

Problems of this type often occur in real-life situations since the structure of the conflict process on large time intervals is not certain, and a forecast should be used.

3.2.2 Combined Truncated Subgame

During the first time interval $[0, \Delta t]$, the players have certain information about the game structure on $[0, \Delta t]$ and a forecast for the time interval $[\Delta t, T]$. At the instant Δt , the information about the game is being updated; on the second time interval $]\Delta t, 2\Delta t]$, the players have certain information about the game structure on the time interval $]\Delta t, 2\Delta t]$ and a forecast for the time interval $[2\Delta t, T]$. To incorporate this feature in the model, the following definition is introduced.

Definition 3.2.1 *Let $0 \leq j \leq l-1$. A combined truncated subgame $\bar{\Gamma}_j(x, j\Delta, T)$ is defined on the time interval $[j\Delta t, T]$ in the following way. The motion equation and*

payoff function on the time interval $[j\Delta t, (j+1)\Delta t]$ coincide with those of the game $\Gamma(x_0, 0, T)$ on the same time interval. However, on the time interval $[(j+1)\Delta t, T]$, the subgame $\bar{\Gamma}_j(x, j\Delta t, T)$ is a stochastic differential game. The motion equation and initial condition of the combined truncated subgame $\bar{\Gamma}_j(x, j\Delta t, T)$ have the following form:

$$dx := \begin{cases} g_j(t, x, u)dt, & t \in [j\Delta t, (j+1)\Delta t], \\ g_j(t, x, u)dt + \sigma(t, x)dz(t), & t \in](j+1)\Delta t, T], \end{cases} \quad x(j\Delta t) = x. \quad (3.57)$$

The payoff function of player i is

$$K_i^j(x, j\Delta t, T; u) := \int_{j\Delta t}^{(j+1)\Delta t} h_i(x(\tau), u(\tau))d\tau + E \left\{ \int_{(j+1)\Delta t}^T h_i(x(\tau), u(\tau))d\tau + q_i(x(T)) \right\}. \quad (3.58)$$

Clearly, the combined truncated subgame $\bar{\Gamma}_j(x, j\Delta t, T)$ for $j = 0, \dots, l-1$ can be defined as a stochastic differential game with the following motion equations and initial conditions:

$$dx = g_j(t, x, u)dt + I(j, t) \cdot \sigma(t, x)dz(t), \quad x(j\Delta t) = x, \quad (3.59)$$

where

$$I(j, t) = \begin{cases} 0, & t \in [j\Delta t, (j+1)\Delta t], \\ 1, & t \in](j+1)\Delta t, T], \end{cases} \quad (3.60)$$

and the payoff function of player i is

$$K_i^j(x, j\Delta t, T; u) = E \left\{ \int_{j\Delta t}^T h_i(x(\tau), u(\tau))d\tau + q_i(x(T)) \right\}. \quad (3.61)$$

3.2.2.1 Solution of Combined Truncated Cooperative Subgame

Consider the combined truncated cooperative subgame $\bar{\Gamma}_j^c(x, j\Delta t, T)$ on the time interval $[j\Delta t, T]$ with the initial condition $x(j\Delta t) = x$. The expected profit of all players to be maximized in this game is

$$\sum_{i \in N} K_i^j(x, j\Delta t, T; u^j) = E \left\{ \sum_{i \in N} \int_{j\Delta t}^T h_i(x(\tau), u(\tau)) d\tau + q_i(x(T)) \right\} \quad (3.62)$$

subject to

$$\begin{aligned} dx &= g_j(t, x, u)dt + I(j, t) \cdot \sigma(t, x)dz(t), \\ x(j\Delta t) &= x. \end{aligned} \quad (3.63)$$

This is a stochastic control problem. Sufficient conditions for the solution and the optimal feedback are given by the following assertion, representing a special case of [189, Theorem A.5].

Theorem 3.2.1 *Assume that there exists a twice continuously differentiable function*

$W^{(j\Delta t)}(t, x) : [j\Delta t, T] \times R^m \rightarrow IR$ *satisfying the partial differential equation*

$$\begin{aligned} -W_t^{(j\Delta t)}(t, x) - \frac{1}{2} \sum_{h, \zeta=1}^m I(j, t) \sigma^{h, \cdot}(t, x(t)) \sigma^{\zeta, \cdot}(t, x(t))^T W_{x^h, x^\zeta}^{(j\Delta t)}(t, x) = \\ = \max_u \left\{ \sum_{i=1}^n h_i(t, x, u) + W_x^{(j\Delta t)}(t, x) g_j(t, x, u) \right\} \end{aligned} \quad (3.64)$$

with the boundary condition

$$W^{(j\Delta t)}(T, x) = \sum_{i=1}^n q_i(x).$$

*Let the maximum in (3.64) be achieved at $u = \psi^{*j}(t, x)$. Then the feedback $u(t, x) = \psi^{*j}(t, x)$ is optimal in the game $\bar{\Gamma}_j^c(x, j\Delta t, T)$ if the closed loop stochastic system has a solution.*

3.2.2.2 Conditionally Cooperative Trajectory

Within the dynamic approach, the combined truncated information about the game structure is only available to the players in the game $\Gamma(x, 0, T)$. This information is not sufficient to design an optimal cooperative control. Instead of the optimal cooperative trajectory in the game $\Gamma(x, 0, T)$, a recursive conditionally optimal cooperative trajectory $\{\hat{x}^*(t)\}_{t=0}^T$ is constructed in the form

$$\{\hat{x}^*(t)\}_{t=0}^T := \begin{cases} x_0^*(t) & t \in [0, \Delta t] \\ \dots \\ x_j^*(t) & t \in]j\Delta t, (j+1)\Delta t] \\ \dots \\ x_{l-1}^*(t) & t \in](l-1)\Delta t, l\Delta t]. \end{cases} \quad (3.65)$$

On the time interval $[0, \Delta t]$, the trajectory $x_0^*(t)$ is optimal in the combined truncated cooperative subgame $\bar{\Gamma}_0^c(x_0, 0, T)$. At the time instant Δt , information about the game structure is updated, and the position of the players is $x_0^*(\Delta t)$. On the time interval $] \Delta t, 2\Delta t]$, the function $x_1^*(t)$ is part of the optimal cooperative trajectory in the combined truncated cooperative subgame $\bar{\Gamma}_1^c(x_0^*(\Delta t), \Delta t, T)$ starting at the instant Δt in the position $x_0^*(\Delta t)$. At the instant $j\Delta t$, information about the game structure is updated, and the system position is $x_{j-1}^*(j\Delta t)$. The function $\hat{x}^*(t)$ on the time interval $]j\Delta t, (j+1)\Delta t]$ is defined as part of the optimal cooperative trajectory $x_j^*(t)$ in the combined truncated cooperative subgame $\bar{\Gamma}_j^c(x_{j-1}^*(j\Delta t), j\Delta t, T)$ starting at the instant $j\Delta t$ in the position $x_{j-1}^*(j\Delta t)$. Therefore, the conditionally cooperative trajectory $\{\hat{x}^*(t)\}_{t=0}^T$ is defined as a composition of the cooperative trajectories $x_j^*(t)$ in the combined truncated cooperative subgames $\bar{\Gamma}_j^c(x_{j-1}^*(j\Delta t), j\Delta t, T)$ defined on the successive time intervals $]j\Delta t, (j+1)\Delta t]$. We denote by $x_{j,0}^* := x_{j-1}^*(j\Delta t) = x_j^*(j\Delta t)$ the boundary vector. Then the combined truncated cooperative subgame is $\bar{\Gamma}_j^c(x_{j,0}^*, j\Delta t, T)$.

3.2.2.3 Characteristic Function for Combined Truncated Subgame

For each coalition $S \subset N$ and $j = 0, \dots, l-1$, we define the characteristic function values as it was done in [190]:

$$V_j(S, x_{j,0}^*, j\Delta t, T) := \begin{cases} \sum_{i=1}^n K_i^j(x_{j,0}^*, j\Delta t, T; u_j^*) & S = N \\ \tilde{V}_j(S) & S \subset N \\ 0 & S = \emptyset, \end{cases} \quad (3.66)$$

where $\tilde{V}_j(S)$ is the value of the noncooperative combined truncated subgame $\bar{\Gamma}_j(x_{j,0}^*, j\Delta t, T)$ defined as the maximized sum of the payoffs of all players in coalition S with respect to the Nash equilibrium strategies for the left-out players. An imputation $\xi^j(x_{j,0}^*, j\Delta t, T)$ in each combined truncated cooperative subgame $\bar{\Gamma}_j^c(x_{j,0}^*, j\Delta t, T)$ is defined as an arbitrary vector function that satisfies the conditions

$$\begin{aligned} \xi_i^j(x_{j,0}^*, j\Delta t, T) &\geq V_j(\{i\}, x_{j,0}^*, j\Delta t, T), \quad i \in N \\ \sum_{i \in N} \xi_i^j(x_{j,0}^*, j\Delta t, T) &= V_j(N, x_{j,0}^*, j\Delta t, T). \end{aligned} \quad (3.67)$$

We denote by $E_j(x_{j,0}^*, j\Delta t, T)$ the set of all possible imputations in the combined truncated subgame. Suppose that a nonempty and optimal solution $W_j(x_{j,0}^*, j\Delta t, T) \subset E_j(x_{j,0}^*, j\Delta t, T)$ is chosen for each combined truncated subgame $\bar{\Gamma}_j^c(x_{j,0}^*, j\Delta t, T)$. It can be the Core, NM-solution, nucleus, or Shapley value. It is easy to suggest that the distribution of the total payoff of players in the game $\Gamma(x_0, 0, T)$ along the conditionally cooperative trajectory $\{\hat{x}^*(t)\}_{t=0}^T$ is organized as a composition of imputations on each time interval $[j\Delta t, (j+1)\Delta t]$, $j = 0, \dots, l-1$, in accordance with the structure of the game $\Gamma(x_0, 0, T)$. This will be formalized below as a new solution concept. The family of sets $W_j(x_{j,0}^*, j\Delta t, T)$ do not compose directly a solution of the game $\Gamma(x_0, 0, T)$. For any $j = 0, \dots, l-1$, the optimal solution of the truncated subgame $\bar{\Gamma}_j^c(x_{j,0}^*, j\Delta t, T)$ is defined on the time interval $[j\Delta t, T]$. This particular solution makes sense on the interval $[j\Delta t, (j+1)\Delta t]$ only because the information about the game structure is updated after every Δt time interval, and it is irrelevant to use a solution based upon the outdated information. The necessary information can be extracted using the imputation distribution

procedure (IDP) [191] for each truncated subgame. Also, the IDP provides the time-consistency property of the new solution concept and an opportunity to define a cooperative solution at any instant on $]j\Delta t, T]$.

3.2.3 Concept of Combined Cooperative Solution

To construct a solution concept for the game $\Gamma(x_0, 0, T)$, an IDP needs to be introduced for all combined truncated cooperative subgames $\bar{\Gamma}_j^c(x_{j,0}^*, j\Delta t, T)$, $j = 0, \dots, l-1$.

Let $0 \leq j \leq l-1$. We define the family of subgames of the game $\bar{\Gamma}_j^c(x_{j,0}^*, j\Delta t, T)$ along its cooperative trajectory $x_j^*(t)$ as $\bar{\Gamma}_j^c(x_j^*(t), t, T)$, where $t \in]j\Delta t, T]$ is the starting instant of the subgame.

The characteristic function along $x_j^*(t)$ in the family of subgames of $\bar{\Gamma}_j^c(x_j^*(t), t, T)$ on the time interval $]t, T]$ is defined as in (4.142). We denote by $E_j(x_j^*, t, T)$ the set of imputations in $\bar{\Gamma}_j^c(x_j^*(t), t, T)$ along x_j^* .

Suppose that for any subgame of the previously defined combined truncated cooperative subgame $\bar{\Gamma}_j^c(x_j^*(t), t, T)$, an optimal solution $W_j(x_j^*, t, T) \neq \emptyset$ along the cooperative trajectory x_j^* is chosen.

Suppose that for any combined truncated subgame $\bar{\Gamma}_j^c(x_{j,0}^*, j\Delta t, T)$, in the starting positions $x_{j,0}^*$ the players agreed to choose the imputations

$$\xi^j(x_{j,0}^*, j\Delta t, T) \in W_j(x_{j,0}^*, j\Delta t, T) \quad (3.68)$$

and the corresponding IDP

$$B_j(t, x_j^*) = [B_1^j(t, x_j^*) \dots B_n^j(t, x_j^*)], \quad (3.69)$$

where $t \in]j\Delta t, T]$, which guarantees the time-consistency of this imputation [192]:

$$\xi^j(x_{j,0}^*, j\Delta t, T) = E \left\{ \int_{j\Delta t}^T B_j(t, x_j^*) dt + q_i(x^*(T)) \right\}. \quad (3.70)$$

The IDP $B_j(t, x_j^*)$ can be obtained from $\xi_t^j(x_j^*, t, T)$ by direct differentiation [189, Theorem 8.1].

Theorem 3.2.2 *If the imputations $\xi^j(x_j^*, t, T)$ are twice continuously differentiable with respect to t and x_j^* , then*

$$\begin{aligned} B_j(\tau, x_j^*) = & - \left[\xi_t^j(x_j^*, t, T) \Big|_{t=\tau} \right] \\ & - \left[\xi_{x_j^*}^j(x_j^*, t, T) \Big|_{t=\tau} \right] g_j \left[\tau, x_j^*, \psi_1^{*j}(\tau, x) \dots \psi_n^{*j}(\tau, x) \right] - \\ & - \frac{1}{2} \sum_{h, \zeta=1}^m \Omega^{h, \zeta}(j, \tau, x_j^*) \left[\xi_{x^h, x^\zeta}^{(j\Delta t)}(x_j^*, t, T) \Big|_{t=\tau} \right], \end{aligned} \quad (3.71)$$

where $\Omega(j, t, x) = I(j, t)\sigma(t, x)\sigma(t, x)^T$ is the covariance matrix, and $I(j, t)$ is the indicator function defined in (3.60).

A new solution concept for the combined differential game $\Gamma(x_0, 0, T)$ consists of the recursive merging of the family of time-consistent solutions $W_j(x_{j,0}^*, j\Delta t, T)$ of the combined truncated subgames $\bar{\Gamma}_c(x_{j,0}^*(j\Delta t), j\Delta t, T)$ for $j = 0, \dots, l-1$, obtained by the dynamic updating approach. For any imputation $\xi_j(x_{j,0}^*, j\Delta t, T) \in W_j(x_{j,0}^*, j\Delta t, T)$, there exists an IDP $B_j(t, x_j^*)$. We define the combined truncated distribution

$$\hat{B}(t, \hat{x}^*) := B_j(t, x_j^*), \quad t \in [j\Delta t, (j+1)\Delta t], \quad 0 \leq j \leq l-1. \quad (3.72)$$

The function $\hat{B}(t, \hat{x}^*)$ determines the combined vector

$$\hat{\xi}^j(x_{j,0}, j\Delta t, T) := \int_{j\Delta t}^T \hat{B}(\tau, \hat{x}^*(\tau)) d\tau + q(\hat{x}^*(T)) \quad (3.73)$$

$$= \sum_{m=j}^{l-1} \left[\int_{m\Delta t}^{(m+1)\Delta t} B_m(\tau, x_m^*(\tau)) d\tau \right] + q(x_l^*(T)) \quad (3.74)$$

for $j = 0, \dots, l-1$. We denote by $\hat{W}_j(x_{j,0}^*, j\Delta t, T)$ the set of all vectors $\hat{\xi}^j(x_{j,0}, j\Delta t, T)$ composed by (3.72), (3.73).

Within this concept, the new solution of the game $\Gamma(x_0, 0, T)$ is defined by $\hat{W} := (\hat{W}_j(x_{j,0}^*, j\Delta t, T))_{j=0}^{l-1}$.

The construction of $\hat{B}(t, \hat{x}^*)$ does not require the function $B_j(t, x_j^*)$ to be defined on the entire time interval where the chosen imputation $\xi^j(x_{j,0}^*, j\Delta t, T)$ is defined: we need it on the time interval $]j\Delta t, (j+1)\Delta t]$ only. The last term in (3.71) equals to 0 on this interval:

$$-\frac{1}{2} \sum_{h,\zeta=1}^m \Omega^{h,\zeta}(j, t, x_j^*) \left[\xi_{x^h, x^\zeta}^{(j\Delta t)}(t, x_j^*) \right] = 0, \quad t \in]j\Delta t, (j+1)\Delta t],$$

because $\Omega(j, t, x) = I(j, t)\sigma(t, x)\sigma(t, x)^T$ and the indicator function $I(j, t) = 0$ for $t \in]j\Delta t, (j+1)\Delta t]$. Therefore, the formula for calculating $\hat{B}(t, \hat{x}^*)$ is valid if the imputation $\xi^j(x_j^*, t, T)$ is just once continuously differentiable with respect to t and x_j^* .

It is easy to see that the solution \hat{W} is time-consistent. However, there is another surprising property of \hat{W} .

Definition 3.2.2 *A solution $W = (W_j(x_{j,0}^*, j\Delta t, T))_{j=0}^{l-1}$ is called strong Δt -time-consistent iff for any $j = 0, \dots, l-1$ and any $\xi \in W_0(x_0, 0, T)$, the IDP $B(t, x^*)$ generated by ξ satisfies the inclusion*

$$\int_0^{j\Delta t} B(\tau, x^*(\tau)) d\tau \oplus W_j(x_{j,0}^*, j\Delta t, T) \subset W_0(x_0, 0, T), \quad (3.75)$$

where $a \oplus A = \{a + a' : a' \in A\}$.

Theorem 3.2.3 *The solution \hat{W} is strong Δt -time-consistent in the game $\Gamma(x_0, 0, T)$.*

Proof Let $0 \leq j \leq l-1$, and assume that an imputation $\hat{\xi}^0(x_0, 0, T) \in \hat{W}_0(x_0, 0, T)$ generates an IDP $\hat{B}(t, \hat{x}^*)$. Then for $0 \leq k < j$, there exists $\xi^k(x_{k,0}^*, k\Delta t, T) \in W_k(x_{k,0}^*, k\Delta t, T)$ with IDPs $B_k(t, x_k^*)$ such that

$$\hat{B}(t, \hat{x}^*) = B_k(t, x_k^*), \quad t \in [k\Delta t, (k+1)\Delta t[, \quad 0 \leq k \leq j-1.$$

Hence,

$$\int_0^{j\Delta t} \hat{B}(\tau, \hat{x}^*(\tau)) d\tau = \sum_{k=0}^{j-1} \int_{k\Delta t}^{(k+1)\Delta t} B_k(t, x_k^*(t)) dt.$$

Assume that $\xi'' \in W_j(x_{j,0}^*, j\Delta t, T)$. Then for $j \leq k \leq l-1$, there exists $\xi^k(x_{k,0}^*, k\Delta t, T) \in W_k(x_{k,0}^*, k\Delta t, T)$ with IDPs $B_k(t, x_k^*)$ such that $\hat{B}(t, \hat{x}^*) = B_k(t, x_k^*)$ for $t \in [k\Delta t, (k+1)\Delta t[$ and

$$\xi'' = \sum_{m=j}^{l-1} \left[\int_{m\Delta t}^{(m+1)\Delta t} B_m(\tau, x_m^*(\tau)) d\tau \right] + q(x_l^*(T)).$$

Therefore,

$$\int_0^{j\Delta t} \hat{B}(\tau, \hat{x}^*(\tau)) d\tau + \xi'' = \sum_{m=0}^{l-1} \left[\int_{m\Delta t}^{(m+1)\Delta t} B_m(\tau, x_m^*(\tau)) d\tau \right] + q(x_l^*(T)) \in \in \hat{W}_0(x_0, 0, T),$$

which completes the proof. □

3.2.4 Cooperative Differential Resource Extraction Game with Dynamic Updating and Uncertain Stochastic Dynamics

The following example of the resource extraction game was considered by D. Yeung and S. Jorgensen ([193]). The problem of time-consistency in this example was studied by D. Yeung ([189]). The dynamic updating approach is applied to the example with the new factor in the solution—the indicator function $I(j, t)$. The resulting subgames form the combined solution.

Consider the differential resource extracting game with two asymmetric extractors (players). The motion equation for the resource stock $x(t) \in X \subset R$ has the

following form:

$$\begin{aligned}\dot{x} &:= a\sqrt{x(t)} - bx(t) - u_1 - u_2 \\ x(0) &:= x_0,\end{aligned}\tag{3.76}$$

where u_i is the harvest rate of player $i = 1, 2$. Suppose that the game dynamics can change, and the parameters of the motion equations switch in the successive time intervals: $a = a_j$ and $b = b_j$ if $t \in [j\Delta t, (j+1)\Delta t)$, for $j = 0, \dots, l-1$, where $T = l\Delta t$.

The payoff function of player i is defined as

$$K_i(x, 0, T; u_1, u_2) := \int_0^T h_i(x(\tau), u(\tau))d\tau + q_i\sqrt{x(T)}, \quad i = 1, 2,\tag{3.77}$$

with

$$h_i(x(\tau), u(\tau)) := \sqrt{u_i(\tau)} - \frac{c_i}{\sqrt{x(\tau)}}u_i(\tau), \quad i = 1, 2,\tag{3.78}$$

where c_1, c_2 are constants such that $c_1 \neq c_2$.

The basic game $\Gamma(x_0, 0, T)$ is defined on the time interval $[0, T]$. Suppose that for any $t \in [j\Delta t, (j+1)\Delta t]$, $j = 0, \dots, l-1$, the players have combined truncated information about the game structure. It includes certain information about the motion equations on the time interval $[j\Delta t, (j+1)\Delta t]$ and a forecast about the motion on the time interval $[(j+1)\Delta t, T]$. The combined truncated information is formalized in the combined truncated subgame $\bar{\Gamma}_j(x, j\Delta t, T)$. The motion equations and initial conditions for this subgame have the following form:

$$\begin{aligned}dx &:= \left[a_j\sqrt{x(t)} - b_jx(t) - u_1 - u_2 \right] dt + I(j, t) \cdot \sigma x(t)dz(t), \\ x(j\Delta t) &:= x,\end{aligned}\tag{3.79}$$

where

$$I(j, t) = \begin{cases} 0, & t \in [j\Delta t, (j+1)\Delta t], \\ 1, & t \in [(j+1)\Delta t, T]. \end{cases}\tag{3.80}$$

The payoff function of player i in the stochastic game $\bar{\Gamma}_j(x, j\Delta t, T)$ is

$$K_i^j(x, j\Delta t, T; u_1, u_2) := E \left\{ \int_{j\Delta t}^T h_i(x(\tau), u(\tau)) d\tau + q_i \sqrt{x(T)} \right\}. \quad (3.81)$$

The resource extraction game was studied in detail in [189; 193]. The combined truncated subgame $\bar{\Gamma}_j(x, j\Delta t, T)$ has a Nash equilibrium defined by the feedback

$$u_i^j(t, x) := \frac{x}{4[c_i + A_i^j(t)/2]^2}, \quad i = 1, 2,$$

where the functions $A_i^j(t)$ satisfy the equations

$$\begin{aligned} \dot{A}_i^j(t) &:= A_i^j(t) \left[\frac{b_j}{2} + \frac{1}{8}\sigma^2 + \frac{1}{8(c_j + A_{3-i}^j(t)/2)^2} \right] - \frac{1}{4(c_i + A_i^j(t)/2)}, \\ \dot{C}_i^j(t) &:= -\frac{a_j}{2} A_i^j(t) \end{aligned}$$

for $i = 1, 2$ with the boundary conditions $A_i^j(T) := q_i$ and $C_i^j(T) := 0$.

The value function of player $i = 1, 2$ in the Nash equilibrium is equal [193] to

$$V_i^j(t, x) := A_i^j(t)\sqrt{x} + C_i^j(t), \quad i = 1, 2. \quad (3.82)$$

Now consider the case when the resources extractors agree to act cooperatively in the combined truncated subgame $\bar{\Gamma}_j^c(x, j\Delta t, T)$. They follow the optimality principle, under which they will maximize their joint expected payoffs and share the excess of the total expected cooperative payoff.

The maximized expected joint payoff in the game $\bar{\Gamma}_j^c(x, j\Delta t, T)$ is derived in the similar way as in [189; 193]:

$$W^j(t, x) := A^j(t)\sqrt{x} + C^j(t), \quad (3.83)$$

where the functions $A^j(t)$, $C^j(t)$ satisfy the equations

$$\begin{aligned} \dot{A}^j(t) &:= \left[\frac{1}{8} I(j, t) \cdot \sigma^2 + \frac{b_j}{2} \right] A^j(t) - \frac{1}{4 \left[c_1 + \frac{A^j(t)}{2} \right]} - \frac{1}{4 \left[c_2 + \frac{A^j(t)}{2} \right]}, \\ \dot{C}^j(t) &:= -\frac{a_j}{2} A^j(t), \quad A^j(T) := \sum_{i=1}^2 q_i, \quad C^j(T) := 0. \end{aligned} \quad (3.84)$$

The optimal cooperative trajectory $x_j^*(t)$ in the stochastic game $\bar{\Gamma}_j^c(x, j\Delta t, T)$ can be represented explicitly [189] on the entire interval $[j\Delta t, T]$. The formula for the beginning part of $x_j^*(t)$ on the interval $[j\Delta t, (j+1)\Delta t]$ is simpler because the motion equation on this interval is not stochastic. The trajectory with the initial data $x = x_{j,0}^*$ is

$$x_j^*(t) = \varpi_j^2(j\Delta t, t) \left[\sqrt{x_{j,0}^*} + \frac{1}{2} a_j \int_{j\Delta t}^t \varpi_j(j\Delta t, \tau)^{-1} d\tau \right]^2, \quad t \in [j\Delta t, (j+1)\Delta t], \quad (3.85)$$

where

$$\varpi_j(j\Delta t, t) = \exp \int_{j\Delta t}^t - \left[\frac{1}{2} b_j + \frac{1}{8 \left[c_1 + \frac{A^j(\tau)}{2} \right]^2} + \frac{A^j(\tau)}{8 \left[c_2 + \frac{A^j(\tau)}{2} \right]^2} \right] d\tau. \quad (3.86)$$

The initial data are defined recursively by the optimal trajectory of the previous game: $x_{0,0}^* := x_0$ and $x_{j,0}^* := x_{j-1}^*(j\Delta t)$ for $1 \leq j \leq l-1$. Within the dynamic updating approach, the conditionally cooperative trajectory $\hat{x}^*(t)$ is defined as

$$\hat{x}_j^*(t) := x_j^*(t), \quad t \in [j\Delta t, (j+1)\Delta t], \quad 0 \leq j \leq l-1.$$

Consider a numerical example with four time intervals, $T = 4$, $\Delta t = 1$, and the following parameters of the motion equations:

$$\begin{aligned} a_0 &:= 10, & a_1 &:= 9, & a_2 &:= 12, & a_3 &:= 8, \\ b_0 &:= 0.5, & b_1 &:= 0.8, & b_2 &:= 0.5, & b_3 &:= 1.6. \end{aligned} \quad (3.87)$$

Let $c_1 := 0.05$, $c_2 := 0.1$, $q_1 := 1.5$, and $q_2 := 1$ for the payoff function and $x_0 := 1$ for the initial data.

The conditionally cooperative trajectory $\hat{x}^*(t)$ is composed from solutions of the combined truncated subgames $\bar{\Gamma}_j^c(x_{j,0}^*, j\Delta t, T)$ with the motion equation (3.79). If $\sigma = 0$ in this equation, then the forecast is deterministic and expands the current information on the entire remaining interval $[j\Delta t, T]$. If $\sigma \neq 0$, then the forecast is stochastic. The next simulation compares the deterministic and stochastic forecasts, also showing the solution without forecast. The stochastic forecast is calculated for $\sigma := 2$.

The trajectory without forecast is calculated on each interval $[j\Delta t, (j+1)\Delta t]$ as the optimal one in the game on this interval with the payoff function

$$\kappa^j(x, u) := \sum_{i=1}^2 \left[\int_{j\Delta t}^{(j+1)\Delta t} h_i(x(\tau), u(\tau)) d\tau + q_i \sqrt{x((j+1)\Delta t)} \right].$$

The initial data on this interval is the value of the trajectory on the previous interval at the boundary point $j\Delta t$. The conditionally cooperative trajectories $\hat{x}^*(t)$ for the deterministic and stochastic forecasts and without forecast are shown in Fig. 3.6. The resource stock $\hat{x}^*(t)$ grows faster without forecast. The deterministic forecast slows this growth down, and the stochastic forecast slows it down even more. However, this does not mean that the payoff becomes less. The instantaneous payoff functions

$$h_{1,2}(\hat{x}^*(t), \hat{u}^*(t)) := \sum_{i=1}^2 h_i(\hat{x}^*(t), \hat{u}^*(t))$$

have been calculated along the conditionally cooperative trajectories $\hat{x}^*(t)$ for the systems with the deterministic and stochastic forecast and without forecast. They are shown in Fig. 3.7. At the beginning part of the game interval, the order of the payoff functions is opposite to that of the resource stock: the greatest value corresponds to the stochastic forecast, and the lowest value is obtained without forecast.

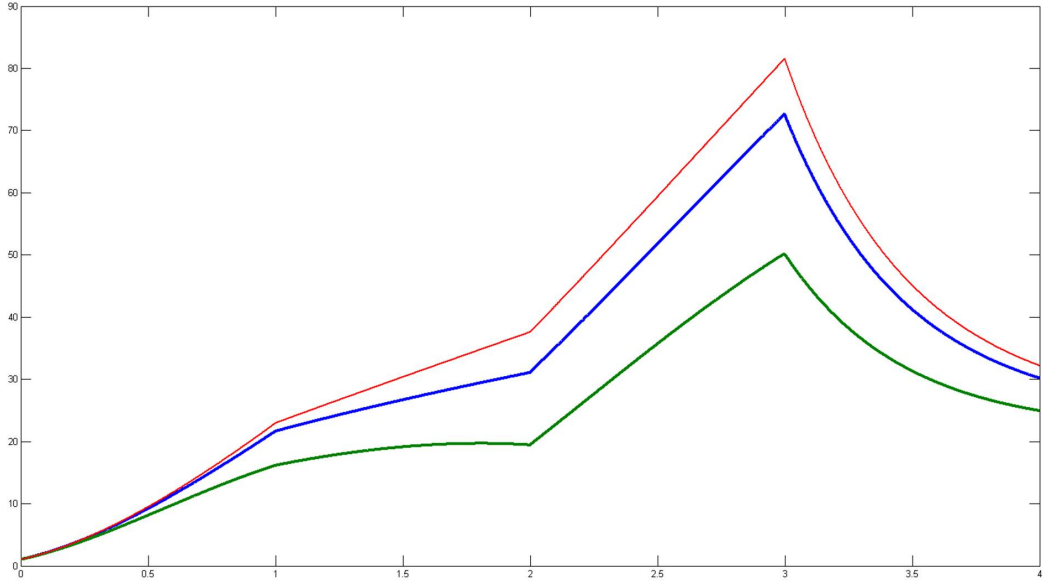


Figure 3.6: Resource stock trajectory composed from cooperative truncated games: without forecast (red), with deterministic forecast (blue), and with stochastic forecast (green).

The residual payoff functions

$$K^c(t, T) := \int_t^T h_{1,2}(\hat{x}^*(\tau), \hat{u}^*(\tau)) d\tau + (q_1 + q_2) \sqrt{\hat{x}^*(T)}, \quad t \in [0, T], \quad (3.88)$$

along the trajectories $\hat{x}^*(t)$ are shown in Fig. 3.8 by solid lines. The total payoff is $K^c(0, T)$. This value is equal to 34.04 (stochastic forecast), 34.0 (deterministic forecast), and 33.54 (without forecast). The dashed lines in Fig. 3.8 show the expected values of the residual payoff function in the current truncated subgame $\bar{\Gamma}_j^c(x_{j,0}, j\Delta t, T)$. These subgames change at the switching points $t = j = 1, 2, 3$, implying jumps of the expected residual payoff values.

Analysis of Figs. 3.7 and 3.8 and similar numerical examples of the resource extraction game have led to the following conclusions. The total cooperative payoff function $K^c(0, T)$ is nearly the same for the three algorithms considered. The main difference consists in the payoff distribution over time (Fig. 3.7) and the growth of the resource stock (Fig. 3.6). The deterministic forecast is "cautious" compared to the algorithm without forecast because it prefers to increase the payoff and give up the resource stock value at the beginning of the game. The stochastic forecast is even more "cautious" for the same reason.

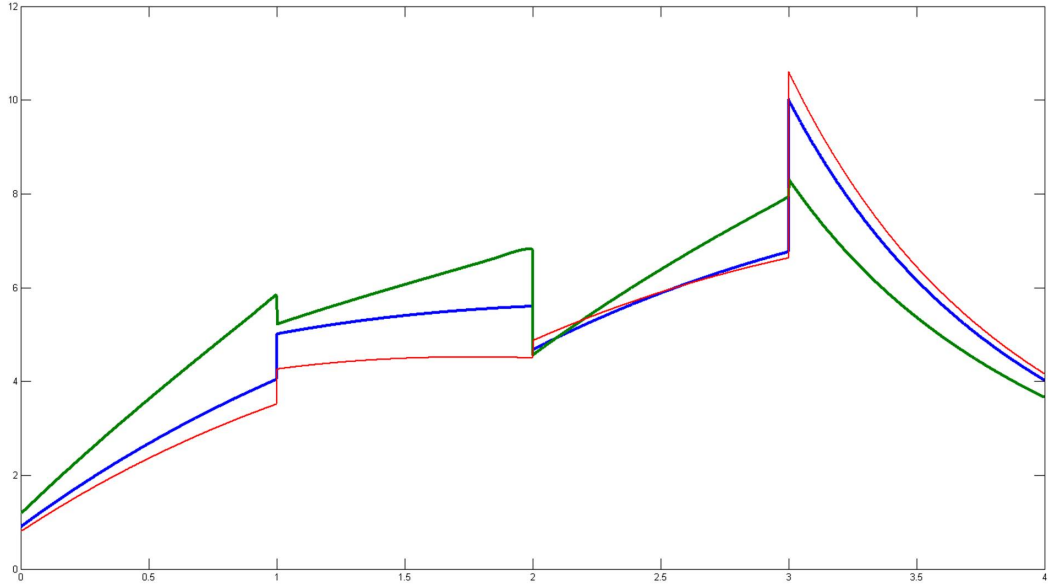


Figure 3.7: Instantaneous cooperative payoff functions $h_{1,2}(\hat{x}^*(t), \hat{u}^*(t))$: without forecast (red), with deterministic forecast (blue), and with stochastic forecast (green).

The continuous time imputation function $\xi_i(t, T)$ must satisfy the group and individual rationality conditions and be time-consistent [189; 193]. Sufficient conditions are obtained for the Nash equilibrium point that gives the proportional imputation for the truncated cooperative subgame $\bar{\Gamma}_j^c(x_{j,0}, j\Delta t, T)$, $0 \leq j \leq l - 1$,

$$\begin{aligned} \xi^{(j)i}(x, t, T) &:= \frac{V_i^j(t, x)}{\sum_{k=1}^2 V_k^j(t, x)} W^j(t, x) = \\ &= \frac{\left[A_i^j(t) \sqrt{x} + C_i^j(t) \right]}{\sum_{k=1}^2 \left[A_k^j(t) \sqrt{x} + C_k^j(t) \right]} \left[A^j(t) \sqrt{x} + C^j(t) \right]. \end{aligned}$$

We denote the corresponding IDP by

$$B_j(t, x_j^*) := \left[B_1^j(t, x_j^*), B_2^j(t, x_j^*) \right]. \quad (3.89)$$

The general formula for the IDP $B_i^j(t, x_j^*)$ in the stochastic resource extraction game

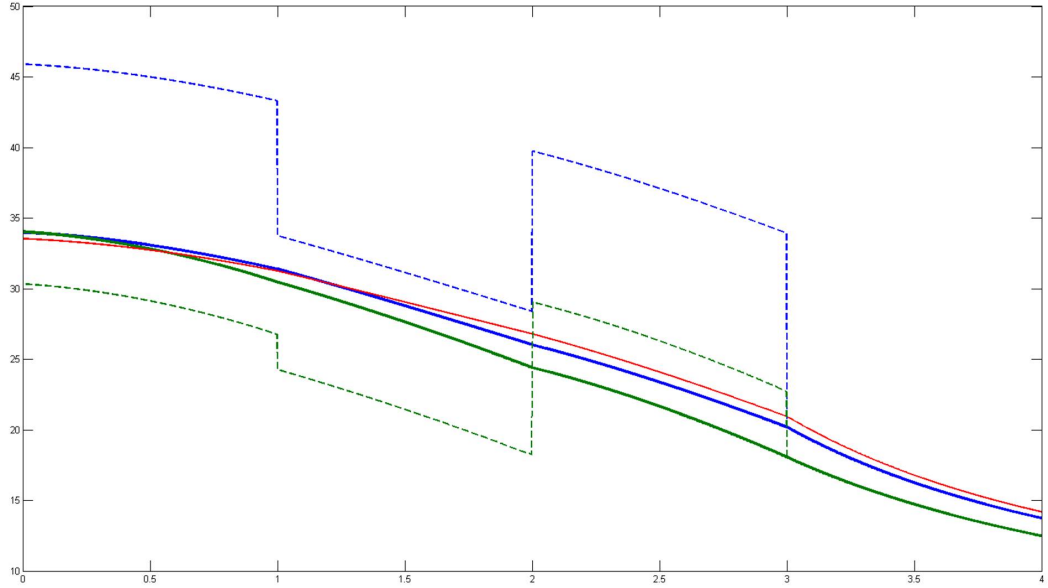


Figure 3.8: Residual payoff functions $K^c(t, T)$: without forecast (red), with deterministic forecast (blue), and with stochastic forecast (green). Expected values of residual payoff function in current truncated subgame $\bar{\Gamma}_j^c(x_{j,0}, j\Delta t, T)$ (dashed lines).

was derived by Yeung and Petrosyan [189]:

$$\begin{aligned}
 B_i^j(\tau, x_j^*) &= - \left[\xi_t^{(j)i}(x_j^*, t, T) \Big|_{t=\tau} \right] \\
 &- \left[\xi_{x_j^*}^{(j)i}(x_j^*, t, T) \Big|_{t=\tau} \right] \left[a_j \sqrt{x_j^*} - b_j x_j^* - \frac{x_j^*}{4[c_1 + A^j(t)]^2} - \frac{x_j^*}{4[c_2 + A^j(t)]^2} \right] - \\
 &- I(j, t) \cdot \frac{1}{2} \sigma^2 (x_j^*)^2 \left[\xi_{x_j^*, x_j^*}^{(j)i}(t, x_j^*) \Big|_{t=\tau} \right], \quad i = 1, 2. \quad (3.90)
 \end{aligned}$$

The combined truncated distribution $\hat{B}_i(t, \hat{x}^*)$, $i = 1, 2$, is composed from $B_i^j(\tau, x_j^*)$ by (3.72) within the dynamic updating approach. The last term in (3.90) is equal to zero on the time interval $]j\Delta t, (j+1)\Delta t]$ because $I(j, t) = 0$. The partial derivatives $\xi_t^{(j)i}$ and $\xi_x^{(j)i}$ are explicitly represented by the functions \hat{x}^* , A^j , C^j , A_1^j , C_1^j , A_2^j , and C_2^j , which are obtained by integration of the differential equations.

The resulting IDPs are shown in Fig 3.9 for the solutions with the deterministic and stochastic forecasts. This is a distribution of the instantaneous payoff function h_{12} from Fig. 3.7 between the players. The sum of the solid and dashed lines in Fig. 3.9 is equal to the line in Fig. 3.7 of the same colour. Each player may obtain greater or less than a half of the full payoff at a time, and this share depends both on the instant and the forecast algorithm.

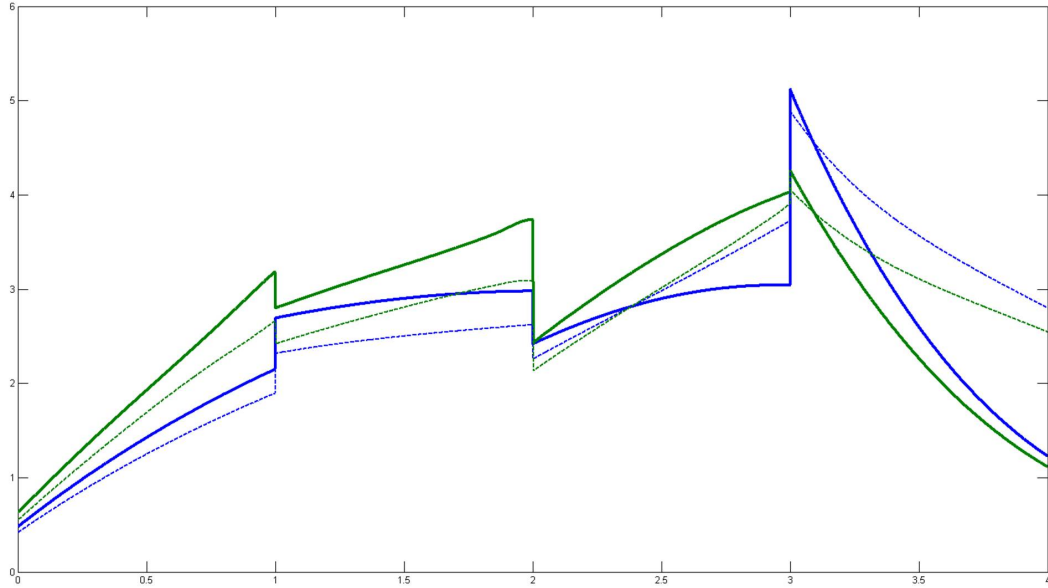


Figure 3.9: IDP for deterministic forecast (blue) and stochastic forecast (green). $\hat{B}_1(t, \hat{x}^*(t))$ — IDP of first player (solid line). $\hat{B}_2(t, \hat{x}^*(t))$ — IDP of second player (dashed line).

3.3 Cooperative Differential Games with Dynamic Updating and Random Horizon

3.3.1 Random Truncated Subgame

Suppose that information for the players is updated at fixed time instants $t = t_0 + j\Delta t$, $j = 0, \dots, l$, where $t_0 < \Delta t < T$, $l = \frac{T-t_0}{\Delta t} - 1$. During the time interval $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, the players have complete information about the game dynamics described by $g(t, x, u)$ and the payoff function described by $h_i(x(t), u(t))$. The problem is that the players are not sure about the period of time when this information is valid: all they know is that it is defined on the time interval $[t_0 + j\Delta t, \bar{T}_j]$, where the duration \bar{T}_j is a random variable with known characteristics. We denote by \bar{t}_j , $j = 0, \dots, l$, the realization of the random variable \bar{T}_j .

As mentioned before, during the time interval $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$ the players have complete information about the game dynamics and payoff function on the time interval $[t_0 + j\Delta t, \bar{T}_j]$, where \bar{T}_j is a random variable taking values from the time interval $[\max(t_0 + (j+1)\Delta t, \bar{t}_{j-1}), T]$, and \bar{t}_{j-1} is a realization of the random variable \bar{T}_{j-1} (\bar{T}_{j-1} is realized at the time instant $t = t_0 + j\Delta t$). At the time

instant $t = t_0 + (j + 1)\Delta t$, the information about the game is being updated, and the random variable \bar{T}_j is realized, i.e., \bar{t}_j becomes known to the players. On the next time interval $(t_0 + (j + 1)\Delta t, t_0 + (j + 2)\Delta t]$, the players have complete information about the game structure on the time interval $(t_0 + (j + 1)\Delta t, \bar{T}_{j+1}]$, where \bar{T}_{j+1} is a random variable taking values from the time interval $[\max(t_0 + (j + 2)\Delta t, \bar{t}_j), T]$. For $j = 0$, suppose that $\bar{t}_{j-1} = 0$.

It may remain unclear why the random variable \bar{T}_j is realized at the time instants $t = t_0 + (j + 1)\Delta t$ but its value \bar{t}_j exceeds $t = t_0 + (j + 1)\Delta t$. An interpretation can be the following. Suppose that at the time instants $t = t_0 + j\Delta t$, the players receive information about the game, but to accurately estimate the duration of information horizon (the random variable \bar{T}_j) they need some time Δt for making calculations, etc. At the time instant $t = t_0 + (j + 1)\Delta t$, the calculations are performed, and a certain value of the information horizon \bar{T}_j becomes known to the players, i.e., \bar{t}_j . Another interpretation is that during the current Δt -time interval $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, the players receive additional information, which helps them estimating the time on which the information about the process is certain, i.e., defining the value of the information horizon \bar{t}_j . At the time instant $t = t_0 + (j + 1)\Delta t$, after the estimation is performed, the players receive new information about the game structure with the random information horizon \bar{T}_{j+1} , and the same procedure continues.

Definition 3.3.1 *Let $j = 0, \dots, l$. A random truncated subgame $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t)$ is defined on the time interval $[t_0 + j\Delta t, \bar{T}_j]$, where \bar{T}_j is a random variable taking values from the time interval $[\max(t_0 + (j + 1)\Delta t, \bar{t}_{j-1}), T]$, and \bar{t}_{j-1} is a realization of the random horizon \bar{T}_{j-1} in the previous truncated subgame $\bar{\Gamma}_{j-1}(x_{j-1,0}, t_0 + (j - 1)\Delta t)$. The realization of \bar{T}_{j-1} occurs at the time instant $t = t_0 + j\Delta t$. The dynamical system and payoff function on the time interval $[t_0 + j\Delta t, \bar{T}_j]$ coincide with those of the game $\Gamma(x_0, T - t_0)$ on the same time interval. The dynamical system and the initial condition of the truncated subgame $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t)$ have the following form:*

$$\dot{x} = g(t, x, u), \quad x(t_0 + j\Delta t) = x_{j,0}. \quad (3.91)$$

The payoff function of player i in random truncated subgame j is

$$K_i^j(x_{j,0}, t_0 + j\Delta t; u) = \int_{t_0 + j\Delta t}^T \int_{t_0 + j\Delta t}^t h_i(x(\tau), u(\tau)) d\tau dF_j(t), \quad (3.92)$$

where $F_j(t)$ is a distribution function of \bar{T}_j :

$$\int_{t_0 + j\Delta t}^T dF_j(t) = \int_{\max(t_0 + (j+1)\Delta t, \bar{t}_{j-1})}^T dF_j(t) = 1. \quad (3.93)$$

By the definition, $F_j(t)$ is a conditional distribution function, i.e., $F_j(t) = F_j(t \mid \bar{T}_{j-1} = \bar{t}_{j-1})$. For the sake of brevity, we will use the notation $F_j(t) = F_j(t \mid \bar{T}_{j-1} = \bar{t}_{j-1})$.

Suppose that the realization of the random horizon \bar{T}_{j-1} in the game $\bar{\Gamma}_{j-1}(x_{j-1,0}, t_0 + (j-1)\Delta t)$ exceeds the time $t = t_0 + (j+1)\Delta t$:

$$\bar{t}_{j-1} > t_0 + (j+1)\Delta t. \quad (3.94)$$

Then the random horizon \bar{T}_j must exceed the realization of \bar{T}_{j-1} because the information about the game structure is already known on the time interval $[t_0 + j\Delta t, \bar{t}_{j-1}]$. That is why in formula (3.93), the probability of \bar{T}_j taking values from the time interval $[t_0 + j\Delta t, \bar{t}_{j-1}]$ equals zero:

$$\int_{t_0 + j\Delta t}^{\max(t_0 + (j+1)\Delta t, \bar{t}_{j-1})} dF_j(t) = 0. \quad (3.95)$$

In the papers on cooperative differential games with random horizon [194; 195; 196; 197; 198], the distribution function of \bar{T}_j was defined on the infinite time interval. In this section, \bar{T}_j takes values from a finite time interval because the original game is defined on the finite time interval $[t_0, T]$.

In [199], the order of integration in the double integral (3.92) was changed

according to Tonelli's theorem:

$$K_i^j(x_{j,0}, t_0 + j\Delta t; u) = \int_{t_0+j\Delta t}^T (1 - F_j(\tau)) h_i(x(\tau), u(\tau)) d\tau. \quad (3.96)$$

3.3.1.1 Solution of Random Truncated Cooperative Subgame

Consider a truncated cooperative subgame $\bar{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t)$ defined on the time interval $[t_0 + j\Delta t, \bar{T}_j]$ with the initial condition $x(t_0 + j\Delta t) = x_{j,0}$, where \bar{T}_j is a random variable with the distribution function (3.93). Classically in the first step of cooperative differential games, the cooperative strategies and corresponding cooperative trajectory are defined. In the second step, a rule for allocating the cooperative payoff between the players along the cooperative trajectory is defined. For this purpose, a characteristic function and a corresponding cooperative solution are defined. The total payoff of the players to be maximized in this game is

$$\sum_{i \in N} K_i^j(x_{j,0}, t_0 + j\Delta t; u) = \sum_{i \in N} \int_{t_0+j\Delta t}^T \int_{t_0+j\Delta t}^t h_i(x(\tau), u(\tau)) d\tau dF_j(t) \quad (3.97)$$

subject to

$$\dot{x} = g(t, x, u), \quad x(t_0 + j\Delta t) = x_{j,0}. \quad (3.98)$$

This is an optimal control problem. Sufficient conditions for the solution and the optimal feedback are given by Theorem 3.3.1, which was first presented in [197]. We denote by $W^{(j\Delta t)}(t, x)$ the maximum joint payoff of the players (3.97):

$$W^{(j\Delta t)}(t, x) = \max_{u \in U} \left\{ \sum_{i \in N} K_i^j(x, t; u) \right\}, \quad (3.99)$$

where x and t are the initial state and time of the random truncated subgame, respectively, and $U = U_1 \times \dots \times U_n$.

Theorem 3.3.1 *Assume that there exists a continuously differential function $W^{(j\Delta t)}(t, x) : [t_0 + j\Delta t, \bar{T}_j] \times R^m \rightarrow R$ satisfying the partial differential equation*

$$\frac{f_j(t)}{1 - F_j(t)} W^{(j\Delta t)}(t, x) = W_t^{(j\Delta t)}(t, x) + \max_{u \in U} \left\{ \sum_{i=1}^n h_i(t, x, u) + W_x^{(j\Delta t)}(t, x) g(t, x, u) \right\}, \quad (3.100)$$

where $\lim_{t \rightarrow T^-} W^{(j\Delta t)}(t, x) = 0$, $f_j(t)$ is the density function of the random variable \bar{T}_j (3.93). Assume that the maximum in (3.100) is achieved under controls $u_j^*(t, x)$. Then $u_j^*(t, x)$ is optimal in the control problem (3.97), (3.98).

Theorem 3.3.1, presented in [197], requires that the function $W^{(j\Delta t)}$ be C^1 . However, it is possible to assume only continuity considering viscosity solutions using the Subbotin approach [183; 184]. But due to the shortage of space, it is not possible to properly introduce and define this solution here. In the example model, the solution $W^{(j\Delta t)}$ will be obtained from C^1 .

3.3.1.2 Conditionally Cooperative Trajectory

During the game $\Gamma(x_0, T - t_0)$, the players possess only truncated information about its structure. Obviously, this is not enough to construct optimal control and the corresponding trajectory for the game $\Gamma(x_0, T - t_0)$. As a cooperative trajectory in the game $\Gamma(x_0, T - t_0)$, a conditionally cooperative trajectory will be used:

Definition 3.3.2 A conditionally cooperative trajectory $\{\hat{x}^*(t)\}_{t=t_0}^T$ is defined as a composition of the cooperative trajectories $x_j^*(t)$ in the truncated cooperative subgames $\bar{\Gamma}_j^c(x_{j-1}^*(t_0 + j\Delta t), t_0 + j\Delta t)$ defined on the successive time intervals $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$:

$$\{\hat{x}^*(t)\}_{t_0}^T = \begin{cases} x_0^*(t), & t \in [t_0, t_0 + \Delta t), \\ \dots, \\ x_j^*(t), & t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t), \\ \dots, \\ x_l^*(t), & t \in [t_0 + l\Delta t, t_0 + (l + 1)\Delta t), \end{cases} \quad (3.101)$$

On the time interval $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, the conditionally cooperative trajectory coincides with the cooperative trajectory $x_j^*(t)$ in the truncated cooperative subgame $\bar{\Gamma}_j^c(x_{j-1}^*(t_0 + j\Delta t), t_0 + j\Delta t)$. At the time instant $t = t_0 + (j + 1)\Delta t$, information about the game structure is updated in the position $x_j^*(t_0 + (j + 1)\Delta t)$. On the time interval $(t_0 + (j + 1)\Delta t, t_0 + (j + 2)\Delta t]$, the trajectory $\hat{x}^*(t)$ coincides with the cooperative trajectory $x_{j+1}^*(t)$ in the truncated cooperative subgame $\bar{\Gamma}_{j+1}^c(x_j^*(t_0 + (j + 1)\Delta t), t_0 + (j + 1)\Delta t)$ starting at the time instant $t = t_0 + (j + 1)\Delta t$ in the position $x_j^*(t_0 + (j + 1)\Delta t)$. For $j = 0$, $x_{j-1}^*(t_0 + j\Delta t) = x_0$.

3.3.1.3 Characteristic Function

For each coalition $S \subset N$ and $j = 0, \dots, l$, we define the values of characteristic function as in [185]:

$$V_j(S; x_{j,0}^*, t_0 + j\Delta t) = \begin{cases} \sum_{i \in N} K_i^j(x_{j,0}^*, t_0 + j\Delta t; u_j^*), & S = N, \\ \tilde{V}_j(S, x_{j,0}^*, t_0 + j\Delta t), & S \subset N, \\ 0, & S = \emptyset, \end{cases} \quad (3.102)$$

where $\tilde{V}_j(S, x_{j,0}^*, t_0 + j\Delta t)$ is defined as the total payoff of the players from coalition S in the Nash equilibrium $u_j^{NE} = (u_1^{NE,j}, \dots, u_n^{NE,j})$ in the game with following set of players: coalition S (acting as one player) and the players from the set $N \setminus S$, i.e., in the game with $|N \setminus S| + 1$ players.

An imputation $\xi_j(x_{j,0}^*, t_0 + j\Delta t)$ in each random truncated cooperative subgame $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t)$ is defined as an arbitrary vector satisfying the conditions

$$\begin{aligned} \xi_i^j(x_{j,0}^*, t_0 + j\Delta t) &\geq V_j(\{i\}, x_{j,0}^*, t_0 + j\Delta t), \quad i \in N, \\ \sum_{i \in N} \xi_i^j(x_{j,0}^*, t_0 + j\Delta t) &= V_j(N, x_{j,0}^*, t_0 + j\Delta t). \end{aligned} \quad (3.103)$$

We denote by $E_j(x_{j,0}^*, t_0 + j\Delta t)$ the set of all possible imputations in the random truncated subgame. As an optimality principle (or solution)

$$W_j(x_{j,0}^*, t_0 + j\Delta t) \subset E_j(x_{j,0}^*, t_0 + j\Delta t) \quad (3.104)$$

in each random truncated cooperative subgame $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t)$, the IDP-core was

introduced in [100]. Construction of this solution is based on a special class of IDPs [187].

Definition 3.3.3 *A function $\beta_j(t, x_j^*)$, $t \in [t_0 + j\Delta t, \bar{T}_j]$ is called an Imputation Distribution Procedure for an imputation $\xi_j(x_{j,0}^*, t_0 + j\Delta t) \in E_j(x_{j,0}^*, t_0 + j\Delta t)$ if*

$$\xi_j(x_{j,0}^*, t_0 + j\Delta t) = \int_{t_0 + j\Delta t}^T (1 - F_j(\tau))\beta_j(t, x_j^*(t))d\tau. \quad (3.105)$$

Using the IDP $\beta_j(t, x_j^*)$, we can define a rule for allocating the imputation $\xi_j(x_{j,0}^*, t_0 + j\Delta t)$ on the time interval $[t_0 + j\Delta t, \bar{T}_j]$, where \bar{T}_j is a random variable. It is obvious that the number of functions $\beta_j(t, x_j^*(t))$ satisfying equation (3.105) is infinite (i.e., the ways to allocate the cooperative payoff between the players is infinite). However, according to the formula for the class of games with random horizons presented in [194; 195],

$$\beta_j(t, x_j^*(t)) = \frac{f_j(t)}{1 - F_j(t)}\xi_j(x_j^*(t), t) - \frac{d}{d\tau}\xi_j(x_j^*(t), t). \quad (3.106)$$

This formula allows defining the unique $\beta_j(t, x_j^*(t))$ that ensures the time-consistency property [187] of the imputation $\xi_j(x_{j,0}^*, t_0 + j\Delta t)$ or cooperative solution $W_j(x_{j,0}^*, t_0 + j\Delta t)$ (in the case of multiple principles of optimality).

Definition 3.3.4 *A solution $W_j(x_{j,0}^*, t_0 + j\Delta t)$ ($\xi_j(x_{j,0}^*, t_0 + j\Delta t)$) is called time-consistent if for any imputation $\xi_j(x_{j,0}^*, t_0 + j\Delta t) \in W_j(x_{j,0}^*, t_0 + j\Delta t)$, there exists an IDP $\beta_j(t, x_j^*)$ satisfying the following inclusion $\forall t \in [t_0 + j\Delta t, T]$:*

$$\left\{ \int_t^T (1 - F_j(t))\beta_i^j(\tau, x_j^*)d\tau \right\} \in W(x_j^*(t), t) \left(\left\{ \int_t^T (1 - F_j(t))\beta_i^j(\tau, x_j^*)d\tau \right\} = \xi_j(x_{j,0}^*, t_0 + j\Delta t) \right).$$

3.3.1.4 IDP-core

The approach proposed in [100] is employed: the functions are constructed that can be used as IDPs for some imputations, and then the corresponding solutions

are composed. Suppose that the characteristic function $V_j(S; x_j^*(t), t)$, $S \subset N$, is continuously differentiable with respect to t , $t \in [t_0 + j\Delta t, T]$ along the cooperative trajectory $x_j^*(t)$. We introduce the following notation:

$$U_j(S; x_j^*(t), t) = -\frac{d}{dt}V_j(S; x_j^*(t), t), \quad (3.107)$$

where $t \in [t_0 + j\Delta t, T]$ and $S \subseteq N$.

We define $B_j(t, x_j^*)$ as a set of integrable vector functions $\beta_j(t, x_j^*)$ satisfying the following inequalities:

$$\begin{aligned} B_j(t, x_j^*) = \{ & \beta_j(t, x_j^*) = (\beta_1^j(t, x_j^*), \dots, \beta_n^j(t, x_j^*)) : \\ & \sum_{i \in S} (1 - F_j(t)) \beta_i^j(t, x_j^*) \geq U_j(S, x_j^*(t), t), \\ & \sum_{i \in N} \beta_i^j(t, x_j^*) = U_j(N, x_j^*(t), t), \forall S \subset N \}. \end{aligned} \quad (3.108)$$

Suppose that $B_j(t, x_j^*) \neq \emptyset$, $\forall t \in [t_0 + j\Delta t, T]$, $j = 0, \dots, l$. Using the set $B_j(t, x_j^*)$, we define the following set of vectors:

Definition 3.3.5 *The set of all possible vectors $\xi_j(x_j^*(t), t)$ for some integrable selectors $\beta_j(t, x_j^*) \in B_j(t, x_j^*)$ will be called the IDP-core and denoted by $\overline{C}_j(x_j^*(t), t)$, where*

$$\overline{C}_j(x_j^*(t), t) = \{\xi_j(x_j^*(t), t), t \in [t_0 + j\Delta t, T]\} \quad (3.109)$$

and for $t \in [t_0 + j\Delta t, T]$,

$$\xi_j(x_j^*(t), t) = \int_t^T (1 - F_j(\tau)) \beta_j(\tau, x_j^*) d\tau. \quad (3.110)$$

It is easy to suggest that the distribution of the total payoff of players in the game $\Gamma(x_0, T - t_0)$ along the conditionally cooperative trajectory $\{\hat{x}^*(t)\}_{t=t_0}^T$ can be organized as a composition of IDPs for each time interval $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, l$, in accordance with the structure of the game $\Gamma(x_0, T - t_0)$. This will be formalized in this section as a new solution concept.

The family of sets $W_j(x_{j,0}^*, t_0 + j\Delta t) = \overline{C}_j(x_{j,0}^*, t_0 + j\Delta t)$ do not compose directly a solution of the game $\Gamma(x_0, T - t_0)$. For any $j = 0, \dots, l$, the optimal solution of the truncated subgame $\overline{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t)$ is defined on the time interval $[t_0 + j\Delta t, t_0 + j\Delta t + \overline{T}_j]$. This particular solution makes sense on the interval $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$ only because the information about the game structure is updated after every Δt time interval, and it is irrelevant to use a solution based on the outdated information. The necessary information can be extracted using the IDP for each truncated subgame. Therefore, to construct the optimal solution for the entire game $\Gamma(x_0, T - t_0)$, we use the set of IDPs $B_j(t)$ instead of the set of imputations $\overline{C}_j(x_{j,0}^*, t_0 + j\Delta t)$.

3.3.2 Cooperative Solution Concept for Game with Dynamic Updating and Random Horizon

To introduce a solution concept for the differential game $\Gamma(x_0, T - t_0)$ with the dynamic updating approach, we use the family of sets $B_j(t, x_j^*)$, $j = 0, \dots, l$. First, the set of IDPs for the entire game $\Gamma(x_0, T - t_0)$ is constructed in the following way: for each fixed composition of IDPs $\beta_j(t, x_j^*) \in B_j(t, x_j^*)$, $j = 0, \dots, l$, the resulting IDP $\hat{\beta}(t, \hat{x}^*)$ is defined.

Definition 3.3.6 *The resulting IDP $\hat{\beta}(t, \hat{x}^*)$ is a function defined as a combination of imputation distribution procedures $\beta_j(t, x_j^*) \in B_j(t, x_j^*)$ in all truncated cooperative subgames $\overline{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t)$, $j = 0, \dots, l$:*

$$\hat{\beta}(t, \hat{x}^*) = \begin{cases} (1 - F_0(t))\beta_0(t, x_0^*), t \in [t_0, t_0 + \Delta t], \\ \dots \\ (1 - F_j(t))\beta_j(t, x_j^*), t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t], \\ \dots \\ (1 - F_l(t))\beta_l(t, x_l^*), t \in [t_0 + l\Delta t, t_0 + (l + 1)\Delta t]. \end{cases} \quad (3.111)$$

The set of all possible resulting IDPs $\hat{\beta}(t, \hat{x}^*)$ (3.111) for different compositions $\beta_j(t, x_j^*) \in B_j(t, x_j^*)$, $j = 0, \dots, l$ will be denoted by $\hat{B}(t, \hat{x}^*)$.

Using the resulting IDP $\hat{\beta}(t, \hat{x}^*) \in \hat{B}(t)$, we determine a resulting imputation that can be used as an imputation in the game $\Gamma(x_0, T - t_0)$ with the dynamic

updating approach. But the question stands: will the resulting imputation actually allocate the joint cooperative payoff along the conditionally cooperative trajectory $\hat{x}^*(t)$? This fact is proved in Theorem 3.3.2.

Definition 3.3.7 *The resulting imputation $\hat{\xi}(x_0, T - t_0)$ is a vector defined in the following way:*

$$\hat{\xi}(x_0, T - t_0) = \int_{t_0}^T \hat{\beta}(\tau, \hat{x}^*(\tau)) d\tau = \sum_{j=0}^l \left[\int_{j\Delta t}^{(j+1)\Delta t} (1 - F_j(\tau)) \beta_j(\tau, x_j^*(\tau)) d\tau \right] \quad (3.112)$$

The resulting solution $\hat{W}(x_0, T - t_0)$ is the set of all resulting imputations $\hat{\xi}(x_0, T - t_0)$ composed by (3.111), (3.112). In the game models with dynamic updating, the set \hat{W} is used as a solution.

Theorem 3.3.2 *With any $\hat{\xi}(x_0, T - t_0) \in \hat{W}(x_0, T - t_0)$ it is possible to allocate the joint payoff of the players (3.97) along the conditionally cooperative trajectory $\hat{x}^*(t)$ during the game $\Gamma(x_0, T - t_0)$, and $\forall t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, l$:*

$$\sum_{i=1}^n \int_{t_0}^t \hat{\beta}_i(\tau, \hat{x}^*(\tau)) d\tau = \sum_{i=1}^n \left[\sum_{k=0}^{j-1} \left[\int_{k\Delta t}^{(k+1)\Delta t} (1 - F_k(\tau)) h_i(\hat{x}^*(\tau), \hat{u}^*(\tau)) d\tau \right] + \int_{t_0+j\Delta t}^t (1 - F_j(\tau)) h_i(\hat{x}^*(\tau), \hat{u}^*(\tau)) d\tau \right]. \quad (3.113)$$

Proof To prove this theorem, we start from the last random truncated subgame $\bar{\Gamma}_l(x_{l,0}^*, t_0 + l\Delta t)$: we show that $\forall t \in [t_0 + l\Delta t, T]$,

$$\sum_{i=1}^n \int_{t_0+l\Delta t}^t \hat{\beta}_i(\tau, \hat{x}^*(\tau)) d\tau = \sum_{i=1}^n \int_{t_0+l\Delta t}^t (1 - F_l(\tau)) h_i(\hat{x}^*(\tau), \hat{u}^*(\tau)) d\tau. \quad (3.114)$$

Indeed, the maximum joint payoff in this game is defined by the function $W^{(l\Delta t)}(t_0 +$

$l\Delta t, x_{i,0}^*$) (3.99). By the definition of this function, $\forall t \in [t_0 + l\Delta t, T]$ we have

$$\begin{aligned} W^{(l\Delta t)}(t, \hat{x}^*(t)) &= \\ \max_{u \in U} \left\{ \sum_{i \in N} K_i^l(\hat{x}^*(t), t; u) \right\} &= \sum_{i=1}^n \int_t^T (1 - F_l(\tau)) h_i(\hat{x}^*(\tau), \hat{u}^*(\tau)) d\tau = \\ \sum_{i=1}^n \int_t^T \xi_i^l(\hat{x}^*(\tau), \tau) d\tau &= \sum_{i=1}^n \int_t^T \beta_i^l(\tau, \hat{x}^*(\tau)) d\tau = \sum_{i=1}^n \int_t^T \hat{\beta}_i(\tau, \hat{x}^*(\tau)) d\tau. \end{aligned} \quad (3.115)$$

However,

$$\begin{aligned} W^{(l\Delta t)}(t_0 + l\Delta t, x_{i,0}^*) - W^{(l\Delta t)}(t, \hat{x}^*(t)) &= \\ \sum_{i=1}^n \int_{t_0 + l\Delta t}^t (1 - F_l(\tau)) h_i(\hat{x}^*(\tau), \hat{u}^*(\tau)) d\tau, \end{aligned} \quad (3.116)$$

where $\forall t \in [t_0 + l\Delta t, T]$. From (3.115) and (3.116) it follows that (3.114) holds $\forall t \in [t_0 + l\Delta t, T]$. Using this result, we prove that it also holds for the random truncated subgame $\bar{\Gamma}_{l-1}(x_{i,l-1,0}^*, t_0 + (l-1)\Delta t)$: we show that

$$\begin{aligned} \sum_{i=1}^n \int_{t_0 + (l-1)\Delta t}^t \hat{\beta}_i(\tau, \hat{x}^*(\tau)) d\tau &= \\ \sum_{i=1}^n \int_{t_0 + (l-1)\Delta t}^t (1 - F_{l-1}(\tau)) h_i(\hat{x}^*(\tau), \hat{u}^*(\tau)) d\tau, \end{aligned} \quad (3.117)$$

where $\forall t \in [t_0 + (l-1)\Delta t, t_0 + l\Delta t]$. Like in the game $\bar{\Gamma}_l(x_{i,0}^*, t_0 + l\Delta t)$, $\forall t \in [t_0 + (l-1)\Delta t, t_0 + l\Delta t]$ we obtain

$$\begin{aligned} W^{((l-1)\Delta t)}(t_0 + (l-1)\Delta t, x_{i,l-1,0}^*) - W^{((l-1)\Delta t)}(t, \hat{x}^*(t)) &= \\ \sum_{i=1}^n \int_{t_0 + (l-1)\Delta t}^t (1 - F_{l-1}(\tau)) h_i(\hat{x}^*(\tau), \hat{u}^*(\tau)) d\tau. \end{aligned} \quad (3.118)$$

Hence, (3.117) is satisfied. We need to proceed to the first random truncated

subgame $\bar{\Gamma}_0(x_0, t_0)$. Finally combining the results (3.114) and (3.117), we show that (3.113) holds $\forall t \in [t_0, T]$. This completes the proof. \square

3.3.2.1 Time-consistency of Solution Concept

It is easy to see that the resulting solution $\hat{W}(x_0, T - t_0)$ is time-consistent, but there is another surprising property of $\hat{W}(x_0, T - t_0)$.

Theorem 3.3.3 *The resulting solution $\hat{W}(x_0, T - t_0)$ is strong time-consistent in the game $\Gamma(x_0, T - t_0)$.*

Proof Suppose that in the game $\Gamma(x_0, T - t_0)$, the players agreed to choose an imputation $\hat{\xi}(x_0, T - t_0) \in \hat{W}(x_0, T - t_0)$. This means that during the game, in each random truncated subgame $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t)$ they agreed on choosing the imputation $\xi_j(x_{j,0}^*, t_0 + j\Delta t) \in \bar{C}_j(x_{j,0}^*, t_0 + j\Delta t)$ with the corresponding IDP $\beta_j(t, x_j^*) \in B_j(t, x_j^*)$, $t \in [t_0 + j\Delta t, T]$. In fact, during the game, the players use the IDP $\hat{\beta}(t, \hat{x}^*) = \beta_j(t, x_j^*)$ and allocate the cooperative payoff in the following way:

$$\int_{t_0}^T \hat{\beta}(\tau, \hat{x}^*(\tau)) d\tau = \sum_{j=0}^l \int_{t_0 + j\Delta t}^{t_0 + (j+1)\Delta t} (1 - F_j(t)) \beta_j(t, x_j^*) dt.$$

Suppose that at a given time instant $t = t_{br}$, where $t_{br} \in [t_0 + k\Delta t, T]$, in the random truncated subgame $\bar{\Gamma}_k^c(x_{k,0}^*, t_0 + k\Delta t)$ the players decide to choose another imputation $\xi'_k(x_k^*(t_{br}), t_{br})$ from the IDP-core $\bar{C}_k(x_k^*(t_{br}), t_{br})$. Therefore, there exists an IDP $\beta'_k(t, x_k^*) \in B_k(t, x_k^*)$, $t \in [t_{br}, T]$, corresponding to the imputation

$$\xi'_k(x_k^*(t_{br}), t_{br}) = \int_{t_{br}}^T (1 - F_j(t)) \beta'_k(t, x_k^*) dt. \quad (3.119)$$

In this case, during the game, the players will allocate the cooperative payoff according to $\hat{\xi}'(x_{t_0}, T - t_0)$ using the following resulting IDP:

$$\hat{\beta}'(t) = \begin{cases} (1 - F_k(t)) \beta_k(t, x_k^*), & t \in [t_0 + k\Delta t, t_{br}), \\ (1 - F_k(t)) \beta'_k(t, x_k^*), & t \in [t_{br}, t_0 + (k+1)\Delta t], \\ (1 - F_j(t)) \beta_j(t, x_j^*), & t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t], \end{cases}$$

where $j \neq k, j = 0, \dots, l$. The corresponding resulting imputation has the following form:

$$\begin{aligned} \hat{\xi}'(x_0, T - t_0) = & \int_{t_0}^T \hat{\beta}'(t, \hat{x}^*) dt = \sum_{\substack{j=0 \\ j \neq k}}^l \int_{t_0+j\Delta t}^{t_0+(j+1)\Delta t} (1 - F_j(t)) \beta_j(t, x_j^*) dt + \\ & \int_{t_0+k\Delta t}^{t_{br}} (1 - F_k(t)) \beta_k(t, x_k^*) dt + \int_{t_{br}}^{t_0+(k+1)\Delta t} (1 - F_k(t)) \beta'_k(t, x_k^*) dt. \end{aligned} \quad (3.120)$$

Since $\beta'_k(t, x_k^*) \in B_k(t, x_k^*)$, $t \in [t_{br}, T]$ the resulting IDP $\hat{\beta}'(t, \hat{x}^*)$ belongs to $\hat{B}(t, \hat{x}^*)$. By the definition of $\hat{W}(x_0, T - t_0)$, all vectors $\hat{\xi}(x_0, T - t_0)$ obtained by formula (3.112) using $\hat{\beta}(t, \hat{x}^*)$ from the set $\hat{B}(t, \hat{x}^*)$ are called the resulting solution $\hat{W}(x_0, T - t_0)$ of the game $\Gamma(x_0, T - t_0)$. In (3.120) the imputation $\hat{\xi}'(x_0, T - t_0)$ is constructed with the IDP $\hat{\beta}'(t, \hat{x}^*)$ from the set $\hat{B}(t, \hat{x}^*)$, and it can be seen as the resulting imputation $\hat{\xi}'(x_{t_0}, T - t_0)$ belonging to the initial solution $\hat{W}(x_0, T - t_0)$. This completes the proof. \square

3.3.3 Random Horizon Dynamic Updating in Cooperative Extraction Differential Game

The following example of the resource extraction game with two players was considered by Jorgensen and Yeung [69]. The problem of time-consistency in this example was studied by Yeung et. al [188]. Three competitive models were implemented: with the stochastic and deterministic forecasts and without forecast. Here the resource extraction game with three players and a special cooperative solution is considered; also, see Section 3.3.1.3 and [100]. An analytical form of the characteristic function for each coalition is derived according to [185] and presented below. Furthermore, the dynamic updating approach with random horizon is applied to the example. In the final part of the example, the strong time-consistency property of the constructed solution is demonstrated.

In the following model, the analytical solution of the problem is derived, but a general analytical solution cannot be found. To apply the dynamic updating approach to the general class of cooperative differential games, we need to solve two main problems. The first one is (3.97) subject to (3.98) for each truncated subgame. Mathematically this is a classical control problem, and there are numerous

methods for solving it. Solving this problem, we obtain the approximate cooperative strategies, cooperative trajectories $x_j^*(t)$, and corresponding joint payoff (3.97). The second problem is defining how to allocate the cooperative payoff between the players. It is necessary to calculate the characteristic function (3.102) for each truncated subgame along the cooperative trajectory. This can be done using coevolutionary algorithms [200] suitable for game-theoretic problems. After calculating the characteristic functions, the solution for each truncated subgame (e.g., the IDP-core) can be determined. Then the corresponding resulting solution (3.112) can be calculated.

3.3.3.1 Original Game

Consider an economy endowed with a single renewable resource and $n \geq 2$ resource extractors (players). We denote by $u_i(t)$ the quantity of the resource extracted by player i at a time instant t , for $i \in N$. Each player controls its rate of extraction. Let $x(t) \in X \subset R$ be the size of the resource stock at a time instant t . The growth dynamics of the renewable resource stock becomes

$$\dot{x} = a\sqrt{x(t)} - bx(t) - \sum_{i=1}^3 u_i, \quad x(t_0) = x_0, \quad (3.121)$$

where $a\sqrt{x(t)} - bx(t)$ is the natural rate of evolution of the resource, and $u_i \in [0, d]$, $d > 0$, $i = \overline{1, 3}$.

The extraction cost for player $i \in N$ depends on the quantity of the resource extracted $u_i(t)$, the resource stock size $x(t)$, and the parameter c_i , $i = \overline{1, 3}$:

$$K_i(x_0, t_0; u) = \int_{t_0}^T \sqrt{u_i(\tau)} - \frac{c_i}{\sqrt{x(\tau)}} u_i(\tau) d\tau, \quad (3.122)$$

where c_i is constant and $c_i \neq c_k$, $\forall i \neq k = \overline{1, 3}$. The set of parameters x_0, T, a, b, d , and c_i , $i = \overline{1, 3}$, is considered such that it is always nonnegative in the corresponding control problem.

3.3.3.2 Random Truncated Subgame

The original game $\Gamma(x_0, T - t_0)$ is defined on the time interval $[t_0, T]$. Suppose that for any $t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, $j = 0, \dots, l$, the players have truncated information about the game structure. It includes information about the dynamical system and payoff function on the time interval $[t_0 + j\Delta t, \bar{T}_j]$, where \bar{T}_j is a truncated exponentially distributed random variable with the distribution $F_j(t)$ and density $f_j(t)$ functions:

$$F_j(t) = \frac{1 - \exp(-\lambda(t - \max(t_0 + (j + 1)\Delta t, \bar{t}_{j-1})))}{1 - \exp(-\lambda(T - \max(t_0 + (j + 1)\Delta t, \bar{t}_{j-1})))}, \quad (3.123)$$

$$f_j(t) = \frac{\lambda \exp(-\lambda(t - \max(t_0 + (j + 1)\Delta t, \bar{t}_{j-1})))}{1 - \exp(-\lambda(T - \max(t_0 + (j + 1)\Delta t, \bar{t}_{j-1})))}. \quad (3.124)$$

The exponential distribution is widely used for describing the time between events in a Poisson process. Under the events the change in the game structure is understood. Also, we introduce the notation

$$\Lambda_j(t) = \begin{cases} \frac{f_j(t)}{1 - F_j(t)}, & t \in [\max(t_0 + (j + 1)\Delta t, \bar{t}_{j-1}), T], \\ 0, & t \in [t_0 + j\Delta t, \max(t_0 + (j + 1)\Delta t, \bar{t}_{j-1})]. \end{cases}$$

The truncated information is formalized in the random truncated subgame $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t)$. The dynamical system and the initial conditions for this subgame have the following form:

$$\dot{x} = a\sqrt{x(t)} - bx(t) - \sum_{i=1}^3 u_i, \quad x(t_0 + j\Delta t) = x_{j,0}. \quad (3.125)$$

According to (3.96), the payoff function of player i is

$$K_i^j(x_{j,0}, t_0 + j\Delta t; u) = \int_{t_0 + j\Delta t}^T (1 - F_j(\tau)) h_i(x(\tau), u(\tau)) d\tau. \quad (3.126)$$

Consider the case when the resource extractors agree to act cooperatively in the random truncated subgame $\bar{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t)$. They follow the optimality principle under which they would maximize their joint payoffs and share the excess of the

total expected cooperative payoff over the sum of individual noncooperative payoffs proportional to the noncooperative payoffs.

3.3.3.3 Cooperative Trajectory

Next, consider the random truncated subgame $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t)$. The maximized joint payoff in the game $\bar{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t)$ has the following form [69]:

$$W^{j\Delta t}(t, x) = A^j(t)\sqrt{x} + C^j(t), \quad (3.127)$$

where the functions $A^j(t)$ and $C^j(t)$ satisfy the equations

$$\begin{aligned} \dot{A}^j(t) &= \left[\Lambda_j(t) + \frac{b}{2} \right] A^j(t) - \sum_{i=1}^3 \left[\frac{1}{4 \left[c_i + \frac{A^j(t)}{2} \right]} \right], \\ \dot{C}^j(t) &= \Lambda_j(t)C^j(t) - \frac{a}{2}A^j(t) \end{aligned} \quad (3.128)$$

with the boundary conditions $\lim_{t \rightarrow T^-} A^j(t) = \lim_{t \rightarrow T^-} C^j(t) = 0$.

The optimal cooperative trajectory $x_j^*(t)$ in the random truncated subgame $\bar{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t)$ can be represented explicitly [69] on the entire interval $[t_0 + j\Delta t, T]$. The trajectory with the initial condition $x = x_{j,0}^*$ is

$$x_j^*(t) = \varpi_j^2(t_0 + j\Delta t, t) \left[\sqrt{x_{j,0}^*} + \frac{1}{2}a \cdot \int_{t_0 + j\Delta t}^t \varpi_j(t_0 + j\Delta t, \tau)^{-1} d\tau \right]^2, \quad (3.129)$$

where $t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t]$,

$$\varpi_j(t_0 + j\Delta t, t) = \exp \int_{t_0 + j\Delta t}^t - \left[\frac{1}{2}b + \sum_{i=1}^3 \left[\frac{1}{4 \left[c_i + \frac{A^j(\tau)}{2} \right]^2} \right] \right] d\tau. \quad (3.130)$$

The initial conditions are defined recursively by the optimal trajectory of the previous game: $x_{0,0}^* = x_0$ and $x_{j,0}^* = x_{j-1}^*(t_0 + j\Delta t)$ for $j = 0, \dots, l$. The conditionally cooperative trajectory $\hat{x}^*(t)$ is defined within the dynamic updating

approach as

$$\hat{x}_j^*(t) = x_j^*(t), \quad t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t], \quad (3.131)$$

for $j = 0, \dots, l$.

3.3.3.4 Characteristic Function

To allocate the cooperative payoff in each random truncated subgame, it is necessary to define the characteristic function values $V_j(S; x_{j,0}, t_0 + j\Delta t)$ ($V_j(S; x_j^*(t), t)$) for each coalition $S \subset N$. According to formula (3.102), the maximized joint payoff $W_j(t_0 + j\Delta t, x_{j,0})$ (3.127) corresponds to the characteristic function value $V_j(N; x_{j,0}, t_0 + j\Delta t)$ of the grand coalition in the random truncated subgame $\bar{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t)$:

$$V_j(N; x_j^*(t), t) = W^{j\Delta t}(t, x_j^*(t)), \quad (3.132)$$

where $t \in [t_0 + j\Delta t, T]$, $j = 0, \dots, l$. Next, the characteristic function values for the following coalitions are defined:

$$\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}. \quad (3.133)$$

According to (3.102), for one player coalitions $\{i\}$, $i = \overline{1, 3}$, the Nash equilibrium and $V_j(\{i\}; x_j^*(t), t)$ are determined.

3.3.3.5 One Player Coalitions

The random truncated subgame $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t)$ has a Nash equilibrium defined by the feedback

$$u_i^j(t, x) = \frac{x}{4[c_i + A_i^j(t)/2]^2}, \quad i = \overline{1, 3}, \quad (3.134)$$

where the functions $A_i^j(t)$ are defined by the equations

$$\begin{aligned} \dot{A}_i^j(t) &= A_i^j(t) \left[\Lambda_j(t) + \frac{b}{2} + \sum_{k \neq i} \frac{1}{8(c_k + A_k^j(t)/2)^2} \right] - \frac{1}{4(c_i + A_i^j(t)/2)}, \\ \dot{C}_i^j(t) &= \Lambda_j(t) C_i^j(t) - \frac{a}{2} A_i^j(t) \end{aligned}$$

for $i = \overline{1, 3}$, with the boundary conditions $\lim_{t \rightarrow T^-} A_i^j(T) = 0$ and $\lim_{t \rightarrow T^-} C_i^j(T) = 0$.

The value function of player $i = \overline{1, 3}$ in the Nash equilibrium is

$$V_i^j(t, x) = A_i^j(t)\sqrt{x} + C_i^j(t), \quad i = \overline{1, 3}. \quad (3.135)$$

Therefore, the characteristic function value for one player coalitions $S = \{i\}$, $i \in N$, can be calculated in the following way:

$$V_j(\{i\}; x_j^*(t), t) = V_i^j(t, x_j^*(t)), \quad (3.136)$$

where $t \in [t_0 + j\Delta t, T]$, $j = 0, \dots, l$.

3.3.3.6 Two Player Coalitions

According to formula (3.102), the characteristic function $V_j(S; x_{j,0}, t_0 + j\Delta t)$ ($V_j(S; x_j^*(t), t)$) for two player coalitions $S = \{1, 2\}, \{1, 3\}, \{2, 3\}$ is defined as the total payoff of the players from coalition S in a Nash equilibrium $u_j^{NE} = (u_1^{NE,j}, u_2^{NE,j}, u_3^{NE,j})$ in the game with the following set of players: coalition S (acting as one player) and the players from the set $N \setminus S$, i.e., in the game with $|N \setminus S| + 1 = 2$ players. This means that the players from coalition S behave as one player, and other players from the set $N \setminus S$ are acting separately. Using this approach, we define the Nash equilibrium between two players: the combined player (coalition S) and the second player (coalition N/S).

Consider calculations of $V_j(S; x_{j,0}, t_0 + j\Delta t)$ in the case $S = \{1, 2\}$. Calculations for other coalitions have the same algorithm. The payoff of the players in this case has the following form:

$$\begin{aligned} V_{\{1,2\}}^j(t, x) &= A_{\{1,2\}}^j(t)\sqrt{x} + C_{\{1,2\}}^j(t), \\ V_3^j(t, x) &= A_3^j(t)\sqrt{x} + C_3^j(t), \end{aligned}$$

where the functions $A_{\{1,2\}}^j(t)$, $A_3^j(t)$, $C_{\{1,2\}}^j(t)$, and $C_3^j(t)$ satisfy the equations

$$\begin{aligned}\dot{A}_{\{1,2\}}^j(t) &= A_{\{1,2\}}^j(t) \left[\Lambda_j(t) + \frac{b}{2} + \frac{1}{8(c_3 + A_3^j(t)/2)^2} \right] - \sum_{k \in S} \frac{1}{4(c_k + A_{\{1,2\}}^j(t)/2)}, \\ \dot{A}_3^j(t) &= A_3^j(t) \left[\Lambda_j(t) + \frac{b}{2} + \sum_{k \in S} \frac{1}{8(c_k + A_{\{1,2\}}^j(t)/2)^2} \right] - \frac{1}{4(c_3 + A_3^j(t)/2)}, \\ \dot{C}_{\{1,2\}}^j(t) &= \Lambda_j(t) C_{\{1,2\}}^j(t) - \frac{a}{2} A_{\{1,2\}}^j(t), \\ \dot{C}_3^j(t) &= \Lambda_j(t) C_3^j(t) - \frac{a}{2} A_3^j(t)\end{aligned}$$

with the initial conditions $\lim_{t \rightarrow T^-} A_{\{1,2\}}^j(t) = \lim_{t \rightarrow T^-} A_3^j(t) = 0$ and $\lim_{t \rightarrow T^-} C_{\{1,2\}}^j(t) = \lim_{t \rightarrow T^-} C_3^j(t) = 0$.

Therefore, the characteristic function value for coalition $S = \{1, 2\}$ can be calculated in the following way:

$$V_j(\{1, 2\}; x_j^*(t), t) = V_{\{1,2\}}^j(t, x_j^*(t)), \quad (3.137)$$

where $t \in [t_0 + j\Delta t, T]$, $j = 0, \dots, l$.

3.3.3.7 IDP-core

Using the characteristic function values $V_j(S; x_{j,0}, t_0 + j\Delta t)$, $\forall S \subset N$, (3.132), (3.136), (3.137), and formula (3.108), we construct the set $B_j(t, x_j^*)$ as a set of

integrable vector functions $\beta_j(t, x_j^*)$ satisfying the following relations:

$$\begin{aligned}
\sum_{i=1}^3 (1 - F_j(t)) \beta_i^j(t, x_j^*) &= -\frac{d}{dt} V_j(\{1, 2, 3\}; x_j^*(t), t), \\
(1 - F_j(t)) (\beta_1^j(t, x_j^*) + \beta_2^j(t, x_j^*)) &\geq -\frac{d}{dt} V_j(\{1, 2\}; x_j^*(t), t), \\
(1 - F_j(t)) (\beta_1^j(t, x_j^*) + \beta_3^j(t, x_j^*)) &\geq -\frac{d}{dt} V_j(\{1, 3\}; x_j^*(t), t), \\
(1 - F_j(t)) (\beta_2^j(t, x_j^*) + \beta_3^j(t, x_j^*)) &\geq -\frac{d}{dt} V_j(\{2, 3\}; x_j^*(t), t), \\
(1 - F_j(t)) \beta_1^j(t, x_j^*) &\geq -\frac{d}{dt} V_j(\{1\}; x_j^*(t), t), \\
(1 - F_j(t)) \beta_2^j(t, x_j^*) &\geq -\frac{d}{dt} V_j(\{2\}; x_j^*(t), t), \\
(1 - F_j(t)) \beta_3^j(t, x_j^*) &\geq -\frac{d}{dt} V_j(\{3\}; x_j^*(t), t).
\end{aligned} \tag{3.138}$$

Then, combining the sets $B_j(t, x_j^*)$, $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, l$ for all random truncated subgames, we construct the set $\hat{B}(t, \hat{x}^*)$. Further, we calculate the set of all possible imputations $\hat{\xi}(x_0, T - t_0) \in \hat{W}(x_0, T - t_0)$ (3.112).

The step-by-step construction of the IDP-core for a linear-quadratic game model of pollution control was presented in the paper [100].

3.3.3.8 Numerical Example

Consider a numerical example in which information about the game structure during the time intervals $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$ is known for the next time interval of a length \bar{T}_j , where \bar{T}_j is a random variable obeying the distribution (3.123) with $\lambda = 0.5$. The total game length is $T = 4$. Information about the game is updated every $\Delta t = 1$ time interval. The dynamical system has the following parameters: $a = 10$ and $b = 0.5$. Assume that $c_1 = 0.15$, $c_2 = 0.65$, and $c_3 = 0.45$ in the payoff function, and the initial conditions are $t_0 = 0$, $x_0 = 200$. During the realization of the game, the information horizon takes the following values:

$$\bar{t}_0 = 2.423, \quad \bar{t}_1 = 3.538, \quad \bar{t}_2 = 3.871, \quad \bar{t}_3 = 4.$$

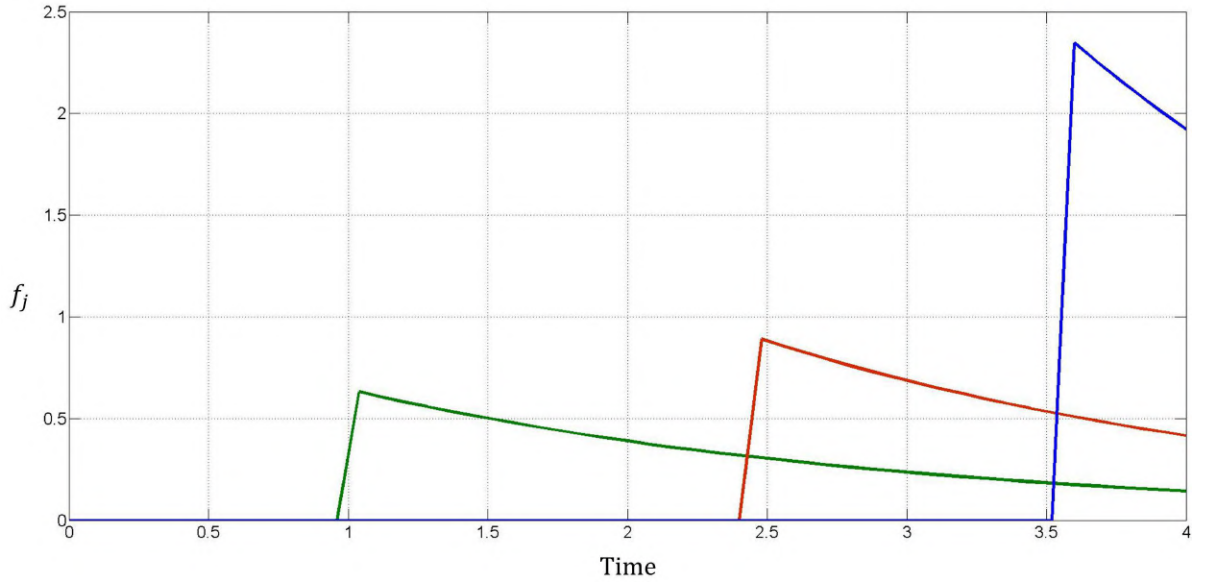


Figure 3.10: Probability density function $f_j(t)$, $j = 0, 1, 2, 3$ (3.124) for each random truncated subgame.

The generated values of the information horizon affect the distribution of the time during which the available truncated information is correct. Figure 3.10 illustrates how the information horizon \bar{T}_j is generated and the probability density function $f_j(t)$ (3.124) changes between the random truncated subgames.

Figure 3.11, 3.12, 3.13 shows the cooperative strategies for each player defined using the dynamic updating approach with random horizon (non-smooth solid line) and the cooperative strategies in the original game [69] (smooth dotted line).

The conditionally cooperative trajectory $\hat{x}^*(t)$ is composed from the solutions of the random truncated subgames $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t)$ for the dynamical system (3.125). In Fig. 3.14 the following comparison is presented: the conditionally cooperative trajectory $\hat{x}^*(t)$ (thick solid line) defined using the dynamic updating approach with random horizon, the conditionally cooperative trajectory $\bar{x}^*(t)$ (thin solid line) defined using the classical approach of dynamic updating [76] (where $\bar{T}_j = 2$ is a deterministic value), and the cooperative trajectory $x^*(t)$ (dotted line) in the original game $\Gamma(x_0, T - t_0)$. The cooperative trajectory $x^*(t)$ was defined in [69]. The other two figures show the conditionally cooperative trajectory $\hat{x}^*(t)$ or $\bar{x}^*(t)$ and the corresponding cooperative trajectories in each truncated subgame.

Next, to allocate the cooperative payoff among the players, it is necessary to define a set of IDPs $\beta_j(t, x_j^*)$ for each random truncated subgame $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t)$,

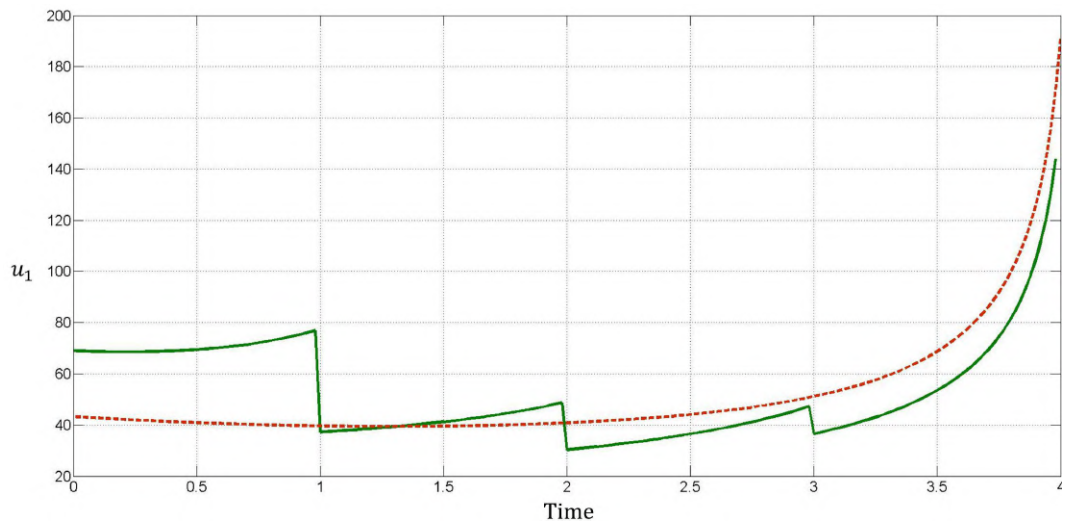


Figure 3.11: Cooperative strategies of player 1 under dynamic updating with random horizon (non-smooth solid line) and cooperative strategies in original game [69] (smooth dotted line).

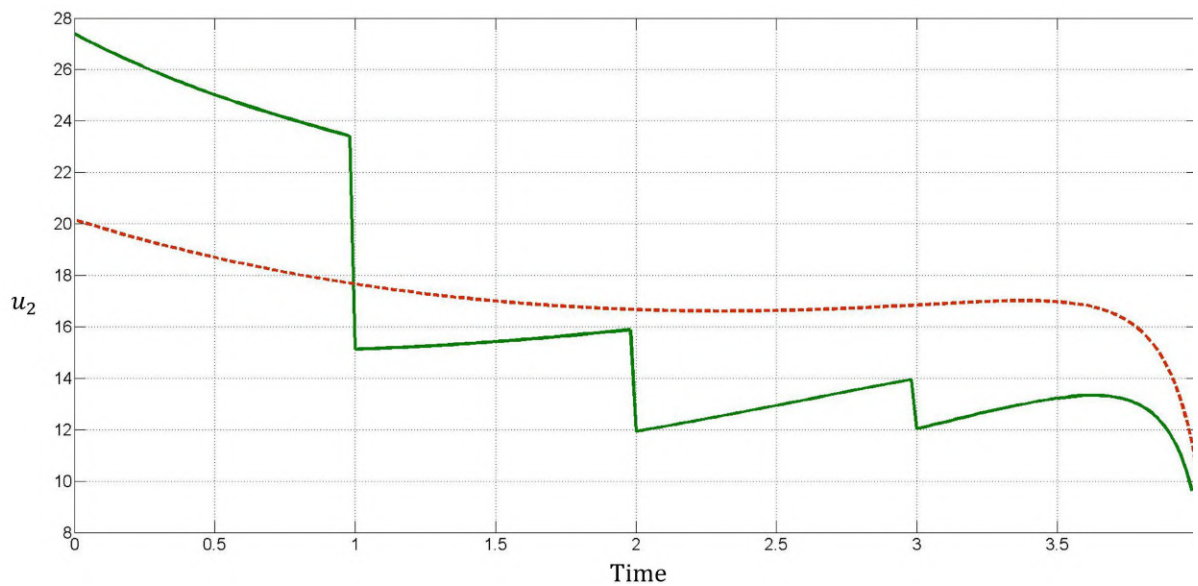


Figure 3.12: Cooperative strategies of player 2 under dynamic updating with random horizon (non-smooth solid line) and cooperative strategies in original game [69] (smooth dotted line).

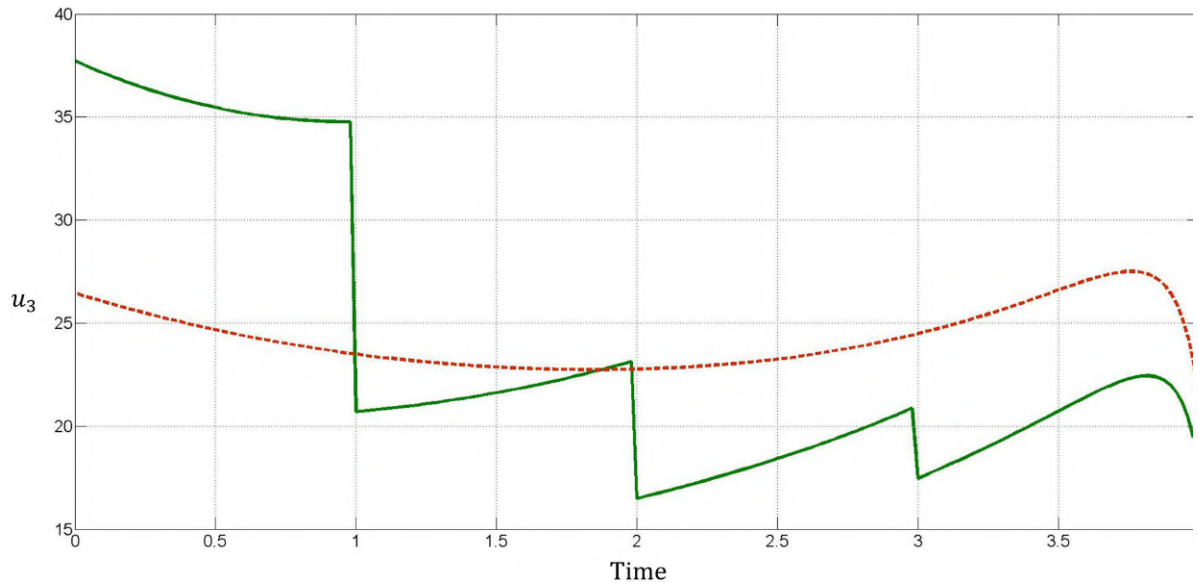


Figure 3.13: Cooperative strategies of player 3 under dynamic updating with random horizon (non-smooth solid line) and cooperative strategies in original game in [69] (smooth dotted line).

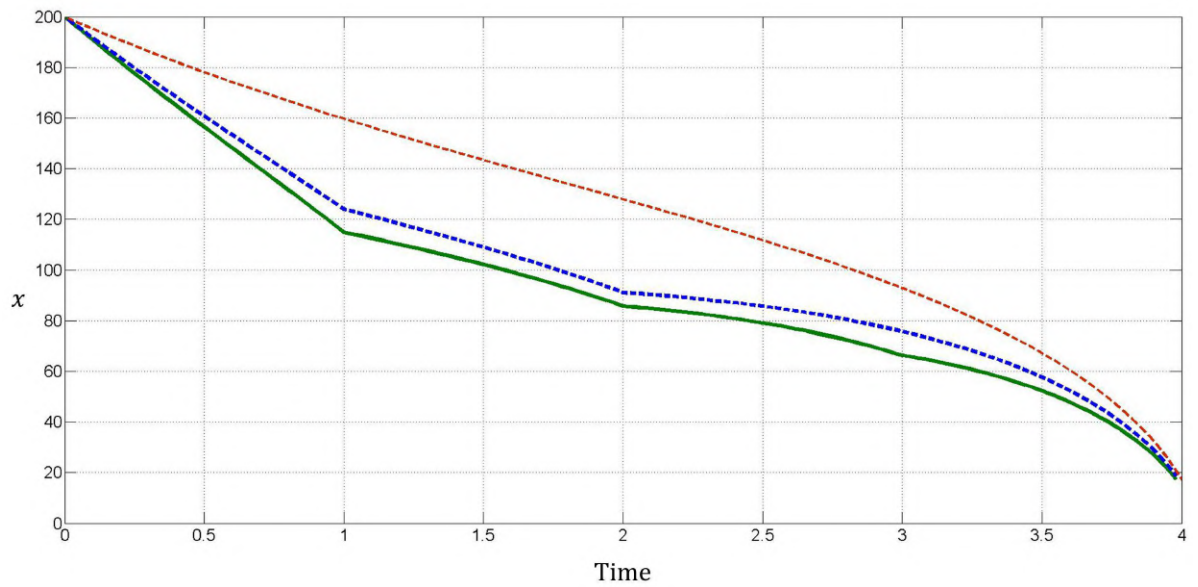


Figure 3.14: Resource stock trajectory $\hat{x}^*(t)$ (thick solid line) under dynamic updating with random horizon, trajectory $\bar{x}^*(t)$ (thick dotted line) under classical dynamic updating, and cooperative trajectory $x^*(t)$ (thin dotted line) in original game $\Gamma(x_0, T - t_0)$.

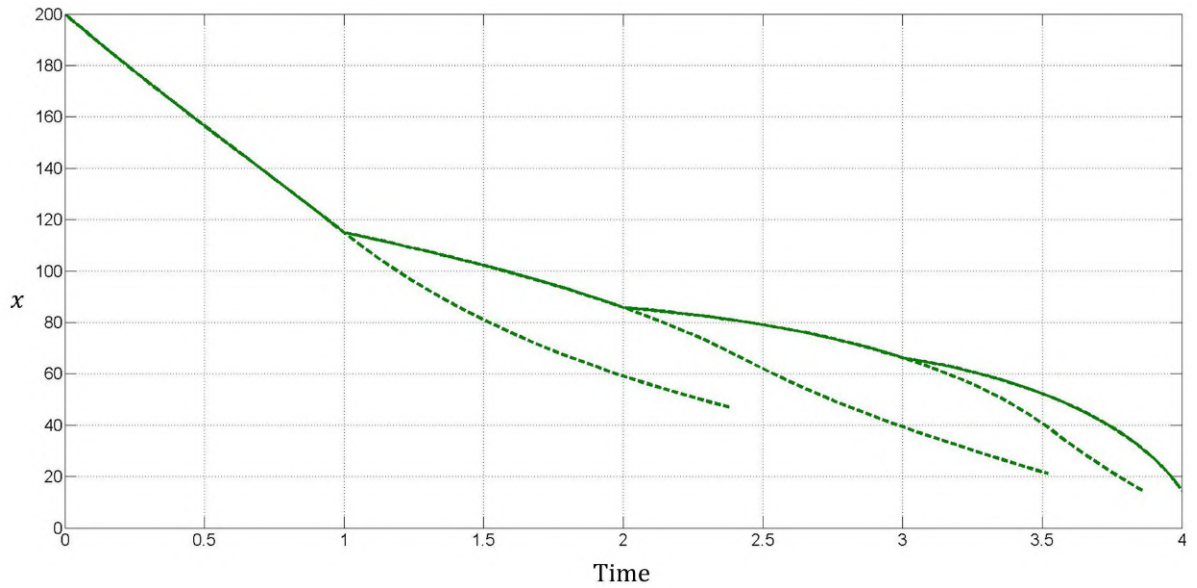


Figure 3.15: Resource stock trajectory $\hat{x}^*(t)$ (thick solid line) under dynamic updating with random horizon and corresponding cooperative trajectories (dotted lines).

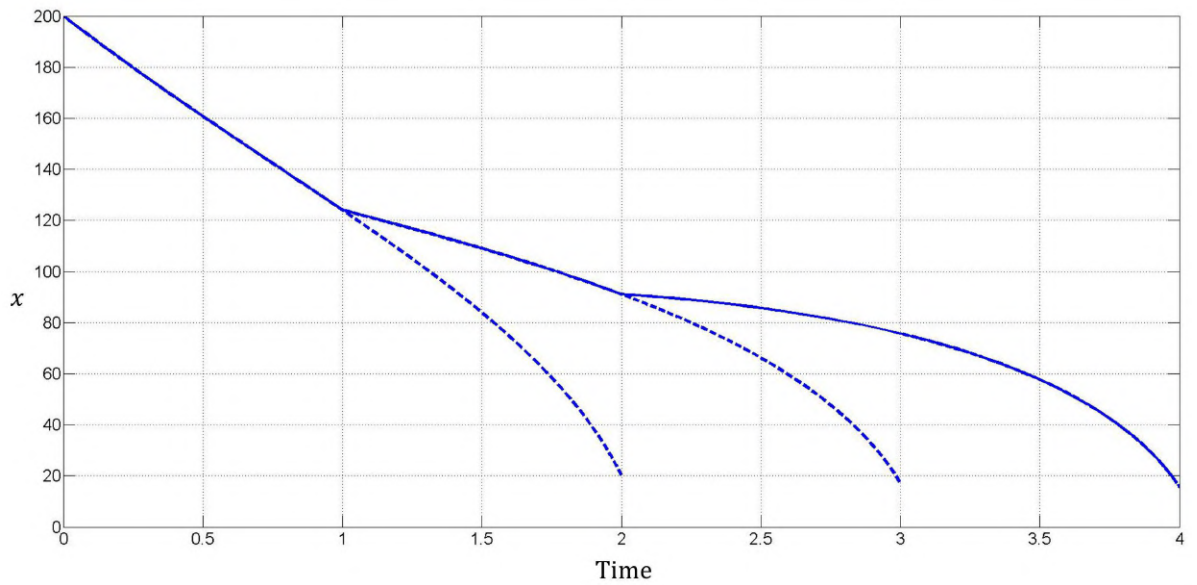


Figure 3.16: Resource stock trajectory $\bar{x}^*(t)$ (thick solid line) under classical dynamic updating and corresponding cooperative trajectories (dotted line).

$j = 0, \dots, l$. For this purpose, we numerically calculate the characteristic function values $V_j(S; x_j^*(t), t)$, $S \subset N$, for each random truncated subgame $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t)$ under the fixed parameters of the model.

Using the characteristic function values $V_j(S; x_j^*(t), t)$, $S \subset N$, we construct the set $B_j(t, x_j^*)$, $j = 0, \dots, l$ (3.108). Combining the sets $B_j(t, x_j^*)$, we construct the set of IDPs for the entire game $\hat{B}(t, \hat{x}^*)$. On the basis of $\hat{B}(t, \hat{x}^*)$, the solution concept

$\hat{W}(x_0, T - t_0)$ is constructed using formula (3.112).

Let us demonstrate the strong time-consistency of the solution concept \hat{W} . Suppose that at the beginning of the game $\Gamma(x_0, T - t_0)$, the players agreed to use the proportional solution. For each random truncated subgame $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t)$, the proportional solution for players $i \in N$ is defined using the IDP in the following way:

$$(1 - F_j(t))\beta_{i,j}^{Prop}(t, x_j^*) = \frac{U_j(\{i\}; x_{j,0}^*, t_0 + j\Delta t)}{\sum_{i \in N} U_j(\{i\}; x_{j,0}^*, t_0 + j\Delta t)} U_j(N; x_{j,0}^*, t_0 + j\Delta t), \quad (3.139)$$

where $U_j(S; x_{j,0}^*, t_0 + j\Delta t)$, $\forall S \subset N$ is defined in (3.107). According to the proportional solution with dynamic updating, the players should allocate the cooperative payoff during the entire game $\Gamma(x_0, T - t_0)$ using the following IDP:

$$\hat{\beta}_{Prop}(t, \hat{x}^*) = (1 - F_j(t))\beta_j^{Prop}(t, x_j^*), \quad t \in [j\Delta t, (j+1)\Delta t], j = 0, \dots, l. \quad (3.140)$$

Integrating $\hat{\beta}_{Prop}(t, \hat{x}^*)$ with respect to t , we define the proportional imputation $\hat{\xi}_{Prop}(\hat{x}^*(t), T - t)$ (3.112). In Fig. 3.18, 3.19 it can be seen that $\hat{\beta}_{Prop}(t, \hat{x}^*)$ belongs to the set $\hat{B}(t, \hat{x}^*)$, meaning that the proportional solution is strong time-consistent with the given parameters.

Suppose that at a time instant $t_{br} \in [t_0, T]$, the players decided that the proportional solution is no longer fair for them, and they chose another imputation from the solution concept $\hat{W}(\hat{x}^*(t_{br}), T - t_{br})$, e.g., the one based on the Shapley value for each random truncated subgame. In each random truncated subgame, the Shapley value is calculated as

$$Sh_i^j(\hat{x}^*(t_{br}), t_{br}) = \sum_{\substack{S \subset N \\ i \in S}} \frac{(|N| - |S|)! (|S| - 1)!}{|N|!} \cdot \left(V_j(S; t_{br}, \hat{x}^*(t_{br})) - V_j(S \setminus \{i\}; t_{br}, \hat{x}^*(t_{br})) \right). \quad (3.141)$$

Using the Shapley value, we define an IDP for each random truncated subgame (3.105). According to the proportional solution with dynamic updating, the players should allocate the cooperative payoff during the entire game using the following

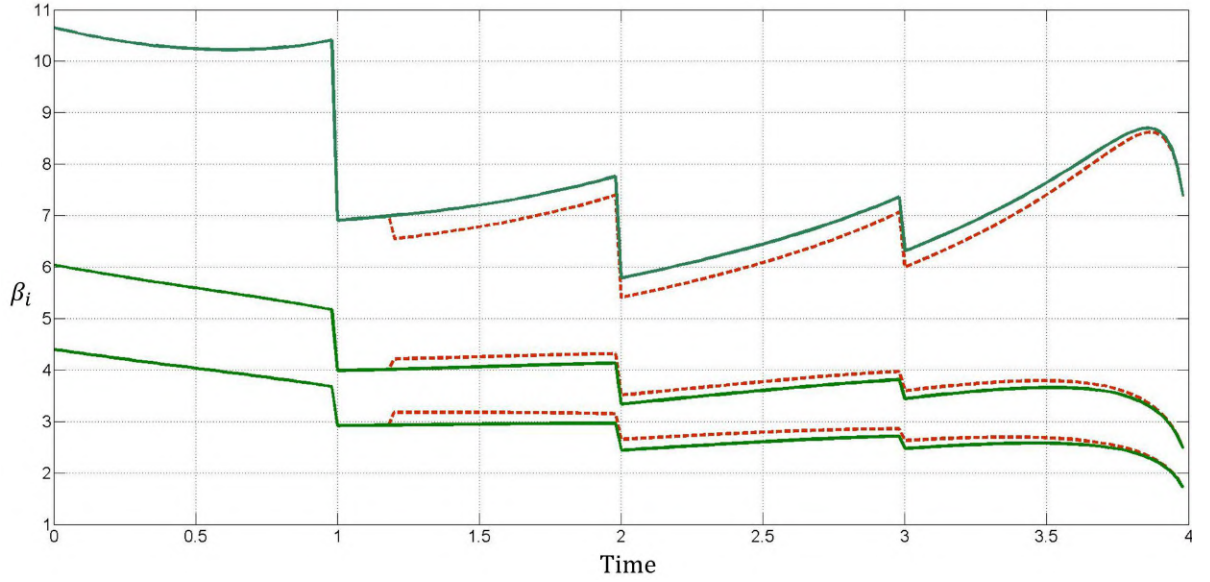


Figure 3.17: IDP $\hat{\beta}_{Prop}(t, \hat{x}^*)$ for proportional solution (3.139) (thick solid line) and IDP $\hat{\beta}(t, \hat{x}^*)$ for combined solution (3.142) (dotted line).

IDP:

$$\hat{\beta}_{Sh}(t, \hat{x}^*) = (1 - F_j(t))\beta_j^{Sh}(t, x_j^*), \quad t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t], \quad j = 0, \dots, l,$$

where $\beta_j^{Sh}(t, x_j^*)$ is defined by formula (3.106). Note that the IDPs for the Shapley value $\beta_j^{Sh}(t, x_j^*)$ and proportional solution $\beta_j^{Prop}(t, x_j^*)$ are calculated in the way to ensure the time-consistency property, i.e., using formula (3.106). A detailed description of the step-by-step solution of the IDP for the Shapley value was presented in [194].

Let $t_{br} = 1.2$ be the time instant when the players decided to reconsider the proportional solution. Then according to (3.111), the IDP for the entire game has the following form:

$$\hat{\beta}(t, \hat{x}^*) = \begin{cases} \hat{\beta}_{Prop}(t, \hat{x}^*), & t \in [t_0, t_{br}], \\ \hat{\beta}_{Sh}(t, \hat{x}^*), & t \in (t_{br}, T]. \end{cases} \quad (3.142)$$

In Fig. 3.17 the IDP $\hat{\beta}_{Prop}(t, \hat{x}^*)$ for the proportional solution (3.140) (thick solid line) and the IDP $\hat{\beta}(t, \hat{x}^*)$ for the combined solution (3.142) (dotted line) are presented.

Directly integrating $\hat{\beta}(t, \hat{x}^*)$ (3.142) with respect to t , we define formula (3.112)

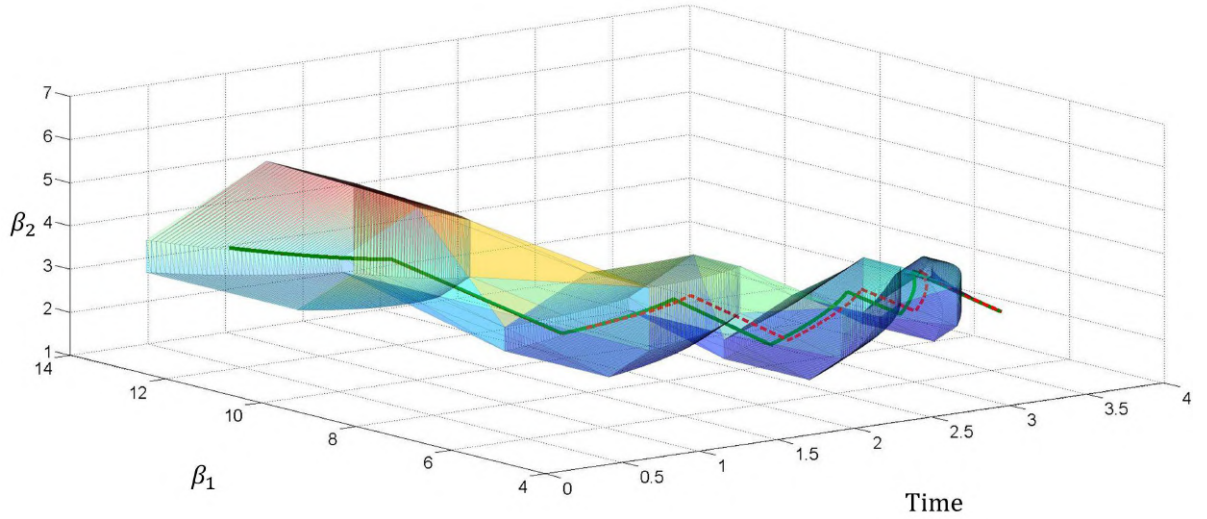


Figure 3.18: Axis: β_1, β_2, t . β_3 can be calculated using (3.108).

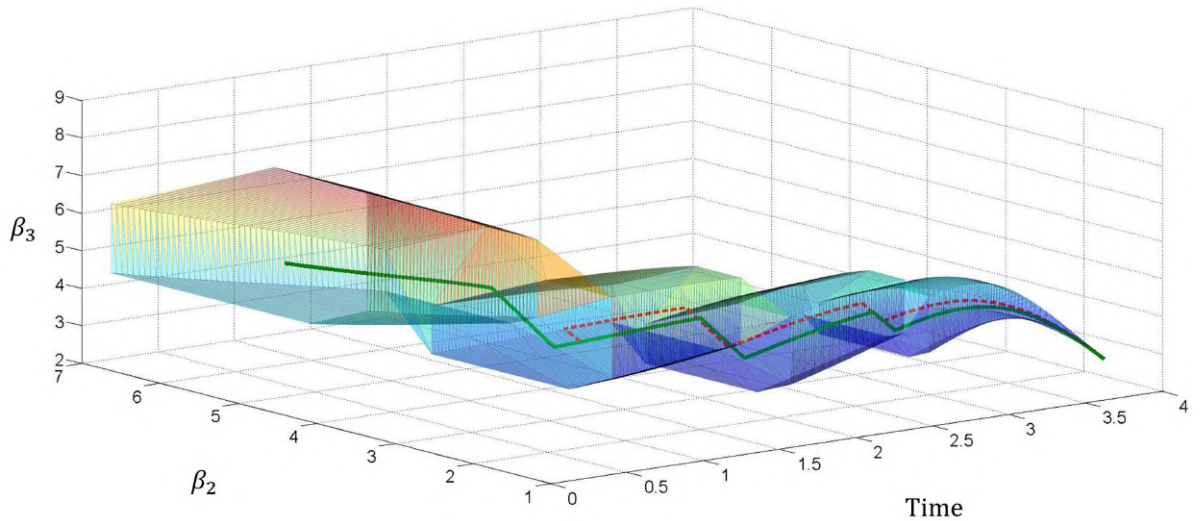


Figure 3.19: Axis: β_2, β_3, t . β_1 can be calculated using (3.108).

for the resulting allocation $\hat{\xi}(\hat{x}^*(t), T-t)$. According to $\hat{\xi}(\hat{x}^*(t), T-t)$, the players allocate the cooperative payoff in the game $\Gamma(x_0, T-t_0)$ in the following way:

$$\hat{\xi}(\hat{x}^*(t), T-t) = (12.3, 30.2, 16.8). \quad (3.143)$$

In Fig. 3.18, 3.19 it can be seen that $\hat{\beta}(t, \hat{x}^*)$ (3.142) belongs to the set $\hat{B}(t, \hat{x}^*)$ (3.108), meaning that the corresponding imputation $\hat{\xi}(\hat{x}^*(t), T-t) \in \hat{W}(\hat{x}^*(t), T-t)$ with the given parameters. This fact demonstrates the strong time-consistency of the solution concept $\hat{W}(x_0, T-t_0)$. Also, in Fig. 3.18, 3.19 it can be seen that the proportional solution $\hat{\beta}_{Prop}(t, \hat{x}^*)$ (3.140) belongs to the set $\hat{B}(t, \hat{x}^*)$.

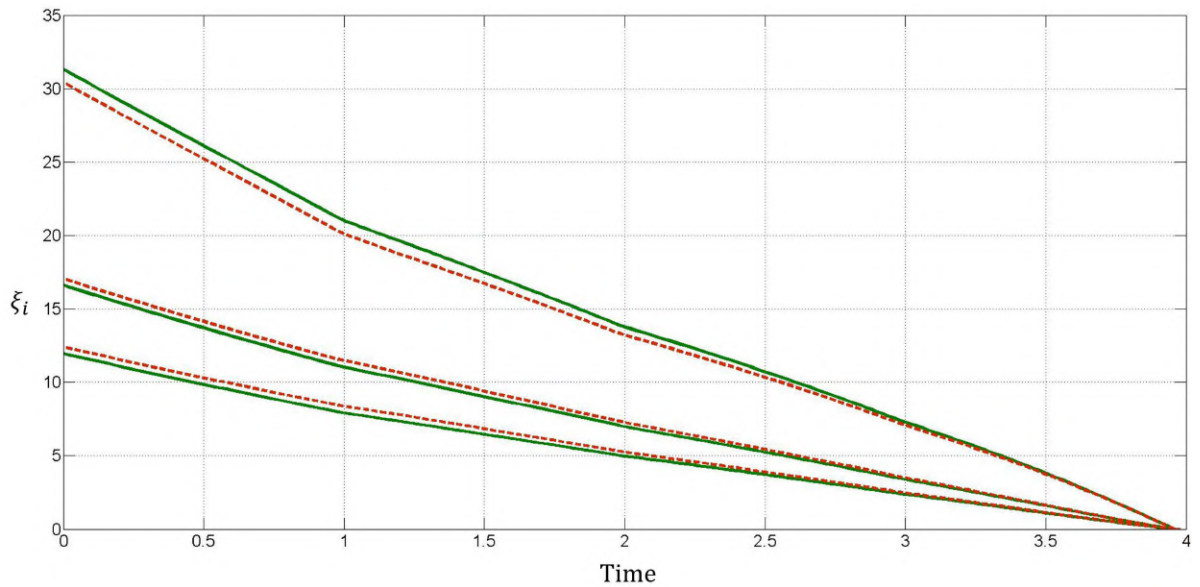


Figure 3.20: Imputation $\hat{\xi}_{Prop}(\hat{x}^*(t), T - t)$ for proportional solution (thick solid line) and imputation $\hat{\xi}(\hat{x}^*(t), T - t)$ for combined solution (dotted line).

In Fig. 3.20 the difference between the $\hat{\xi}(\hat{x}^*(t), T - t)$ and $\hat{\xi}_{Prop}(\hat{x}^*(t), T - t)$ is presented.

3.4 Differential Game Model of Oil Market with Dynamic Updating

3.4.1 Introduction

This section is devoted to constructing a game-theoretic model for the world oil market using the dynamic updating approach. Game models with dynamic updating allow taking into account the variability of market demands, an adaptation of participants' actions to the changing environment, and actual planning horizons for demand. The oil market has highly volatile prices; therefore, the dynamic updating approach is applied to the oil market, and the resulting model is studied. It is supposed that countries do not have or cannot use long-term forecasts for the parameters of the oil market. Therefore, their behavior can be modeled using the approach proposed.

This section aims to simulate the oil market dynamics during two particular periods, from December 2015 to November 2016 and from December 2016 to May

2017. The OPEC countries and eleven non-OPEC countries reached a six-month-long agreement about reducing oil production for the latter period at the summit in Vienna on November 30, 2016. The agreement was aimed to result in the growth of oil prices. These countries, taken together, account for over 60% of oil production in the world. To show the effect of the agreement, we consider two game-theoretic models corresponding to the two periods mentioned. It is supposed that the countries did not cooperate officially on the amounts of oil to be produced before November 30, 2016. Therefore, their behavior can be simulated using a non-cooperative game model. Hence, first, we construct a non-cooperative game for the time interval from December 2015 to November 2016; then, we adapt this model to the real oil price data. As an optimality principle, the feedback Nash equilibrium is used. Making an agreement on the quantity of oil production seems to be a case of purely cooperative behavior. Therefore, a coalitional game model is constructed for the time interval from December 2016 to May 2017. The cooperation is modeled using a non-transferable utility game model. Subsequently, a classical cooperative solution for this type of games is constructed, i.e., the subgame consistent solution proposed in [165]. Furthermore, the cooperative agreement is modeled which was actually used by the countries (the historical solution). According to the historical agreement, the solution is based on the oil production quantities in November 2016. It prescribes players 1 and 2 to linearly lower the quantities of oil production to the level agreed previously. Obviously, the second solution is not subgame consistent, though it allows us to model the latest real-life market agreement.

We present statistical simulation results and compare the historical data with the model data over both periods. In particular, a comparison of the theoretical oil price trajectory is given between December 2015 and November 2016 with the statistical data. Over the period between December 2016 and May 2017, the oil price trajectory corresponding to the statistical data, the subgame consistent cooperative solution defined in [165], and the cooperative solution corresponding to the agreement signed by the largest oil exporters at the end of 2016 are compared. To define the parameters of the model, open source data on the world oil market is used. The International Energy Agency for monthly data on crude oil supply are used from January 2015 to May 2017, the Finam agency for monthly data on Brent and light oil prices from January 2015 to May 2017, Rystad Energy Ucube, and oil market news for the cost of producing a barrel of oil in 2016.

The largest oil exporters reached the agreement to reduce oil production, aiming at raising prices after November 30, 2016. In all our models, the exporters are considered in groups of five to one players. This is not the first time the OPEC has restricted oil extraction hoping to control oil prices. These actions of the members seem rational in a sense, as oil prices started to fall steadily in 2014. Thus, in our game models, all the members of the OPEC together are called player 1. The other eleven countries that signed the agreement clearly have common interests, although with more freedom in decision-making than the OPEC countries; they are called player 2. At the same time, the steady fall in oil prices was largely due to the US shale oil and gas revolution, which started in 2012. Therefore, our models include US shale and non-shale oil producing companies as the main market rivals, and they are players 3 and 4, respectively. All the other oil-exporting countries constitute player 5. According to the described rule, the countries are combined into players from 1 to 5 since in our models only the price change process is considered. Furthermore, such a combination greatly simplifies the computational process without exerting any serious effect on the plausibility of the results.

The global oil market is one of the most significant markets and a crucial component in growth rate and budget scenarios for some resource-dependent countries. In all fairness, this has been given detailed coverage in literature. Most of the existing models consider the oil market as a market with imperfect competition. Market power models for the oil market explain the OPEC's presence by cartel behavior, by a dominant firm, or by target behavior in most cases. Thus, Dahl and Yucel in [201] described OPEC behavior as loose coordination or duopoly. In the paper [202], the pattern of extraction in the oil market was shown to be inconsistent with either the patterns predicted by the competitive theory or the dominant firm-competitive fringe theory. Danielsen and Kim [203] provided significant evidence of cooperation among the OPEC countries. Smith in [204] asserted that OPEC is much more than simply a non-cooperative oligopoly but less than a frictionless cartel. Youhanna in [205] affirmed that a partial market sharing cartel model dominates over all other models. Gulen in [206] provided evidence of output coordination and suggested that OPEC acted like a cartel in the 1980s (1982-1993). The author of [207] underlined that the OPEC behavior is consistent with the cartel theory. Bockem in [208] stated that the crude oil market is best described as a price leader model, with the OPEC appearing to be the leader and all non-OPEC countries being

regarded as price takers. The paper [209] claimed that OPEC behaves more like an oligopoly, with Saudi Arabia as a price leader and the largest producer. In the paper [210], the author asserted that the Nash–Cournot non-cooperative model can potentially explain the oil market better than the competitive one. Moran in [211] and Krasner in [212] analyzed the main features of oil oligopoly. The authors in [213] gave a selective survey of oligopoly models for energy production. In the paper [214], the author examined cartel formation in the world oil market under the Cournot setting. The authors in [215] also used the quantity competitive environment to model collisions and proportionate adjustment of production levels. Following this well-established tradition, the oligopoly quantity setting is used below to model the oil market.

In the first part of this section, the dynamic updating approach is applied to the non-cooperative oligopoly differential model [216] of the oil market with the largest oil exporters and other oil-producing countries. In the second part of the section, a partially cooperative differential game with non-transferable utility is considered, which uses the non-cooperative game described above as the original model. It is assumed that players 1 and 2 cooperate, and the rest use the Nash equilibrium strategies against them.

3.4.2 Non-cooperative Game Model

3.4.2.1 Original Game Model

Let us consider a differential game model of Cournot oligopoly [216] on the oil market. An oligopolistic market of n asymmetrical countries (players) from the set $N = \{1, \dots, n\}$, which produce oil and compete for the quantity produced q_i under price stickiness, is given by the differential game $\Gamma(p_0, T - t_0)$ with prescribed duration $T - t_0$ and an initial state $p(t_0) = p_0 \in P \subset R$.

Within the model, the market price p_i evolves according to the differential equation

$$\dot{p}(t) = s(\hat{p}(t) - p(t)), \quad p(t_0) = p_0, \quad (3.144)$$

where $\hat{p}(t) \in P \subset R$ is the notional level of the price at a time instant t , $p(t)$ is its current level, and the parameter $s : 0 < s < 1$ is the speed of adjustment. Thus,

the prices adjust to the difference between the notional and current levels.

Further, it is assumed that the notional prices at any time instant t are defined by the linear inverse demand function

$$\hat{p}(t) = a - d \sum_{i \in N} q_i(t). \quad (3.145)$$

Each player $i \in N$ chooses a quantity $q_i(t) \in U_i \subset R$ produced to maximize its profit:

$$K_i(p_0, T - t_0; q_1, \dots, q_n) = \int_{t_0}^T e^{-\rho(t-t_0)} [q_i(t)(p(t) - c_i - g_i q_i(t))] dt, \quad (3.146)$$

where $0 \leq \rho \leq 1$ is a positive discount rate, taken the same for all the periods and all the players to simplify the model and equalize the players as symmetrical participants in the global capital market. Also, $C_i(t) = c_i q_i(t) + g_i q_i^2(t)$ is the total cost function of each player i .

3.4.2.2 Truncated Subgame

Suppose that the information for the players is updated at fixed time instants $t = t_0 + j\Delta t$, $j = 0, \dots, l$, where $l = \frac{T}{\Delta t} - 1$. During the time interval $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, the players have certain information about the game dynamics (3.144) and payoff function (3.148) within the time interval $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, where $\Delta t \leq \bar{T} \leq T$. At the time instant $t = t_0 + (j+1)\Delta t$, the information about the game is updated, and the same procedure repeats for the time interval with number $j+1$.

To model this kind of behavior, the following definition is introduced, where $p_{j,0} = p(t_0 + j\Delta t)$ and $p_{j,1} = p(t_0 + (j+1)\Delta t)$.

Definition 3.4.1 *Let $j = 0, \dots, l$. A truncated subgame $\bar{\Gamma}_j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ is defined on the time interval $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$. The motion equation and the initial condition of the truncated subgame $\bar{\Gamma}_j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ have the following form:*

$$\dot{p}(t) = s \left(a_j - d_j \sum_{i \in N} q_i^j(t) - p(t) \right), \quad p(t_0 + j\Delta t) = p_{j,0}. \quad (3.147)$$

The payoff function of player i in truncated subgame j is

$$\begin{aligned} K_i^j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^j, \dots, q_n^j) &= \\ &= \int_{t_0 + j\Delta t}^{t_0 + j\Delta t + \bar{T}} e^{-\rho(t-t_0)} \left[q_i^j(t)(p(t) - c_i - g_i q_i^j(t)) \right] dt. \end{aligned} \quad (3.148)$$

The motion equation and the payoff function on the time interval $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ coincide with those of the game $\Gamma(p_0, T - t_0)$ on the same time interval.

3.4.2.3 Non-cooperative Outcome in Truncated Subgame

According to [1; 217], a non-cooperative Nash equilibrium solution of the game $\bar{\Gamma}_j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ can be defined by the Hamilton-Jacobi-Bellman partial differential equations. Consider a family of subgames $\bar{\Gamma}_j(p(t), t, t_0 + j\Delta t + \bar{T})$ with the payoff structure (3.146) and the dynamics (3.147), starting at a time instant $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ from an initial state $p(t)$. For $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, we denote by $q_j^{NE}(t, p) = (q_1^{NE}(t, p), \dots, q_n^{NE}(t, p))$ the set of feedback strategies that constitutes a Nash equilibrium solution of the game $\bar{\Gamma}_j(p(t), t, t_0 + j\Delta t + \bar{T})$. Also, we denote by $V_i^j(\tau, p) : [t, T] \times R^n \rightarrow R$ the value function of player $i \in N$ that satisfies the corresponding Hamilton-Jacobi-Bellman equations [1; 217].

Theorem 3.4.1 *Assume that there exists a continuously differentiable function $V_i^j(t, p) : [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}] \times R \rightarrow R$ satisfying the partial differential equation*

$$\begin{aligned} -V_t^{j,i}(t, p) &= \max_{q_i^j} \left\{ e^{-\rho(t-t_0)} \left[q_i^j(p - c_i - g_i q_i^j) \right] + \right. \\ &\quad \left. + V_p^{j,i}(t, p) s \left(a - d \left[q_i^j(t) + \sum_{k \neq i} q_k^{NE}(t) \right] - p(t) \right) \right\}, \quad i = 1, \dots, n. \end{aligned} \quad (3.149)$$

where $V_i^j(t_0 + j\Delta t + \bar{T}, p) = 0$. Let $q_j^{NE}(t)(t, p)$ be the controls maximizing the right-hand side of (3.149). Then $q_j^{NE}(t)(t, p)$ provides a feedback Nash equilibrium in the truncated subgame $\bar{\Gamma}_j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$.

Since the considered differential game is an LQ differential game, the feedback

Nash equilibrium is unique [217].

In this game model, the Bellman function $V_i^j(t, p)$ can be obtained in the form

$$V_i^j(t, p) = e^{-\rho(t-t_0)} \left[A_i^j(t)p^2 + B_i^j(t)p + C_i^j(t) \right], \quad i = \overline{1, n}. \quad (3.150)$$

Substituting (3.150) into (3.149), we determine the Nash equilibrium strategies

$$q_i^{jNE}(t, p) = -\frac{(c_i - p) + sd \left[B_i^j(t) + 2A_i^j(t)p \right]}{2g_i}, \quad i = \overline{1, n}, \quad (3.151)$$

where the functions $A_i^j(t)$, $B_i^j(t)$, and $C_i^j(t)$, $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, are defined by the system of differential equations

$$\begin{aligned} \dot{A}_i^j(t) &= A_i^j(t) [\rho + 2s] + \frac{(2A_i^j(t)d_j s - 1)^2}{4g_i} - \sum_{k \neq i} \frac{A_i^j(t)d_j s - 2A_i^j(t)A_k^j(t)d_j^2 s^2}{g_k} \\ \dot{B}_i^j(t) &= B_i^j(t) [\rho + s] - \frac{c_i}{2g_i} - 2A_i^j(t)a_j s - \sum_{k \neq i} \frac{A_i^j(t)B_k^j(t)d_j^2 s^2}{g_k} - \\ &\quad - \sum_{k \in N} \frac{B_i^j(t)d_j s - A_i^j(t)c_k d_j s - A_k^j(t)B_i^j(t)d_j^2 s^2}{g_k} \\ \dot{C}_i^j(t) &= C_i^j(t)\rho - B_i^j(t)a_j s + \frac{c_i^2 + (B_i^j(t)d_j s)^2}{4g_i} + \sum_{k \neq i} \frac{B_i^j(t)B_k^j(t)d_j^2 s^2}{2g_k} + \\ &\quad + \sum_{k \in N} \frac{B_i^j(t)c_k d_j s}{2g_k} \end{aligned}$$

with the boundary conditions $A_i^j(t_0 + j\Delta t + \bar{T}) = 0$, $B_i^j(t_0 + j\Delta t + \bar{T}) = 0$ and $C_i^j(t_0 + j\Delta t + \bar{T}) = 0$.

Substituting $q_i^{jNE}(t, p)$ (3.151) into (3.147) yields the following dynamics of a Nash equilibrium trajectory:

$$\dot{p}(t) = s \left(a - d \sum_{i \in N} q_i^{jNE}(t, p) - p(t) \right), \quad p(t_0 + j\Delta t) = p_{j,0}. \quad (3.152)$$

Let $p_j^{NE}(t)$ be the solution of the system (3.152).

3.4.2.4 Conditionally Non-cooperative Trajectory

Suppose that each truncated subgame $\bar{\Gamma}_j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ evolves along $p_j^{NE}(t)$. Then the entire non-cooperative game with dynamic updating evolves along the following trajectory.

Definition 3.4.2 *The conditionally non-cooperative trajectory $\{\hat{p}_{NE}(t)\}_{t=t_0}^T$ is a combination of $p_j^{NE}(t)$ in each truncated subgame $\bar{\Gamma}_j(p_{j,0}^{NE}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$:*

$$\{\hat{p}_{NE}^*(t)\}_{t=t_0}^T = \begin{cases} p_0^{NE}(t), & t \in [t_0, t_0 + \Delta t), \\ \dots, \\ p_j^{NE}(t), & t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t), \\ \dots, \\ p_l^{NE}(t), & t \in [t_0 + l\Delta t, t_0 + (l+1)\Delta t]. \end{cases} \quad (3.153)$$

Along the conditionally non-cooperative trajectory the players obtain the payoff according to the following formula.

Definition 3.4.3 *The resulting non-cooperative outcome for player $i = 1, \dots, n$ in the subgame of the game $\Gamma(p_0, T - t_0)$ with dynamic updating starting at $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ has the following form:*

$$\begin{aligned} \hat{V}_i(t, \hat{p}_{NE}(t)) = & \sum_{m=j+1}^l \left[V_i^m(t_0 + m\Delta t, p_{m,0}^{NE}) - V_i^m(t_0 + (m+1)\Delta t, p_{m,1}^{NE}) \right] \\ & + \left[V_i^j(t, p_j^{NE}(t)) - V_i^j(t_0 + (j+1)\Delta t, p_{j,1}^{NE}) \right], \quad i \in N. \end{aligned} \quad (3.154)$$

3.4.2.5 Numerical Simulation

The first game starts in December 2015 and lasts till the summit in Vienna in November 2016. The oil as a homogeneous product is considered, and the demand function is appraised with the parameters of an average world oil price and the total world oil supply. The average oil prices are calculated for each period based only on two major trading classifications, namely, those of Brent crude and light crude, which are available from on the Finam agency data source. As the initial price, the

average price in December 2015 is taken, which is equal to $p_0 = 34.51$.

Period	Total World Supply, MMBD	Brent, \$ for barrel	Light, \$ for barrel	Average price, \$ for barrel
12.2015	96.411	35.910	33.110	34.510
01.2016	95.875	36.640	33.990	35.315
02.2016	95.420	40.140	37.820	38.980
03.2016	95.294	47.320	45.990	46.655
04.2016	95.400	49.520	48.750	49.135
05.2016	95.187	49.740	48.640	49.190
06.2016	95.954	43.270	41.760	42.515
07.2016	96.891	46.970	45.000	45.985
08.2016	95.894	49.990	48.050	49.020
09.2016	96.001	48.510	46.970	47.740
10.2016	97.362	44.520	43.120	43.820
11.2016	97.241	47.034	45.908	46.471

To ensure that the slope of demand is negative, it is decided not to regress the historical oil prices for the quantity produced, taking into account the average price and the total world supply (in the table above) for only one previous period and a fixed choke price. Thus, it is assumed that the demand parameter a_j is equal to 300 at each period. Therefore, the parameter d_j can be obtained from (3.145):

$$d_j = (a_j - \hat{p}(t-1)) \sum_{i \in N} q_i(t-1). \quad (3.155)$$

The length of each Δt -time interval is one month. The players use the appraised demand with the parameters a_j and d_j as the forecast for the next $\bar{T} = 3$ periods. The values are set upon the cost function parameters using the total cost of producing a barrel and the average volumes of oil production for the players in 2016 under a fixed value $g_i = 0.7$ for each player and each period. Both the parameters of c_i and g_i remain unchanged during the game. It is assumed that the speed of adjustment is $s = 0.2$, and the discount factor is $r = 10\%$.

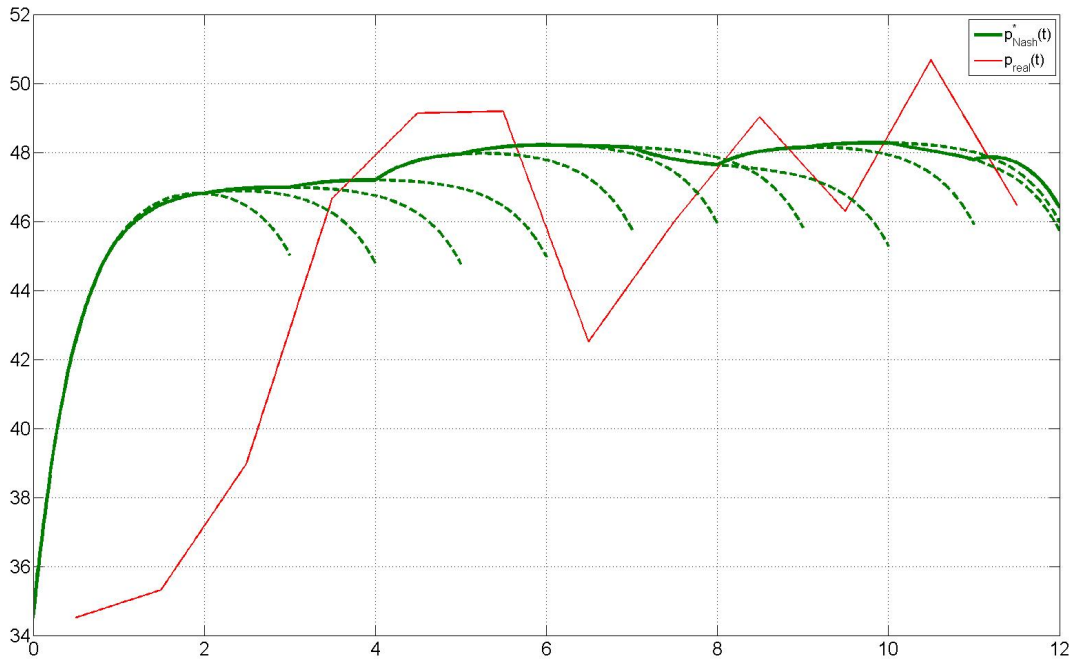


Figure 3.21: Conditionally non-cooperative trajectory of oil price $\hat{p}_{NE}(t)$ (thick solid line) with dynamic updating and historical oil price trajectory (thick dotted line).

Period	a	d
12.2015	300	2.717
01.2016	300	2.754
02.2016	300	2.761
03.2016	300	2.735
04.2016	300	2.659
05.2016	300	2.630
06.2016	300	2.635
07.2016	300	2.683
08.2016	300	2.622
09.2016	300	2.617
10.2016	300	2.628
11.2016	300	2.631

i	Producer	c	g
1	OPEC	3.169	0.7
2	Non-OPEC	17.333	0.7
3	US shale	20.238	0.7
4	US non-shale	18.182	0.7
5	Others	20.867	0.7

Fig. 2 shows the comparison of the conditionally non-cooperative trajectory $\{\hat{p}_{NE}(t)\}_{t=t_0}^T$ and the historical average oil price dynamics. In Fig. 3 the Nash feedback strategies (3.151) corresponding to $\{\hat{p}_{NE}(t)\}_{t=t_0}^T$ and the historical oil production quantities are presented for each group of countries.

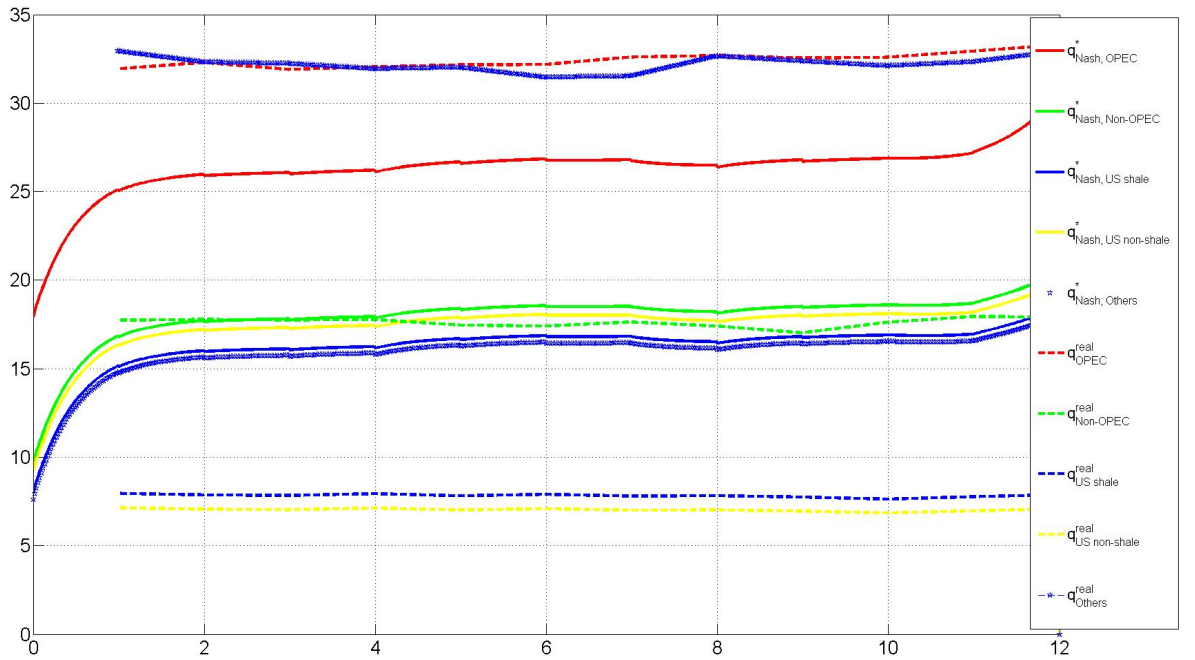


Figure 3.22: Nash feedback strategies under dynamic updating (solid lines) and corresponding historical oil production quantities (dashed lines).

According to Figs. 2 and 3, the parameters c_i and g_i from the table below can be used for modeling the cooperative agreement on the time interval from December 2016 to May 2017.

The payoffs of players $\hat{V}_i(t, \hat{p}_{NE}(t))$, $i \in N$, corresponding to the Nash equilibrium strategies along the trajectory $\hat{p}_{NE}(t)$ are shown in Fig. 4.

3.4.3 Partially Cooperative Game Model

As mentioned in the introduction, at the end of November 2016 two groups of countries, OPEC and non-OPEC (players 1 and 2) signed the cooperative agreement. Using the model from Section 1, we construct a model of this partial cooperation by introducing a 4-person game, where players 1 and 2 act as one player and the remaining players (players 3, 4, 5) act independently. As a historical cooperative solution, a linear reduction of the quantities produced to the level agreed is used. Also, a cooperative solution suggested in the paper [165] is constructed, which possesses the subgame consistency property. In the simulation example, the difference between the historical solution and the solution proposed in [165] is discussed.

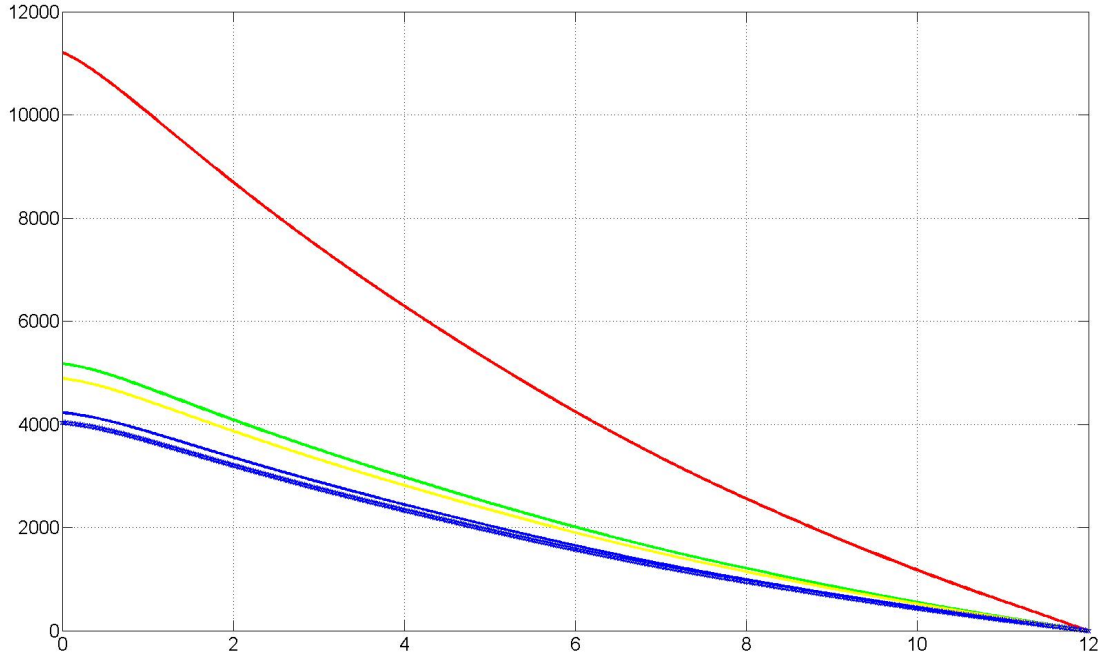


Figure 3.23: Payoffs of players corresponding to Nash equilibrium $\hat{V}_i(t, \hat{p}_{NE}(t))$.

3.4.3.1 Partially Cooperative Truncated Subgame

To model the behavior with players 1,2 cooperating and the others acting independently, first we construct a 4-person differential game. The set of players in this game is denoted by $\tilde{N} = \{(1, 2), 3, 4, 5\}$, $|\tilde{N}| = 4$. Players 1 and 2 cooperate by combining their strategies sets, but the actual payoffs remain individual. Thus, for players 1 and 2, an analog of a cooperative differential game with non-transferable payoffs is constructed. Players 1, 2 orient themselves towards the payoff defined as a linear combination of their payoffs with fixed coefficients α_j :

$$\begin{aligned} \tilde{K}_{1,2}^j(\alpha_j, p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^j, \dots, q_5^j) &= \\ &= \sum_{i \in \{1,2\}} \alpha_i^j K_i^j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^j, \dots, q_5^j), \end{aligned} \quad (3.156)$$

where $K_i^j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^j, \dots, q_5^j)$ are defined in (3.146), and

$$\alpha_j \in \Lambda_j = \{\alpha : \alpha_1^j + \alpha_2^j = 1, \alpha_i^j \geq 0, i = 1, 2\}. \quad (3.157)$$

The individual payoffs $\tilde{K}_i^j(\alpha_j, p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^j, \dots, q_5^j)$, $i = 1, 2$, which the players obtain during the game, are calculated by formula (3.146). The payoffs

of players $i = 3, 4, 5$ are calculated by formula (3.146).

As it turns, for any $\alpha_j \in \Lambda_j$ a unique differential game is obtained.

Definition 3.4.4 *Let $j = 0, \dots, l$. A partially cooperative truncated subgame $\tilde{\Gamma}_j^{\alpha_j}(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ is defined on the time interval $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$. The motion equation and the initial condition of the truncated subgame $\tilde{\Gamma}_j^{\alpha_j}(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ have the following form:*

$$\dot{p}(t) = s \left(a_j - d_j \sum_{i \in N} q_i^j(t) - p(t) \right), \quad p(t_0 + j\Delta t) = p_{j,0}. \quad (3.158)$$

The payoff function $\tilde{K}_{1,2}^j(\alpha_j, p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^j, \dots, q_5^j)$ of players $\{1, 2\}$ in truncated subgame j is defined like in (3.156); the payoff function $\tilde{K}_i^j(\alpha_j, p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^j, \dots, q_5^j)$ of players $i = 3, 4, 5$ is defined like in (3.146).

3.4.3.2 Pareto Optimal Outcomes

Suppose that players 1 and 2 can agree on the form of a joint payoff function, i.e., the coefficients $\alpha_j \in \Lambda_j$. The set of outcomes for all possible $\alpha_j \in \Lambda_j$ for players 1 and 2 defines the set of Pareto optimal outcomes for players 1 and 2. Consider the feedback Nash equilibrium $q_{\alpha_j}^*(t, p)$ in the game $\tilde{\Gamma}_j^{\alpha_j}(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ of 4 persons. For all possible values of α_j , the Nash equilibrium payoffs of players 1 and 2 form a set of Pareto-optimal outcomes $\tilde{K}_1^j(\alpha_j, p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^{*\alpha_j}, \dots, q_5^{*\alpha_j})$ and $\tilde{K}_2^j(\alpha_j, p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; q_1^{*\alpha_j}, \dots, q_5^{*\alpha_j})$. The feedback Nash equilibrium strategies of players are defined like in [217].

Let t and p be the time and initial state of the partially cooperative truncated subgame $\tilde{\Gamma}_j^{\alpha_j}(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$. In this subgame, we denote by $W_{1,2}^{\alpha_j}(t, p)$ the value of players $\{1, 2\}$ in the feedback Nash equilibrium with a determined α_j . Also, we denote by $W_i^{\alpha_j}(t, p)$, $i = 3, 4, 5$, the payoff of players $i = 3, \dots, 5$ using the Nash equilibrium strategies $q_{\alpha_j}^*(t, p)$. Sufficient conditions for the solution and the feedback Nash equilibrium are given below.

Theorem 3.4.2 *Assume that there exist continuously differential functions $W_{1,2}^{\alpha_j}(t, p)$, $W_3^j(t, p)$, $W_4^j(t, p)$, $W_5^j(t, p)$: $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}] \times R \rightarrow R$ satisfying the partial*

differential equation

$$\begin{aligned}
& -W_t^{1,2, \alpha_j}(t, p) = \max_{q_1^j, q_2^j} \left\{ \sum_{i=1}^2 e^{-\rho(t-t_0)} \alpha_i^j \left[q_i^j (p - c_i - g_i q_i^j) \right] + \right. \\
& \left. + W_p^{1,2, \alpha_j}(t, p) s \left(a - d \left[\sum_{i=1}^2 q_i^j(t) + \sum_{k=3}^5 q_k^{*\alpha_j}(t) \right] - p(t) \right) \right\}, \\
& -W_t^{i, \alpha_j}(t, p) = \max_{q_i^j} \left\{ e^{-\rho(t-t_0)} \left[q_i^j (p - c_i - g_i q_i^j) \right] + \right. \\
& \left. + W_p^{i, \alpha_j}(t, p) s \left(a - d \left[q_i^j(t) + \sum_{k \neq i}^5 q_k^{*\alpha_j}(t) \right] - p(t) \right) \right\}, \quad i = 3, \dots, 5, \quad (3.159)
\end{aligned}$$

where $W_{1,2}^{\alpha_j}(t_0 + j\Delta t + \bar{T}, p) = 0$, $W_i^{\alpha_j}(t_0 + j\Delta t + \bar{T}, p) = 0$, $i=3, \dots, 5$. Then $q_{\alpha_j}^*(t, p)$ provides a feedback Nash equilibrium in the partial cooperative truncated subgame $\tilde{\Gamma}_j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$.

In this game model, the Bellman function $W_{1,2}^{\alpha_j}(t, p)$, $W_i^{\alpha_j}(t, p)$, $i = 3, \dots, 5$ can be obtained in the form

$$\begin{aligned}
W_{1,2}^{\alpha_j}(t, p) &= e^{-\rho(t-t_0)} \left[A_j^{1,2}(t) p^2 + B_j^{1,2}(t) p + C_j^{1,2}(t) \right], \\
W_i^{\alpha_j}(t, p) &= e^{-\rho(t-t_0)} \left[A_j^i(t) p^2 + B_j^i(t) p + C_j^i(t) \right], \quad i = 3, 4, 5.
\end{aligned} \quad (3.160)$$

The feedback Nash equilibrium strategies in the partial cooperative truncated subgame are

$$\begin{aligned}
q_i^{*\alpha_j}(t, p) &= -\frac{(c_i - p) + ds \left[B_j^{1,2}(t) + 2A_j^{1,2}(t)p \right]}{2\alpha_i g_i}, \quad i = 1, 2, \\
q_i^{*\alpha_j}(t, p) &= -\frac{(c_i - p) + sd \left[B_j^i(t) + 2A_j^i(t)p \right]}{2g_i}, \quad i = 3, 4, 5.
\end{aligned} \quad (3.161)$$

The functions $A_j^{1,2}(t)$, $B_j^{1,2}(t)$, $C_j^{1,2}(t)$, $A_j^i(t)$, $B_j^i(t)$, and $C_j^i(t)$, $i = 1, 2$, $t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, are defined by the following systems of differential equations:

for $A_j^{1,2}(t_0 + j\Delta t + \bar{T})$, $B_j^{1,2}(t_0 + j\Delta t + \bar{T})$, and $C_j^{1,2}(t_0 + j\Delta t + \bar{T})$,

$$\begin{aligned}
\dot{A}_j^{1,2}(t) &= A_j^{1,2}(t)(\rho + 2s) + \sum_{k \in \{1,2\}} \left(\frac{\alpha_k^j}{4g_k} + \frac{(A_j^{1,2}(t)d_j s)^2}{\alpha_k^j g_k} \right) + \\
&+ \sum_{k \in \{3,4,5\}} \frac{2A_j^k(t)A_j^{1,2}(t)d_j^2 s^2}{g_k} - \sum_{k \in N} \frac{A_j^{1,2}(t)d_j s}{g_k}, \\
\dot{B}_j^{1,2}(t) &= B_j^{1,2}(t)(\rho + s) - 2A_j^{1,2}(t)a_j s - \sum_{k \in \{1,2\}} \left(\frac{\alpha_k^k c_k}{2g_k} - \frac{A_j^{1,2}(t)B_j^{1,2}(t)d_j^2 s^2}{\alpha_k^j g_k} \right) \\
&- \sum_{k \in \{3,4,5\}} \left(\frac{A_j^{1,2}(t)B_j^k(t)d_j^2 s^2}{g_k} + \frac{A_j^k(t)B_j^{1,2}(t)d_j^2 s^2}{g_k} \right) + \\
&- \sum_{k \in N} \left(\frac{B_j^{1,2}(t)d_j s}{2g_k} - \frac{A_j^{1,2}(t)c_k d_j s}{g_k} \right) \\
\dot{C}_j^{1,2}(t) &= C_j^{1,2}(t)\rho - B_j^{1,2}(t)a_j s + \sum_{k \in \{1,2\}} \left(\frac{\alpha_k^k c_k^2}{4g_k} + \frac{(B_j^{1,2}(t)d_j s)^2}{4\alpha_k^j g_k} \right) + \\
&+ \sum_{k \in \{3,4,5\}} \frac{B_j^{1,2}(t)B_j^k(t)d_j^2 s^2}{2g_k} + \sum_{k \in N} \frac{B_j^{1,2}(t)c_k d_j s}{2g_k},
\end{aligned}$$

where $A_j^{1,2}(t_0 + j\Delta t + \bar{T}) = 0$, $B_j^{1,2}(t_0 + j\Delta t + \bar{T}) = 0$, and $C_j^{1,2}(t_0 + j\Delta t + \bar{T}) = 0$;

for $A_j^i(t_0 + j\Delta t + \bar{T})$, $B_j^i(t_0 + j\Delta t + \bar{T})$, and $C_j^i(t_0 + j\Delta t + \bar{T})$, $i = 3, \dots, 5$,

$$\begin{aligned}
\dot{A}_j^i(t) &= A_j^i(t)(\rho + 2s) + \sum_{k \in \{1,2\}} \left(\frac{1}{4g_i} + \frac{2A_j^i(t)A_j^{1,2}(t)d_j^2s^2}{\alpha_k^j g_k} \right) + \\
&+ \sum_{k \in \{3,4,5\}} \frac{2A_j^k(t)A_j^i(t)d_j^2s^2}{g_i} - \sum_{k \in N} \frac{A_j^i(t)d_j s}{g_k}, \\
\dot{B}_j^i(t) &= B_j^i(t)(\rho + s) - 2A_j^i(t)a_j s - \frac{c_i}{2g_i} - \sum_{k \in N} \left(\frac{B_j^i(t)d_j s}{2g_k} - \frac{A_j^i(t)c_k d_j s}{g_k} \right) \\
&+ \sum_{k \in \{1,2\}} \left(\frac{A_j^i(t)B_j^{1,2}(t)d_j^2s^2}{\alpha_k^j g_k} + \frac{A_j^{1,2}(t)B_j^i(t)d_j^2s^2}{\alpha_k^j g_k} \right) + \\
&+ \sum_{k \in \{3,4,5\}} \left(\frac{A_j^i(t)B_j^k(t)d_j^2s^2}{g_k} + \frac{A_j^k(t)B_j^i(t)d_j^2s^2}{g_k} \right) \\
\dot{C}_j^i(t) &= C_j^i(t)\rho + \frac{c_i^2}{4g_i} - B_j^i(t)a_j s + \sum_{k \in \{1,2\}} \frac{B_j^i(t)B_j^{1,2}(t)d_j^2s^2}{2\alpha_k^j g_k} + \\
&+ \sum_{k \in \{3,4,5\}} \frac{B_j^i(t)B_j^k(t)d_j^2s^2}{2g_k} + \sum_{k \in N} \frac{B_j^k(t)c_k d_j s}{2g_k},
\end{aligned}$$

where $A_j^i(t_0 + j\Delta t + \bar{T}) = 0$, $B_j^i(t_0 + j\Delta t + \bar{T}) = 0$, and $C_j^i(t_0 + j\Delta t + \bar{T}) = 0$.

Substituting $q_{\alpha_j}^*(t)(t, p)$ (3.161) into (3.158) yields the dynamics of the feedback Nash equilibrium trajectory corresponding to the vector of weights for players 1, 2:

$$\dot{p}(t) = s \left(a - d \sum_{i \in N} q_i^{*\alpha_j}(t, p) - p(t) \right), \quad p(t_0 + j\Delta t) = p_{j,0}. \quad (3.162)$$

We denote by $p_{\alpha_j}^*(t)$ the solution of the system (3.162) corresponding to the fixed weights $\alpha_j = (\alpha_1^j, \alpha_2^j)$. Further $p_{\alpha_j}^*(t)$ will be called the Pareto optimal trajectory.

3.4.3.3 Individual Players Payoffs under Cooperation

In the previous section, the payoffs of players 3, 4, 5 have been defined, but the individual payoff of players 1, 2 have not. The next step is to define the individual payoffs of the players under cooperation. This can be achieved by substituting the Pareto optimal trajectory $p_{\alpha_j}^*(t)$ (3.162) and the strategies $q_{\alpha_j}^*(t, p)$ (3.161) into the

formula for each player's payoff (3.146). The payoff of players 1, 2 in the truncated subgame $\tilde{\Gamma}_j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ along the Pareto optimal trajectory $p_{\alpha_j}^*(t)$ with the strategies $q_{\alpha_j}^*(t, p)$ involved will be denoted by $W_1^{\alpha_j}(t, p)$ and $W_2^{\alpha_j}(t, p)$, respectively. Another (and more preferable) way is to apply the results from the paper [165], where a special form of the Hamilton-Jacobi-Bellman equation was solved for each player with fixed strategies $q_{\alpha_j}^*(t, p)$ (3.161). (So no maximization in the right-hand side of the equation was needed.)

3.4.3.4 Subgame Consistent Solutions

In the paper [165], the notion of subgame consistency of a cooperative solution in a game with non-transferable payoffs was defined as follows:

Definition 3.4.5 *A payoff vector $W(t, p^*(t)) = (W_1(t, p^*(t)), \dots, W_n(t, p^*(t)))$ along a trajectory $p^*(t)$ is subgame consistent in the game $\Gamma_c(p_0, T - t_0)$ with non-transferable payoffs if it satisfies the following conditions:*

1 *The payoff vector*

$$W(t, p^*(t)) = (W_1(t, p^*(t)), \dots, W_n(t, p^*(t))), \quad t \in [t_0, T] \quad (3.163)$$

is Pareto optimal.

2 *The payoff vector satisfies the individual rationality property*

$$W_i(t, p^*(t)) \geq V_i(t, p^*(t)), \quad t \in [t_0, T], \quad i = \overline{1, n}, \quad (3.164)$$

where $V_i(t, p^(t))$ is an individual payoff of player $i = \overline{1, n}$.*

Definition 3.4.6 *Let us define the set $S_j^\tau = \bigcap_{t_0 + j\Delta t \leq t < \tau} S_t^j$ for $\tau \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t)$, $j = 0, \dots, l$, where S_t^j is the set of $\alpha_j = (\alpha_1^j, \alpha_2^j)$ that satisfy individual rationality at the time instant t for $i = \overline{1, n}$:*

$$W_i^{\alpha_j}(t, p_{\alpha_j}^*(t)) \geq V_i^j(t, p_{\alpha_j}^*(t)), \quad (3.165)$$

where $V_i(t, p^(t))$ is the payoff of player $i \in N$ in the non-cooperative game with Nash equilibrium strategies calculated along the Pareto optimal trajectory $p_{\alpha_j}^*(t)$.*

$S_j^{t_0+(j+1)\Delta t}$ is the set of $\alpha_j = (\alpha_1^j, \alpha_2^j)$ that satisfy the individual rationality (3.165) throughout the time interval before the information about the game structure is updated, at the time instant $t = t_0 + (j + 1)\Delta t$. The set $S_j^{t_0+(j+1)\Delta t}$ provides the corresponding set of subgame consistent solutions:

$$\left\{ W^{\alpha_j}(t, p_{\alpha_j}^*(t)) = (W_1^{\alpha_j}(t, p_{\alpha_j}^*(t)), \dots, W_n^{\alpha_j}(t, p_{\alpha_j}^*(t))), \right. \\ \left. t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t], \alpha_j \in S_j^{t_0+(j+1)\Delta t}, i = \overline{1, n}, j = 0, \dots, l \right\}. \quad (3.166)$$

A more general case was considered in [165], where individual rationality property (3.165) holds on the entire time interval. The following result was proved.

Theorem 3.4.3 *A solution optimality principle under which the players agree to choose the same weights $\alpha = (\alpha_1, \dots, \alpha_n) \in \Lambda$ in all the subgames $\Gamma_c(p_\alpha^*(t), T - t)$, $t \in [t_0, T]$, such that*

$$W_i^\alpha(t, p_\alpha^*(t)) \geq V_i(t, p_\alpha^*(t)), \quad t \in [t_0, T], \quad i = \overline{1, n}. \quad (3.167)$$

yields a subgame consistent solution of the cooperative game $\Gamma_c(p_0, T - t_0)$.

3.4.3.5 Solution Concept

Suppose that at the beginning of each truncated subgame $\tilde{\Gamma}_j(p_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, some weights $\alpha_j \in S_j^{t_0+(j+1)\Delta t}$ are chosen. Then each truncated subgame evolves along the Pareto optimal trajectory $p_{\alpha_j}^*(t)$ and consequently, the entire game $\Gamma_c(p_0, T - t_0)$ evolves along the following trajectory, where $p_{j,0}^* = p_{\alpha_{j-1}}^*(t_0 + j\Delta t)$.

Definition 3.4.7 *The conditionally Pareto optimal trajectory $\{\hat{p}^*(t)\}_{t=t_0}^T$ is the combination of Pareto optimal trajectories $p_{\alpha_j}^*(t)$ for each truncated subgame*

$\tilde{\Gamma}_j(p_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ and the chosen weights $\alpha_j \in S_j^{t_0+(j+1)\Delta t}$:

$$\{\hat{p}^*(t)\}_{t=t_0}^T = \begin{cases} p_{\alpha_0}^*(t), & t \in [t_0, t_0 + \Delta t), \\ \dots, \\ p_{\alpha_j}^*(t), & t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t), \\ \dots, \\ p_{\alpha_l}^*(t), & t \in [t_0 + l\Delta t, t_0 + (l+1)\Delta t]. \end{cases} \quad (3.168)$$

Now let us define the solution of the entire game $\Gamma_c(p_0, T - t_0)$. Let $p_{j,1}^* = p_{\alpha_j}^*(t_0 + (j+1)\Delta t)$.

Definition 3.4.8 *The resulting solution of the game $\Gamma_c(p_0, T - t_0)$ with dynamic updating corresponding to the chosen weights $\alpha_j \in S_j^{t_0+(j+1)\Delta t}$ for each truncated subgame $\tilde{\Gamma}_j(p_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, $j = 0, \dots, l$, $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, is defined in the following way:*

$$\begin{aligned} \hat{W}_i(t, \hat{p}^*(t)) = & \sum_{m=j+1}^l \left[W_i^{\alpha_m}(t_0 + m\Delta t, p_{m,0}^*) - W_i^{\alpha_m}(t_0 + (m+1)\Delta t, p_{m,1}^*) \right] + \\ & + \left[W_i^{\alpha_j}(t, p_{\alpha_j}^*(t)) - W_i^{\alpha_j}(t_0 + (j+1)\Delta t, p_{j,1}^*) \right]. \end{aligned} \quad (3.169)$$

Therefore, the set of all possible solutions of the game $\Gamma_c(p_0, T - t_0)$ has the following form:

$$\left\{ \hat{W}(t, \hat{p}^*(t)) = (\hat{W}_1(t, \hat{p}^*(t)), \dots, \hat{W}_n(t, \hat{p}^*(t))), t \in [t_0, T], \right. \\ \left. \alpha_j \in S_j^{t_0+(j+1)\Delta t}, \forall j = 0, \dots, l \right\}. \quad (3.170)$$

In the game with a moving information horizon, a solution is called conditionally Pareto optimal if it is constructed using the Pareto optimal solutions in each truncated subgame. In this sense, the resulting solution $\hat{W}(t, \hat{p}^*(t))$ is called conditionally Pareto optimal.

3.4.3.6 Subgame Consistency of Solution Concept

In [111], the subgame consistency property of the resulting solution (3.169) was proved for the case of full cooperation. A similar result can be obtained for partial cooperation.

Theorem 3.4.4 *The resulting solution $\hat{W}(t, \hat{p}^*(t))$ is subgame consistent in the game $\Gamma_c(\hat{p}^*(t), T - t)$ with dynamic updating if the following condition is satisfied $\forall t \in [t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, $j = 0, \dots, l$, $i = \overline{1, n}$:*

$$\begin{aligned} W_i^{\alpha_j}(t, p_{\alpha_j}^*(t)) - V_i^j(t, p_{\alpha_j}^*(t)) &\geq \\ &\geq W^{\alpha_j}(t_0 + (j + 1)\Delta t, p_{j,1}^*) - V_i^j(t_0 + (j + 1)\Delta t, p_{j,1}^*), \end{aligned} \quad (3.171)$$

3.4.4 Historical Cooperative Agreement

To make a comprehensive comparison of the theoretical results and the real situation, we should mention the complexity of such an agreement at the international level. It affects not only oil market situations inside the countries but also their balance and all macroeconomic indicators. The agreement is of both economic and political nature. According to the agreement of November 30, 2016, the OPEC countries and 11 non-OPEC countries decided to reduce their production by 1.18 – 7.69% for different countries. Herewith, Iran, Libya, and Nigeria were allowed to boost their output. After the agreement was reached, oil prices did start to increase. However, by the end of May 2017, they had fallen to almost the same level. Countries-outsiders, notably the US shale producers, increased their output. Most of the countries inside the agreement succeeded in fulfilling it. After six months, all these countries decided to prolong the agreement for the next year.

To model this cooperative agreement, we construct linear strategies for players 1 and 2. The Nash equilibrium strategies are used in the non-cooperative differential game for the other players. The optimal strategies for players 3, 4, 5 are obtained in the same way as in Section 1 using the Hamilton-Jacobi-Bellman equation with the strategies (3.172) involved.

According to the statistical data, the following strategies of players $i = 1, 2$ under

the cooperative agreement are used:

$$\bar{q}_1^*(t) = 3.5\%q_1^{avr}(t_{nov})\left(1 - 3.5\%t/T\right) \quad (3.172)$$

$$\bar{q}_2^*(t) = 3.1\%q_2^{avr}(t_{nov})\left(1 - 3.1\%t/T\right), \quad (3.173)$$

where $q_{avr}(t_{nov}) = (q_1^{avr}(t_{nov}), q_2^{avr}(t_{nov})) = (39.838, 18.771)$ is the average quantity of produced oil at t_{nov} , which is November 2016, and T corresponds to May 2017.

3.4.4.1 Numerical Simulation

The second game starts immediately after the summit and lasts from December 2016 until May 2017. Like on the previous time interval, the average oil prices for each period are calculated based only on the two major trading classifications of Brent crude and light crude, which are available from the Finam agency data source. The average price in December 2016 is taken as the initial price $p_0 = 53.73$ for the second game.

Period	Total World Supply, MMBD	Brent, \$ for barrel	Light, \$ for barrel	Average price, \$ for barrel
12.2016	97.324	55.008	52.449	53.728
01.2017	96.627	55.607	52.951	54.279
02.2017	96.412	56.122	53.661	54.892
03.2017	96.411	52.592	49.873	51.232
04.2017	97.075	53.841	51.216	52.529
05.2017	97.529	51.533	48.744	50.139

To test the constructed model with dynamic updating, we choose the same players' parameters c_i and g_i as for the previous time interval, supposing that the demand parameter a_j is equal to 300 at each Δt -period. The players use the appraised demand with the parameters a_j and b_j as the forecast for the next $\bar{T} = 3$ periods. The parameter d_j is obtained using formula (3.155).

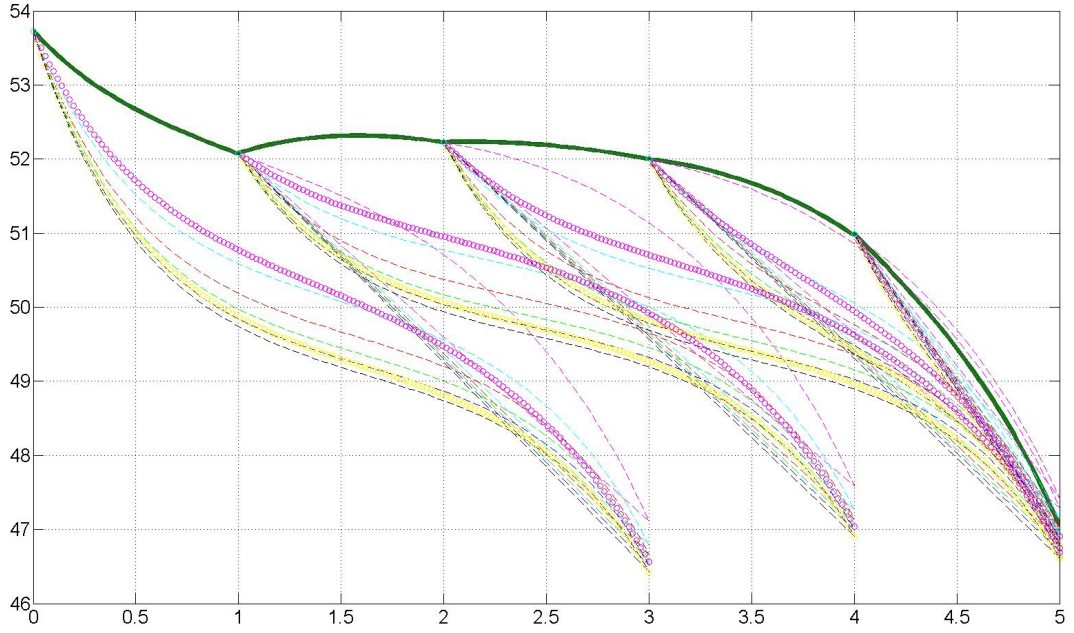


Figure 3.24: Conditionally Pareto optimal trajectory $\hat{p}^*(t)$ (solid line) and set of Pareto optimal trajectories $p_{\alpha_j}^*(t)$ (dashed lines) corresponding to weights $\alpha_j \in S_j^{t_0+(j+1)\Delta t}$ in each truncated subgame $\tilde{\Gamma}_j(p_{j,0}^*, t_0 + \Delta t, t_0 + \Delta t + \bar{T})$.

Period	a	d	i	Producer	c	g
01.2017	300	2.530	1	OPEC	3.169	0.7
02.2017	300	2.543	2	Non-OPEC	17.333	0.7
03.2017	300	2.542	3	US shale	20.238	0.7
04.2017	300	2.580	4	US non-shale	18.182	0.7
05.2017	300	2.549	5	Others	20.867	0.7

In Fig. 5 we can see the conditionally Pareto optimal trajectory $\hat{p}^*(t)$ (solid line) and the set of Pareto optimal trajectories $p_{\alpha_j}^*(t)$ (dashed lines) corresponding the weights $\alpha_j \in S_j^{t_0+(j+1)\Delta t}$ in each truncated subgame $\tilde{\Gamma}_j(p_{j,0}^*, t_0 + \Delta t, t_0 + \Delta t + \bar{T})$. In Fig. 6 the optimal strategies corresponding to $\{\hat{p}_{\alpha_j}^*(t)\}_{t=t_0}^T$ and the historical quantities of oil production are presented for each group of the countries.

In Fig. 7 we can see the individual payoff of player $i = 1$ under cooperation $W_1^j(t, p_{\alpha_j}^*(t))$ (solid lines) for all possible weights $\alpha_j \in S_j^{t_0+(j+1)\Delta t}$ and $\alpha_j^* \in S_j^{t_0+(j+1)\Delta t}$, $j = 3$. In Fig. 8 we can see the payoff of players $\hat{W}_i(t, \hat{p}^*(t))$, $i \in N$ (solid lines) and the payoff of players $\hat{V}_i(t, \hat{p}^*(t))$, $i \in N$ (dashed lines) corresponding to the Nash equilibrium strategies along the conditionally Pareto optimal trajectory

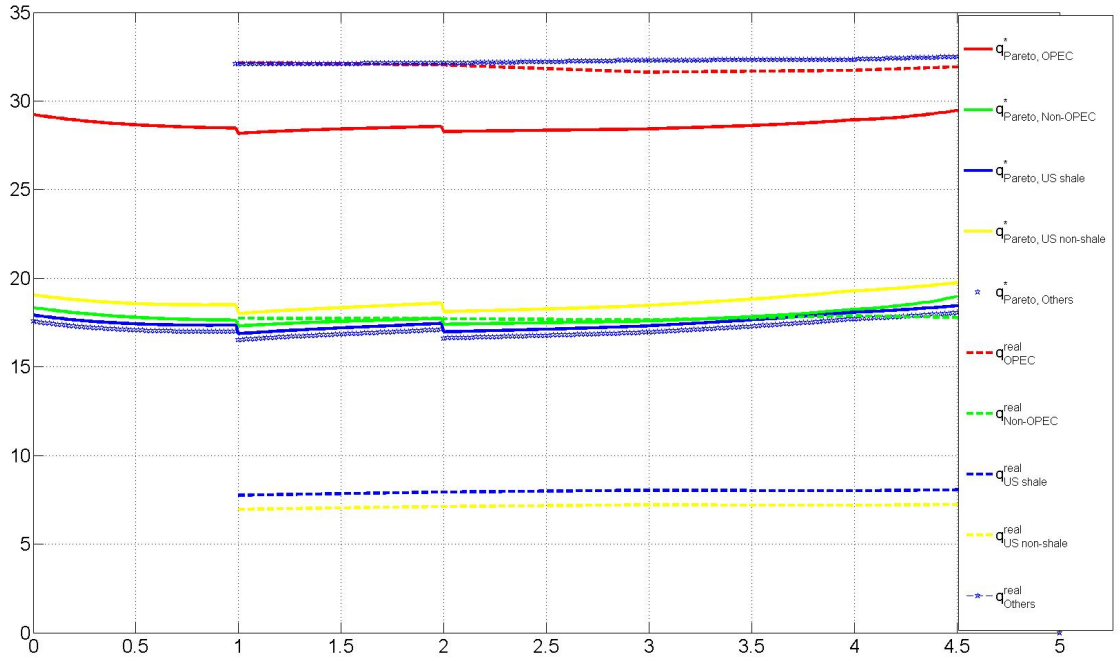


Figure 3.25: Optimal strategies under dynamic updating (solid lines) and corresponding historical quantities of oil production (dashed lines).

$\hat{p}^*(t)$.

3.4.4.2 Comparison of Historical Cooperative Solution and Proposed Solution

This section compares the historical cooperative agreement proposed in (3.172) with the strategies $\bar{q}^*(t)$ involved as well as the subgame consistent cooperative solution proposed in (3.169). In Fig. 9 we can see the conditionally Pareto optimal trajectory $\hat{p}_{\alpha_j}^*(t)$ corresponding to the chosen weights $\alpha_j^* \in S_j^{t_0+(j+1)\Delta t}$ in all the truncated subgames $\tilde{\Gamma}_j(p_{j,0}^*, t_0 + \Delta t, t_0 + \Delta t + \bar{T})$, $j = 0, \dots, l$ and the trajectory $\bar{p}^*(t)$ corresponding to the historical cooperative agreement and the historical oil price dynamics.

In Fig. 10 the optimal strategies (3.151) corresponding to the $\hat{p}_{\alpha_j}^*(t)$, the strategies $\bar{q}^*(t)$ (3.172) corresponding to the historical cooperative agreement are presented for each group of countries and the historical quantities of oil production for each group of countries. In Fig. 11 we can see the payoff of players $\hat{W}_i(t, \hat{p}^*(t))$, $i \in N$ (solid lines), along the trajectory $\hat{p}^*(t)$ and the payoffs corresponding to the historical cooperative agreement (dashed lines).

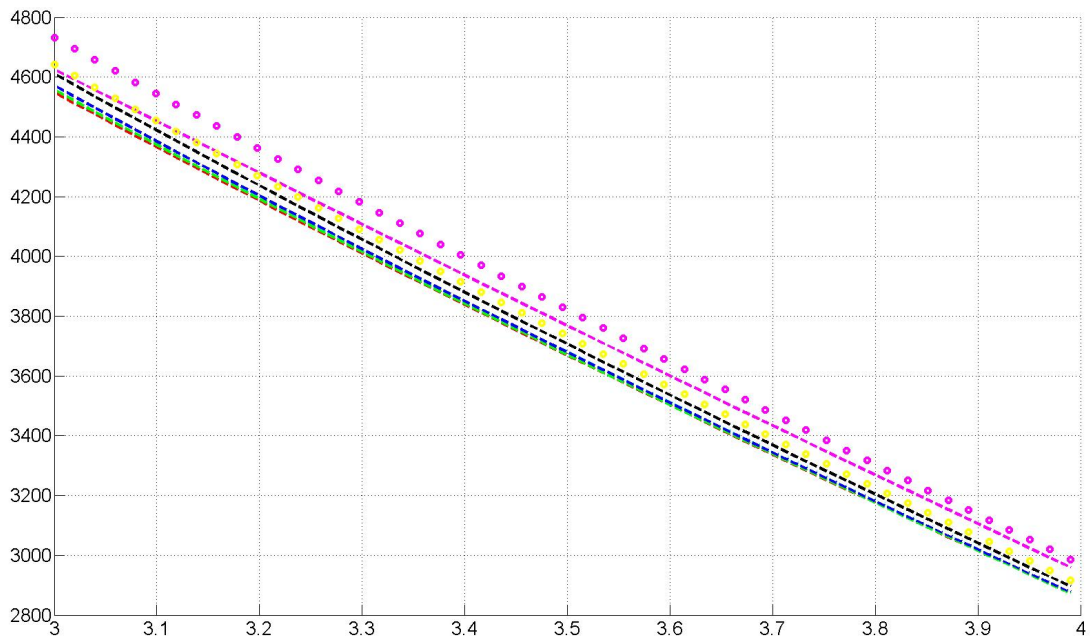


Figure 3.26: Set of individual payoffs $W_1^j(t, p_{\alpha_j}^*(t))$ of player 1 for all possible weights $\alpha_j \in S_j^{t_0+(j+1)\Delta t}$ in truncated subgame $\tilde{\Gamma}_j(p_{j,0}^*, t_0 + \Delta t, t_0 + \Delta t + \bar{T})$, $j = 3$.

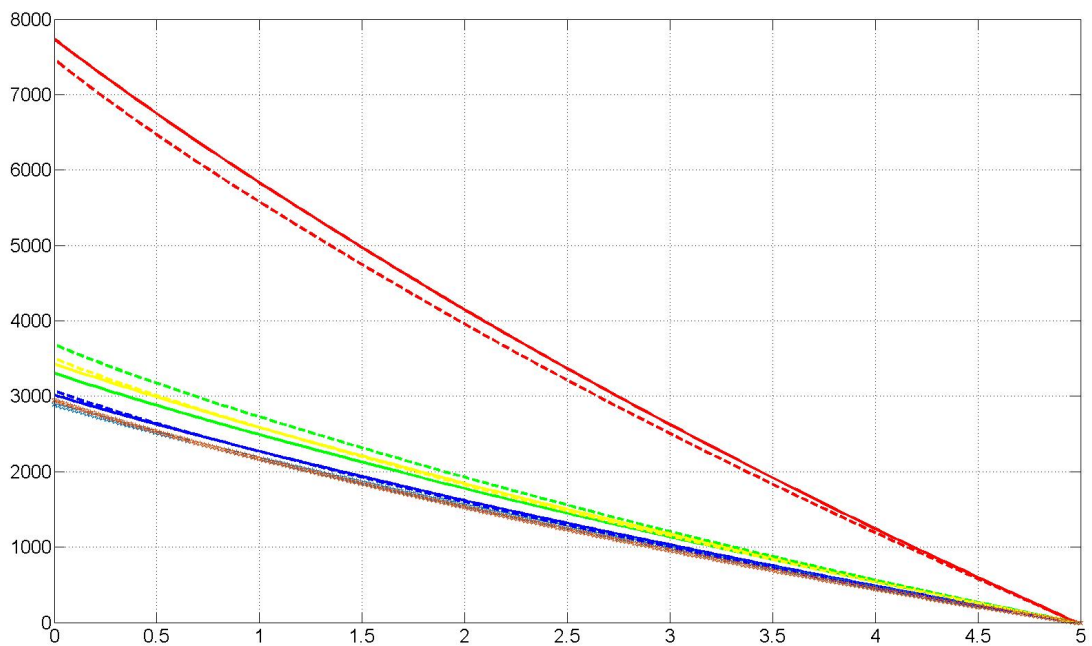


Figure 3.27: $\hat{W}_i(t, \hat{p}^*(t))$, $i \in N$ (solid lines) and $\hat{V}_i(t, \hat{p}^*(t))$, $i \in N$ (dashed lines).

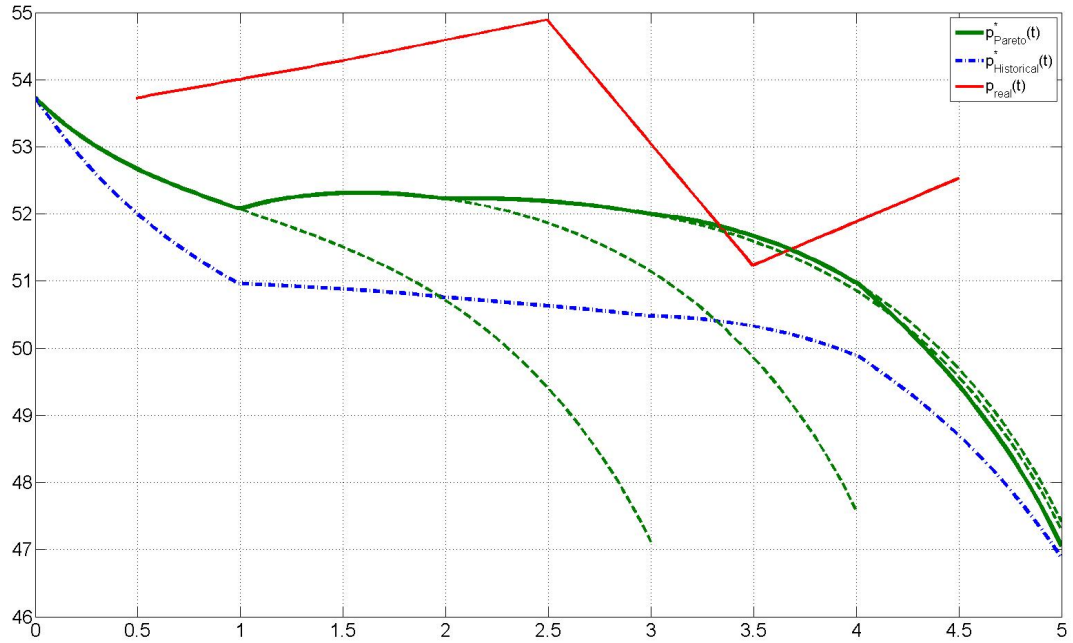


Figure 3.28: Conditionally Pareto optimal trajectory $\hat{p}_{\alpha_j}^*(t)$ (thick solid line) corresponding to chosen weights $\alpha_j^* \in S_j^{t_0+(j+1)\Delta t}$, $j = 0, \dots, l$, trajectory $\bar{p}^*(t)$ (dashed line) corresponding to historical cooperative agreement (3.172) and historical oil price dynamics (thin solid line).

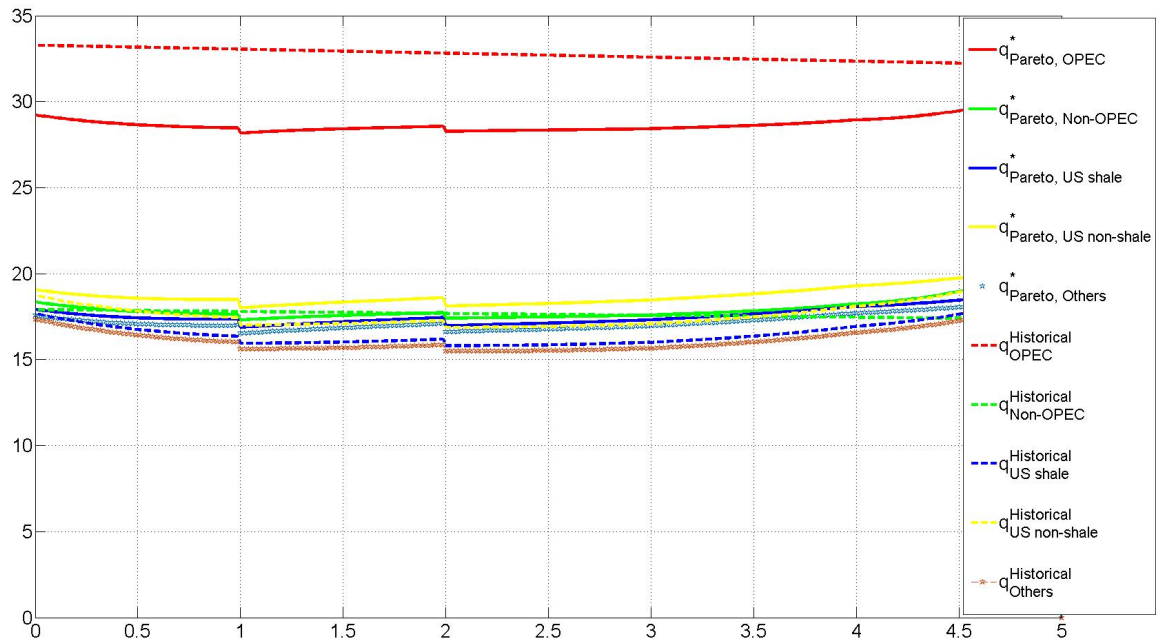


Figure 3.29: Optimal strategies (3.151) under dynamic updating (solid lines) and strategies (3.172) corresponding to historical cooperative agreement (dashed lines).

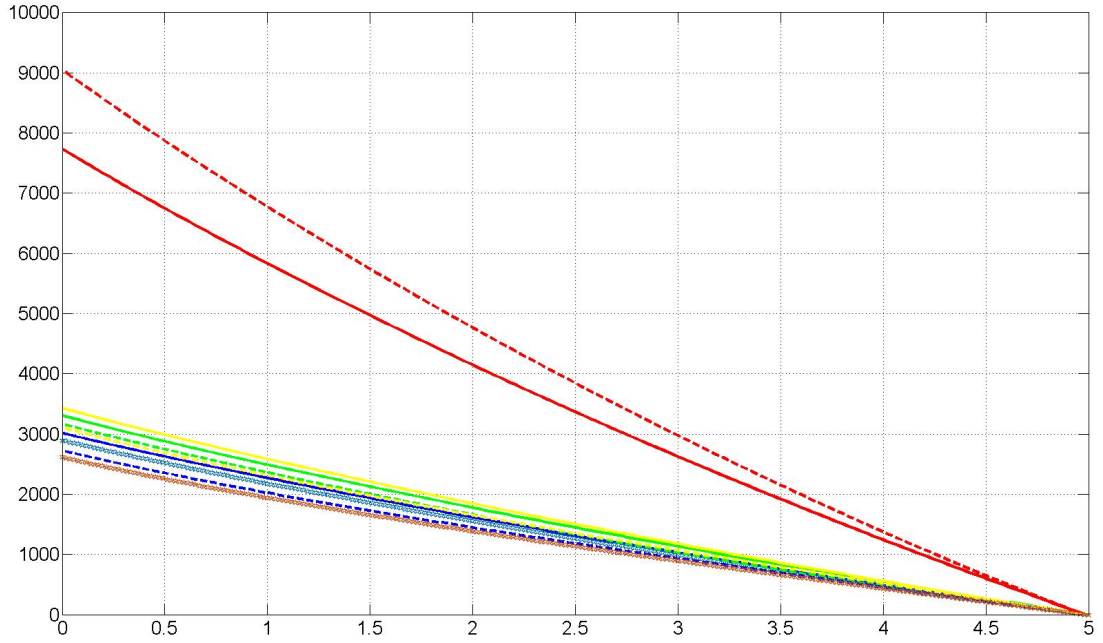


Figure 3.30: $\hat{W}_i(t, \hat{p}^*(t))$, $i \in N$ (solid lines) and payoffs corresponding to historical cooperative agreement (dashed lines).

3.4.4.3 Analysis of Optimality of Corporate Agreement

Generally speaking, the decision to cooperate was justified for the countries under consideration. It can be seen by comparing the cooperative and non-cooperative solutions on the second time interval. In Fig. 12 we can see the conditionally Pareto optimal trajectory $\hat{p}_{\alpha_j}^*(t)$ (solid line) corresponding to the chosen weights $\alpha_j^* \in S_j^{t_0+(j+1)\Delta t}$ in all the truncated subgames $\tilde{\Gamma}_j(p_{j,0}^*, t_0 + \Delta t, t_0 + \Delta t + \bar{T})$, $j = 0, \dots, l$, and the conditionally non-cooperative trajectory of the oil price dynamics $\hat{p}_{NE}(t)$ corresponding to the Nash equilibrium (dashed line).

Nevertheless, the analysis of individual countries' profits confirms the profitability of the agreement only for player 1. In Fig. 13 we can see the payoff of players $\hat{W}_i(t, \hat{p}^*(t))$, $i \in N$ (solid lines), along the trajectory $\hat{p}^*(t)$ and the payoffs of players corresponding to the Nash equilibrium $\hat{V}_i(t, \hat{p}_{NE}(t))$ (dashed lines).

This analysis corresponds to the historical solution. In fact, only the OPEC countries gained benefits. The solution proposed by the model suggests that the countries should reduce oil extraction to an even greater extent. This would lead to even greater rise in prices (Fig.9) and, therefore, to profits for all the market participants. Thus, the countries have rational incentives for such cooperative

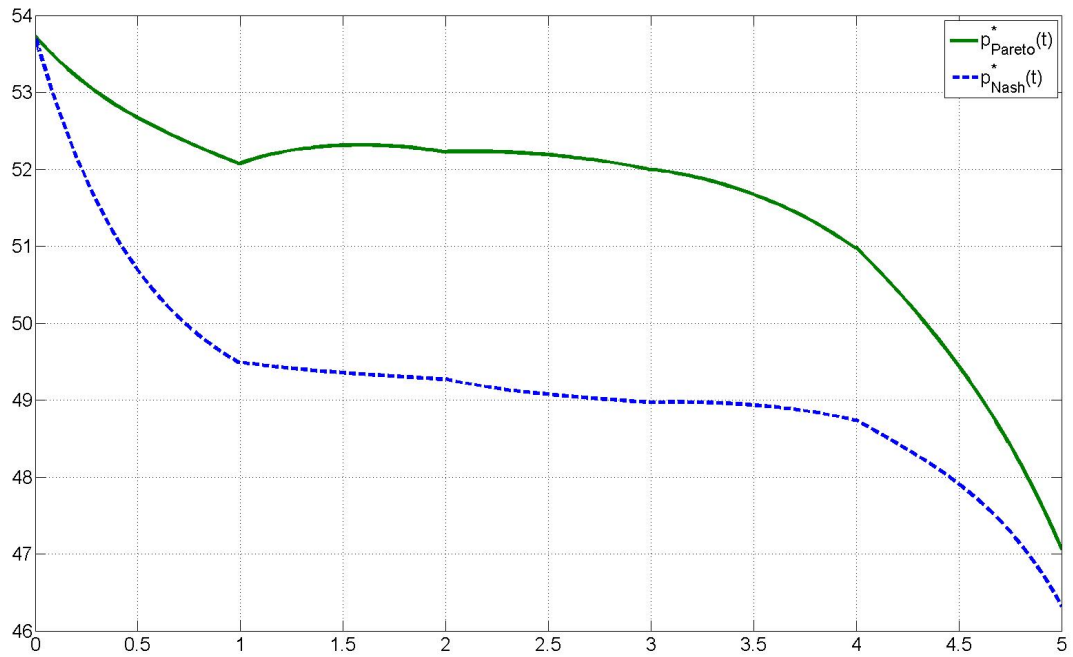


Figure 3.31: Conditionally Pareto optimal trajectory $\hat{p}_{\alpha_j}^*(t)$ (solid line) and conditionally non-cooperative trajectory of oil price $\hat{p}_{NE}(t)$ (dashed line).

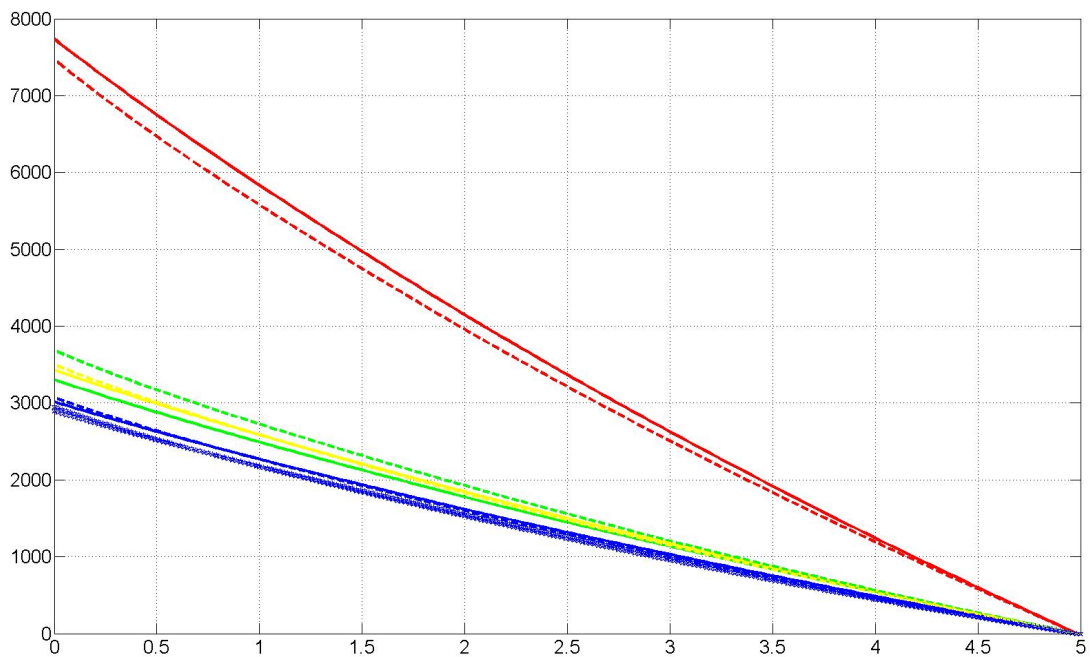


Figure 3.32: $\hat{W}_i(t, \hat{p}^*(t))$, $i \in N$ (solid lines) and payoffs of players corresponding to Nash equilibrium $\hat{V}_i(t, \hat{p}_{NE}(t))$ (dashed lines).

behavior. These incentives for collusion and compliance with the agreements reached stimulated the participants to extend the agreement at the meeting in May 2017.

Chapter 4

Class of Differential Games with Continuous Updating

This chapter is devoted to a new class of differential games with continuous updating. Unlike the previous chapters, it is assumed that at each time instant, players have or use information about the game defined on a closed time interval. However, as the time evolves, information about the game is updated: there is a continuous shift of the time interval that determines the information available to all players. Information about the game includes the motion equations and payoff functions of the players. For this class of games, direct application of classical approaches to determining optimality principles, such as cooperative strategies or Nash equilibrium, is not possible. Chapter 4 studies Hamilton-Jacobi-Bellman equations, Pontryagin's Maximum principle, and conditions for linear-quadratic game models with continuous updating both for cooperative and noncooperative cases.

4.1 Differential Game Model with Continuous Updating

4.1.1 Original Differential Game Model

Consider a differential n -player ($|N| = n$) game with prescribed duration $\Gamma(x_0, T - t_0)$ defined on a time interval $[t_0, T]$.

The motion equation has the form

$$\begin{aligned} \dot{x}(t) &= f(t, x, u), \\ x(t_0) &= x_0, \\ x, x_0 &\in \mathbb{R}^l, \quad u = (u_1, \dots, u_n), \quad u_i = u_i(t, x) \in U_i \subset \text{comp} \mathbb{R}^k, \quad t \in [t_0, T]. \end{aligned} \tag{4.1}$$

The payoff function of player i is defined in the following way:

$$K_i(x_0, T - t_0; u) = \int_{t_0}^T g^i[t, x(t), u(t, x)] dt, \quad i \in N, \tag{4.2}$$

where $g^i[t, x(t), u(t, x)]$ and $f(t, x, u)$ are integrable functions, and $x(t)$ is the solution of the Cauchy problem (4.1) with a fixed $u(t, x) = (u_1(t, x), \dots, u_n(t, x))$. A strategy profile $u(t, x) = (u_1(t, x), \dots, u_n(t, x))$ is called admissible if problem (4.1) has a unique and continuable solution on the interval $[t_0, T]$ (see section 1.1.1). Here and later in the definition of a truncated subgame, we use the conditions of existence, uniqueness and continuability of A. F. Filippov [135].

Using the initial differential game with prescribed duration T , we construct the corresponding differential game with continuous updating.

4.1.2 Game Model with Continuous Updating

Consider an n -player differential game $\Gamma(x_0, t_0, \bar{T})$ defined on a time interval $[t_0, t_0 + \bar{T}]$, where $0 < \bar{T} < +\infty$.

The motion equation has the form

$$\begin{aligned} \dot{x}_{t_0}(s) &= f(s, x_{t_0}, u^{t_0}), \\ x_{t_0}(t_0) &= x_0, \\ x_{t_0} &\in \mathbb{R}^l, \quad u^{t_0} = (u_1^{t_0}, \dots, u_n^{t_0}), \quad u_i^{t_0} = u_i^{t_0}(s, x_{t_0}) \in U_i \subset \text{comp} \mathbb{R}^k. \end{aligned} \tag{4.3}$$

The payoff function of player i is defined in the following way:

$$K_i^{t_0}(x_0, t_0, \bar{T}; u^{t_0}) = \int_{t_0}^{t_0 + \bar{T}} g^i[s, x_{t_0}(s), u^{t_0}(s, x_{t_0})] ds, \quad i \in N, \quad (4.4)$$

where $x_{t_0}(s)$ and $u^{t_0}(s, x)$ are the trajectory and strategies in the game $\Gamma(x_0, t_0, \bar{T})$, and $\dot{x}_{t_0}(s)$ denotes the derivative with respect to s .

Subgame of differential game with continuous updating. Consider an n -player differential game $\Gamma(x, t, \bar{T})$, $t \in [t_0, +\infty)$, defined on the interval $[t, t + \bar{T}]$, where $0 < \bar{T} < +\infty$.

The motion equation of the subgame $\Gamma(x, t, \bar{T})$ has the form

$$\begin{aligned} \dot{x}_t(s) &= f(s, x_t, u^t), \\ x_t(t) &= x, \\ x_t, x &\in \mathbb{R}^l, \quad u^t = (u_1^t, \dots, u_n^t), \quad u_i^t = u_i^t(s, x_t) \in U_i \subset \text{comp} \mathbb{R}^k, \quad s \in [t, t + \bar{T}]. \end{aligned} \quad (4.5)$$

The payoff function of player i in the subgame $\Gamma(x, t, \bar{T})$ has the form

$$K_i^t(x, t; u^t) = \int_t^{t + \bar{T}} g^i[s, x_t(s), u^t(s, x_t)] ds, \quad i \in N, \quad (4.6)$$

where $x_t(s)$ and $u^t(s, x_t)$ are the trajectories and strategies in the game $\Gamma(x, t, \bar{T})$, and $\dot{x}_t(s)$ denotes the derivative with respect to s .

The differential game with continuous updating has the following dynamics: *the current time $t \in [t_0, +\infty)$ changes, and the players continuously obtain new information about the motion equations and payoff functions in the game $\Gamma(x, t, \bar{T})$.*

The strategy profile $u(t, x)$ in the differential game with continuous updating has the form

$$u(t, x) = u^t(s, x)|_{s=t}, \quad t \in [t_0, +\infty), \quad (4.7)$$

where $u^t(s, x)$, $s \in [t, t + \bar{T}]$, is a set of strategies in the subgame $\Gamma(x, t, \bar{T})$.

The trajectory $x(t)$ in the differential game with continuous updating is

determined by

$$\begin{aligned}\dot{x}(t) &= f(t, x, u), \\ x(t_0) &= x_0, \\ x &\in \mathbb{R}^l,\end{aligned}\tag{4.8}$$

where $u = u(t, x)$ is a set of strategies in the game with continuous updating (4.7), and $\dot{x}(t)$ denotes the derivative with respect to t . Suppose that the strategy with continuous updating obtained using (4.7) is admissible, i.e., problem (4.8) has a unique and continuable solution on the interval $[t_0, +\infty)$ (see section 1.1.1). The corresponding conditions of existence, uniqueness and continuability of A. F. Filippov for the system (4.1)-(4.2) were presented in [135].

The essential difference between the game model with continuous updating and the classical differential game with prescribed duration $\Gamma(x_0, T - t_0)$ is as follows: in the latter (original) game, the players are guided by the payoffs they will eventually obtain on the interval $[t_0, T]$; in the game with continuous updating, at a time instant t they orient themselves on the expected payoffs (4.6), which are calculated based on the information defined on the interval $[t, t + \bar{T}]$ (the information they have at the instant t).

4.2 Non-cooperative Differential Game Model with Continuous Updating

4.2.1 Nash Equilibrium in Game with Continuous Updating

Under continuously updated information, it is important to model the behavior of players. To do this, we use the concept of Nash equilibrium in feedback strategies. However, for the class of differential games with continuous updating, we would like to have it in the following form:

- for any fixed $t \in [t_0, +\infty)$, $u^{NE}(t, x) = (u_1^{NE}(t, x), \dots, u_n^{NE}(t, x))$ coincides with the Nash equilibrium in the game (4.5),(4.6) defined on the interval $[t, t + \bar{T}]$ in the instant t .

However, direct application of the classical approaches for defining the Nash equilibrium in feedback strategies is not possible. Consider two intervals $[t, t + \bar{T}]$ and $[t + \epsilon, t + \bar{T} + \epsilon]$, $\epsilon \ll \bar{T}$. According to the problem statement, we have:

- $u^{NE}(t, x)$ at the instant t coincides with the feedback Nash equilibrium in the game defined on the interval $[t, t + \bar{T}]$.
- $u^{NE}(t + \epsilon, x)$ at the instant $t + \epsilon$ must coincide with the feedback Nash equilibrium in the game defined on the interval $[t + \epsilon, t + \epsilon + \bar{T}]$.

To construct such strategies, we consider the concept of generalized Nash equilibrium in feedback (open-loop) strategies as a principle of optimality:

$$\tilde{u}_{fb}^{NE}(t, x; s, x_t) = (\tilde{u}_{fb,1}^{NE}(t, x; s, x_t), \dots, \tilde{u}_{fb,n}^{NE}(t, x; s, x_t)), \quad t \in [t_0, T], \quad s \in [t, t + \bar{T}], \quad (4.9)$$

$$(\tilde{u}_{ol}^{NE}(t, x; s) = (\tilde{u}_{ol,1}^{NE}(t, x; s), \dots, \tilde{u}_{ol,n}^{NE}(t, x; s))), \quad t \in [t_0, T], \quad s \in [t, t + \bar{T}], \quad (4.10)$$

Using this concept, we will construct the strategies $u_{fb}^{NE}(t, x)$ ($u_{ol}^{NE}(t, x)$).

Definition 4.2.1 *A strategy profile $\tilde{u}_{fb}^{NE}(t, x; s, x_t) = (\tilde{u}_{fb,1}^{NE}(t, x; s, x_t), \dots, \tilde{u}_{fb,n}^{NE}(t, x; s, x_t))$ ($\tilde{u}_{ol}^{NE}(t, x; s) = (\tilde{u}_{ol,1}^{NE}(t, x; s), \dots, \tilde{u}_{ol,n}^{NE}(t, x; s))$) is called a feedback-based (open-loop-based) generalized Nash equilibrium with continuous updating if for any fixed $t \in [t_0, +\infty)$, the strategy profile $\tilde{u}_{fb}^{NE}(t, s, x)$ ($\tilde{u}_{ol}^{NE}(t, s, x)$) is a feedback (open-loop) Nash equilibrium in the game $\Gamma(x, t, \bar{T})$.*

Note that the generalized feedback-based (open-loop-based) Nash equilibrium $\tilde{u}_{fb}^{NE}(t, x; s, x_t)$ ($\tilde{u}_{ol}^{NE}(t, x; s)$) for a fixed t is a function of s and x , where s is defined on the interval $[t, t + \bar{T}]$. However, for each fixed value of the parameter t , the strategy profile \tilde{u}^{NE} is a Nash equilibrium in the game $\Gamma(x, t, \bar{T})$ (without updating information). At the same time, we emphasize that in the game $\Gamma(x, t, \bar{T})$ the Nash equilibrium is understood in the sense that the payoff of the player i does not increase when deviating to another admissible strategy profile, which differs from the previous one with only a strategy of player i (that is, not a player changes his strategy independently of other players, but a change in strategy is possible only in the class of admissible strategy profiles). Using generalized Nash equilibrium, we define a solution concept for the game model with continuous updating.

Definition 4.2.2 A strategy profile $u_{fb}^{NE}(t, x)$ ($u_{ol}^{NE}(t, x)$) is called a feedback-based (open-loop-based) Nash equilibrium with continuous updating if it is defined in the following way:

$$\begin{aligned} u_{fb}^{NE}(t, x) &= \tilde{u}_{fb}^{NE}(t, x; s, x_t)|_{s=t} = (\tilde{u}_{fb,1}^{NE}(t, x; s, x_t)|_{s=t}, \dots, \tilde{u}_{fb,n}^{NE}(t, x; s, x_t)|_{s=t}), \\ (u_{ol}^{NE}(t, x) &= \tilde{u}_{ol}^{NE}(t, x; s)|_{s=t} = (\tilde{u}_{ol,1}^{NE}(t, x; s)|_{s=t}, \dots, \tilde{u}_{ol,n}^{NE}(t, x; s)|_{s=t})), \end{aligned} \quad (4.11)$$

where $\tilde{u}_{fb}^{NE}(t, x; s, x_t)$ ($\tilde{u}_{ol}^{NE}(t, x; s)$) is a generalized feedback-based (open-loop-based) Nash equilibrium in the sense of Definition 4.2.1.

A trajectory $x_{fb}^{NE}(t)$ ($x_{ol}^{NE}(t)$) corresponding to a Nash equilibrium with continuous updating $\tilde{u}_{fb}^{NE}(t, x; s, x_t)$ ($\tilde{u}_{ol}^{NE}(t, x; s)$) can be obtained from the system

$$\begin{aligned} \dot{x}(t) &= f(t, x, u_{fb}^{NE}) \quad (\dot{x}(t) = f(t, x, u_{ol}^{NE})) \\ x(t_0) &= x_0, \\ x &\in \mathbb{R}^l. \end{aligned} \quad (4.12)$$

Unlike the generalized feedback-based (open-loop-based) Nash equilibrium, $u_{fb}^{NE}(t, x)$ ($u_{ol}^{NE}(t, x)$) does not contain feedback-based (open-loop-based) Nash equilibrium strategies for any $s \in [t, t+\bar{T}]$. The strategy profile $u_{fb}^{NE}(t, x)$ ($u_{ol}^{NE}(t, x)$) only contains the strategies performed by the players according to the procedure described in Section 4.1.2, i.e., the continuous updating procedure with $s = t$. $u_{fb}^{NE}(t, x)$ ($u_{ol}^{NE}(t, x)$) will be used as a solution concept in the noncooperative game with continuous updating. The notion of Nash equilibrium with continuous updating $u_{fb}^{NE}(t, x)$ ($u_{ol}^{NE}(t, x)$) is close to the subgame perfect Nash equilibrium. In the sense of the expected payoff (4.6), for any fixed current time t , an individual deviation from a Nash equilibrium with continuous updating is not beneficial for the players, i.e., for any subgame of the game with continuous updating, an individual deviation is not beneficial due to the information structure.

4.2.2 Hamilton-Jacobi-Bellman Equations with Continuous Updating

In this section, we will only consider the feedback-based Nash equilibrium with continuous updating $u_{fb}^{NE}(t, x)$. Therefore, we will use the simplified notation

$u^{NE}(t, x)$.

To define the strategy profile $u^{NE}(t, x)$, it is necessary to determine the generalized Nash equilibrium in feedback strategies $\tilde{u}^{NE}(t, x; s, x_t)$ in the game with continuous updating $\Gamma(x_0, t_0, \bar{T})$. To do this, we will use a modernized version of dynamic programming. Within this approach, the Bellman function $V^i(t; s, x)$ is defined as the payoff of player i in the feedback Nash equilibrium $\tilde{u}^{NE}(t, x; s, x_t)$ in the subgame starting at the instant s from the state x in the game defined on the interval $[t, t + \bar{T}]$:

$$V^i(t; s, x) = \int_s^{t+\bar{T}} g^i[\tau, x_t(\tau), \tilde{u}^{NE}(t, \tau, x_t(\tau))] d\tau, \quad i \in N \quad (4.13)$$

при условии

$$\begin{aligned} \dot{x}_t(\tau) &= f(\tau, x, \tilde{u}^{NE}), \\ x_t(s) &= x. \end{aligned} \quad (4.14)$$

The following theorem holds.

Theorem 4.2.1 $\tilde{u}^{NE}(t, x; s, x_t)$ is a generalized Nash equilibrium in feedback strategies in the differential game with continuous updating $\Gamma(x_0, t_0, \bar{T})$ if there exist functions $V^i(t; s, x) : [t_0, +\infty) \times [t, t + \bar{T}] \times R \rightarrow R$, $i \in N$, continuously differentiable with respect to s and x , satisfying the system of partial differential equations

$$\begin{aligned} -V_s^i(t; s, x) &= \max_{\phi_i} \{g^i(s, x, \tilde{u}_{-i}^{NE}) + V_x^i(t; s, x)f(s, x, \tilde{u}_{-i}^{NE})\} = \\ &= g^i(s, x, \tilde{u}^{NE}) + V_x^i(t; s, x)f(s, x, \tilde{u}^{NE}), \\ V^i(t; t + \bar{T}, x) &= 0, \quad i \in N, \end{aligned} \quad (4.15)$$

$$\partial_e \tilde{u}_{-i}^{NE}(\phi_i) = (\tilde{u}_1^{NE}, \dots, \phi_i, \dots, \tilde{u}_n^{NE}).$$

Proof According to the definition of generalized Nash equilibrium, $\tilde{u}^{NE}(t, x; s, x_t)$ should be the feedback Nash equilibrium for any fixed t .

Fixing t in the statement of Theorem 4.2.1 (particularly, in (4.15)), we obtain the classical sufficient conditions for feedback Nash equilibrium in the differential game with prescribed duration $[t, t + \bar{T}]$ presented in [80]. Therefore, for any fixed t , the conditions required by the definition of generalized Nash equilibrium are satisfied. The theorem is proved. \square

If it is possible to obtain a generalized Nash equilibrium $\tilde{u}^{NE}(t, x; s, x_t)$ using equations (4.15), then the procedure (4.11) yields the desired strategy profile $u^{NE}(t, x)$.

4.2.2.1 Differential Game of Investment in Public Goods

As an illustrative example, consider a two-player differential game of investment in public goods. Two players are investing in a public stock of knowledge. It is assumed that knowledge is a pure public good, and every individual has access to it. The model of such type was first formulated in [218]; also, see [50; 219]. In these models, it was assumed that each player has a quadratic utility, and the cost of investment increases quadratically with the investment effort. Below, we will use other functions to define players' payoffs.

4.2.2.1.1 Initial Game Model

Let $x(t)$ be the stock of knowledge at a time t , and $u_i(t, x)$ be the investment of player i in public knowledge at the time t . Assume that the stock of knowledge evolves according to the accumulation equation

$$\dot{x}(t) = -\beta x(t) + u_1(t, x) + u_2(t, x), \quad x(0) = x_0, \quad (4.16)$$

where β is the depreciation rate, and $x_0 > 0$ is a given initial stock of knowledge. Let possible states at each time t satisfy the condition $x(t) > 0$. Assume that each individual i bears costs of investment, which are state-dependent and described by the function $\frac{u_i^2}{x}$. Suppose that the utility of each player is a linear function of the knowledge stock consumption: $q_i x(t)$, where $q_i > 0$. That is, the payoff function of

each player has the form

$$K_i(x_0, T; u) = \int_0^T \left(q_i x(t) - \frac{u_i^2(t, x)}{x(t)} \right) dt. \quad (4.17)$$

To define the feedback Nash equilibrium $u^*(t, x)$, we use the sufficient conditions in the form of HJB equations:

$$\begin{aligned} -V_t^i(t, x) &= \max_{u_i} \left(q_i x - \frac{u_i^2}{x} + V_x^i(t, x)(-\beta x + u_i + u_j^*) \right), \quad i \neq j, \\ V_i(T, x) &= 0, \quad i = 1, 2. \end{aligned} \quad (4.18)$$

Suppose that the Bellman function has the following form:

$$V^i(t, x) = a_i(t)x, \quad i = 1, 2.$$

The maximization problem in (4.18) yields a strategy of player i :

$$u_i^*(t, x) = \frac{xV_x^i(t, x)}{2} = \frac{a_i(t)x}{2}. \quad (4.19)$$

Substituting (4.19) into (4.18), we obtain the following system of differential equations for $a_i(t)$:

$$\begin{aligned} \dot{a}_i(t) &= \beta a_i(t) + \frac{a_i^2(t)}{4} - a_i(t)a_j(t) - q_i, \quad i \neq j, \quad i = 1, 2, \\ a_i(T) &= 0. \end{aligned} \quad (4.20)$$

As an example consider the symmetric case $q_1 = q_2 = q$. Then the solution of (4.20) is

$$a_i(t) = \frac{2q(e^{v(t-T)} - 1)}{(\beta - v)e^{v(t-T)} - \beta - v}, \quad i = 1, 2,$$

where $v = \sqrt{\beta^2 - 3q}$ (by assumption, $\beta^2 > 3q$). The solution of (4.20) is valid on the interval $(-\infty; T + \frac{1}{v} \ln \frac{\beta+v}{\beta-v})$. Note that $[0; T] \subset (-\infty; T + \frac{1}{v} \ln \frac{\beta+v}{\beta-v})$. Then the feedback Nash equilibrium strategies are

$$u_i^*(t, x) = \frac{q(e^{v(t-T)} - 1)}{(\beta - v)e^{v(t-T)} - \beta - v} x(t), \quad i = 1, 2.$$

The resulting equilibrium trajectory is

$$x^*(t) = x_0 e^{-\frac{1}{3}(\beta+2v)t} \left(\frac{(\beta-v)e^{vt} - (\beta+v)e^{vT}}{\beta-v - (\beta+v)e^{vT}} \right)^{\frac{4}{3}}.$$

The state constraint $x^*(t) > 0$ is satisfied for all $t \in [0; T]$ under $x_0 > 0$.

4.2.2.1.2 Game Model with Continuous Updating

Now consider this model as a game with continuous updating. It is assumed that information about the motion equations and payoff functions is updated continuously in time. At every instant $t \in [0, +\infty)$, the players have information only the interval $[t, t + \bar{T}]$. This means that due to possible changes in the number of investors or their utilities, at each time instant they can count for the process stability only over the period \bar{T} .

According to Theorem 4.2.1, the Hamilton-Jacobi-Bellman equations for generalized feedback Nash equilibrium are

$$\begin{aligned} -V_s^i(t; s, x) &= \max_{u_i} \left(q_i x - \frac{u_i^2}{x} + V_x^i(t; s, x)(-\beta x + u_i + \tilde{u}_j^{NE}) \right), \quad i \neq j, \\ V_i(t; t + \bar{T}, x) &= 0, \quad i = 1, 2. \end{aligned} \quad (4.21)$$

Suppose that the Bellman function is defined in the form

$$V_i(t; s, x) = a_i(t, s)x.$$

The maximization problem in (4.21) yields a strategy of player i :

$$\tilde{u}_i^{NE}(t, x; s, x_t) = \frac{xV_x^i(t; s, x)}{2} = \frac{xa_i(t, s)}{2}. \quad (4.22)$$

Substituting (4.22) into (4.21), we obtain the following system of differential equations for $a_i(t, s)$:

$$\begin{aligned} \dot{a}_i(t, s) &= \beta a_i(t, s) + \frac{a_i^2(t, s)}{4} - a_i(t, s)a_j(t, s) - q_i, \quad i \neq j, \quad i = 1, 2, \\ a_i(t, t + \bar{T}) &= 0. \end{aligned} \quad (4.23)$$

Taking into account the symmetry of players, we obtain the solution

$$a_i(t, s) = \frac{2q(e^{v(s-t-\bar{T})} - 1)}{(\beta - v)e^{v(s-t-\bar{T})} - \beta - v}, \quad i = 1, 2,$$

where $v = \sqrt{\beta^2 - 3q}$.

Finally we get the generalized feedback Nash equilibrium strategies

$$\tilde{u}_i^{NE}(t, x; s, x_t) = \frac{q(e^{v(s-t-\bar{T})} - 1)}{(\beta - v)e^{v(s-t-\bar{T})} - \beta - v} x(t).$$

Following the procedure (4.11), we construct the feedback Nash equilibrium with continuous updating

$$u_i^{NE}(t, x) = \tilde{u}_i^{NE}(t, x; s, x_t) \Big|_{s=t} = \frac{q(e^{-v\bar{T}} - 1)}{(\beta - v)e^{-v\bar{T}} - \beta - v} x(t).$$

The equilibrium trajectory $x^{NE}(t)$ with continuous updating is

$$x^{NE}(t) = x_0 e^{\left(-\beta + \frac{2q(e^{-v\bar{T}} - 1)}{(\beta - v)e^{-v\bar{T}} - \beta - v}\right)t}.$$

The state constraint $x^{NE}(t) > 0$ is satisfied for all $t \in [0; T]$ under $x_0 > 0$.

Figures 4.1, 4.2 present the solutions obtained for the following parameter values:

$$\beta = 1/2, \quad q = 1/15, \quad T = 10, \quad \bar{T} = 1/2, \quad x_0 = 5.$$

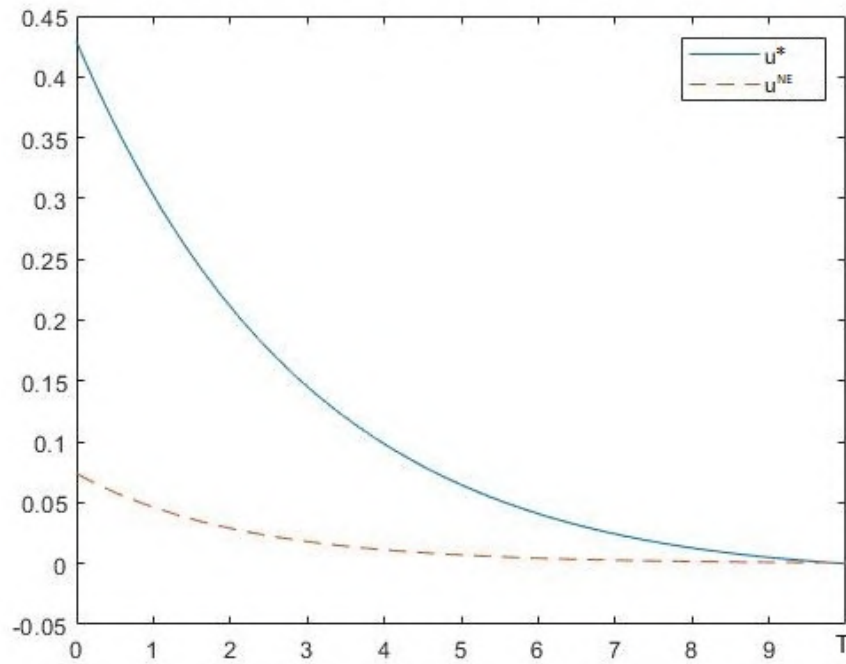


Figure 4.1: Nash equilibrium with continuous updating $u^{NE}(t, x)$ (red dashed line) and Nash equilibrium strategies in original game $u^*(t)$ (blue solid line).

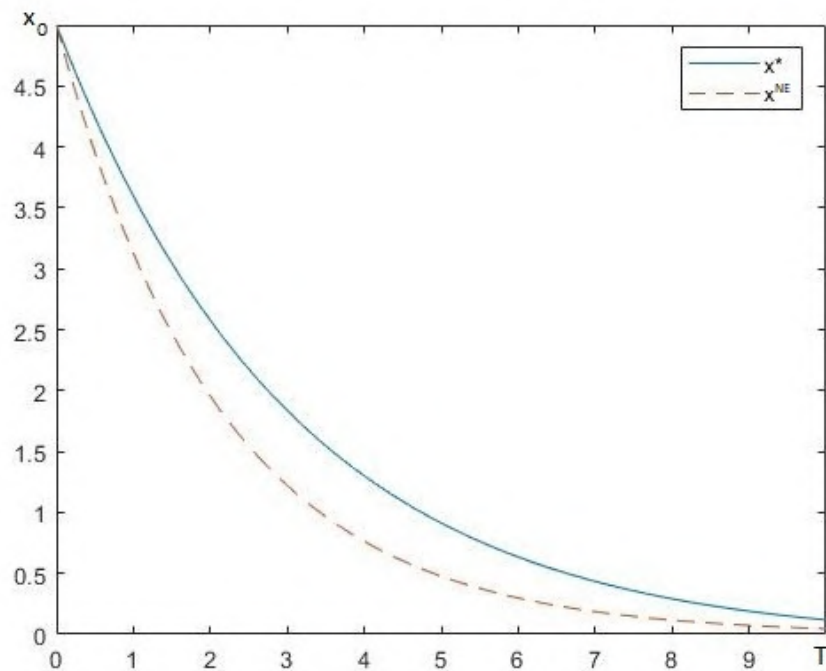


Figure 4.2: Resulting equilibrium trajectory $x^{NE}(t)$ with continuous updating (red dashed line) and resulting equilibrium trajectory $x^*(t)$ in original game (blue solid line).

The numerical simulation results show that in the case of continuous updating, the players are more cautious: their investments are less than in the original game (Figure 4.1). As a result, the knowledge stock is reduced compared to the original game (Figure 4.2).

4.2.3 Pontryagin's Maximum Principle with Continuous Updating

In this section, we will only consider the open-loop-based Nash equilibrium with continuous updating $u_{ol}^{NE}(t, x)$. Therefore, we will use the simplified notation $u^{NE}(t, x)$.

To define the strategy profile $u^{NE}(t, x)$, it is necessary to determine the generalized Nash equilibrium in open-loop strategies $\tilde{u}^{NE}(t, s, x)$ in the game with continuous updating. To do this, we will use a modernized version of Pontryagin's maximum principle. Let us start by defining a real-valued function of the form, $t \in [t_0, +\infty)$, $\tau \in [t, t + \bar{T}]$

$$H_i^t(\tau, x^t, u^t, \lambda^t) = \bar{T}g^i(\bar{T}\tau + t, x^t, u^t) + \lambda_i^t \bar{T}f(\bar{T}\tau + t, x^t, u^t). \quad (4.24)$$

The function H_i^t , $i \in N$ is called the (current-value) Hamiltonian function and plays a prominent role in Pontryagin's Maximum Principle. The variable λ_i^t is called the (current-value) costate variable associated with the state variable x^t , or the (current-value) adjoint variable. Note that like in the Hamilton-Jacobi-Bellman equations, the function H_i^t explicitly depends on the current time t .

By analogy with the case of Hamilton-Jacobi-Bellman equations, it is possible to derive necessary conditions for the generalized Nash equilibrium with continuous updating in the following form:

Theorem 4.2.2 *Let $f(s, \cdot, u^t)$ be continuously differentiable on R^l , $\forall s \in [t, t + \bar{T}]$, and $g^i(s, \cdot, u^t)$ be continuously differentiable on R^l , $\forall s \in [t, t + \bar{T}]$, $i \in N$. If $\tilde{u}^{NE}(t, s, x)$ provides a generalized open-loop Nash equilibrium in the differential game with continuous updating, and for all $t \in [t_0, +\infty)$ $\tilde{x}^t(s)$ with $s \in [t, t + \bar{T}]$ is the corresponding state trajectory in the game $\Gamma(x, t, t + \bar{T})$, then for all $t \in [t_0, +\infty)$ there exist n costate functions $\lambda_i^t(\tau, x)$, where $\tau \in [0, 1]$, $i \in N$, such that the*

following relations are satisfied:

1 for all $\tau \in [0, 1]$

$$H_i^t(\tau, \tilde{x}^t, \tilde{u}^{NE}(t, \tau, x), \lambda^t) = \max_{\phi_i} \{H_i^t(\tau, \tilde{x}^t, \tilde{u}_{-i}^{NE}(t, \tau, x), \lambda^t)\}, i \in N, \quad (4.25)$$

where $\tilde{u}_{-i}^{NE} = (\tilde{u}_1^{NE}, \dots, \phi_i, \dots, \tilde{u}_n^{NE})$,

2 $\lambda_i^t(\tau, x)$ is the solution of the system of adjoint equations

$$\begin{aligned} \frac{d\lambda_i^t(\tau, x)}{d\tau} &= -\frac{\partial H_i^t(\tau, \tilde{x}^t(\tau), \tilde{u}^{NE}(t, \tau, x), \lambda^t)}{\partial x^t} = \\ &= -\bar{T} \frac{\partial g^i(\bar{T}\tau + t, \tilde{x}^t, \tilde{u}^{NE})}{\partial x^t} - \lambda_i^t(\tau, x) \bar{T} \frac{\partial f(\bar{T}\tau + t, \tilde{x}^t, \tilde{u}^{NE})}{\partial x^t}, \quad i \in N, \end{aligned} \quad (4.26)$$

where the transversality conditions are

$$\lambda_i^t(1, x) = 0, \quad i \in N, \quad (4.27)$$

3 for all $t \in [t_0, +\infty)$

$$\dot{\tilde{x}}^t(\tau) = \bar{T} f(\bar{T}\tau + t, \tilde{x}^t, \tilde{u}^{NE}), \quad \tilde{x}^t(0) = x, \quad \tau \in [0, 1]. \quad (4.28)$$

Proof We fix $t \geq t_0$ and consider the game $\Gamma(x, t, t + \bar{T})$. Using the change of variables

$$\tau = \frac{s - t}{\bar{T}}, \quad (4.29)$$

we reduce the motion equation (4.14) to

$$\dot{x}^t(\tau) = \bar{T} f(\bar{T}\tau + t, x^t, u^t), \quad x^t(0) = x, \quad \tau \in [0, 1]. \quad (4.30)$$

The payoff function of player $i \in N$ has the form

$$K_i^t(x, t, \bar{T}; u^t) = \int_0^1 \bar{T} g^i[\bar{T}\tau + t, x^t(\tau), u^t(\tau, x)] d\tau, \quad i \in N. \quad (4.31)$$

For the optimization problem (4.30)-(4.31), the Hamiltonian has the form

$$H_i^t(\tau, x^t, u^t, \lambda^t) = \bar{T}g^i(\bar{T}\tau + t, x^t(\tau), u^t(\tau, x)) + \lambda_i^t(\tau, x)\bar{T}f(\bar{T}\tau + t, x^t(\tau), u^t(\tau, x)). \quad (4.32)$$

If $\tilde{u}^{NE}(t, \tau, x)$ is a generalized open-loop Nash equilibrium in the differential game with continuous updating, then by Definition 4.2.1, for every fixed $t \geq t_0$, $\tilde{u}^{NE}(t, \tau, x)$ is an open-loop Nash equilibrium in the game $\Gamma(x, t, t + \bar{T})$. Therefore, for any fixed $t \geq t_0$, conditions 1-3 of the theorem are satisfied as necessary conditions for Nash equilibrium in open-loop strategies; see [220]. The theorem is proved. \square

Also, note that if for every $t \geq t_0$ the functions H_i^t are concave in (x_t, u_t) for all $i \in N$, then the conditions of this theorem are sufficient for generalized open-loop Nash equilibrium [221].

4.2.3.1 Differential Game of Logarithmic Advertising Game Model with Continuous Updating

As an illustrative example, we consider the logarithmic excess-advertising model of a duopoly proposed by Jørgensen in [222]. There are two firms (players) operating in a market. It is assumed that the market potential is constant over time. The only marketing instrument used by the players is advertising. Advertising has diminishing returns since it suffers from increasing marginal costs. The Nash optimal open-loop advertising strategies were determined in [222]. Below, we obtain open-loop Nash equilibrium with continuous updating using Theorem 4.2.2.

4.2.3.1.1 Original Game Model

Consider the model investigated in [222]. Let $x_i(t)$ denote the rate of sales of firm (player) i at a time instant t , ($i = 1, 2$). Assume that $x_1 + x_2 = M$, implying that the market potential is fully exhausted by the players at each time instant. The game is played on an interval $[0, T]$, where T is an arbitrary fixed positive number.

Due to the assumption $x_1 + x_2 = M$, $\dot{x}_2 = -\dot{x}_1$. The state equation is

$$\begin{aligned}\dot{x}_1 &= k \log \frac{u_1}{u_2} = k(\log u_1 - \log u_2), \\ \dot{x}_2 &= -\dot{x}_1 = k(\log u_2 - \log u_1), \\ x_1(0) &= x_1^0, x_2(0) = x_2^0,\end{aligned}\tag{4.33}$$

where k is a positive constant, and $x_i(0)$ is a given initial rate of sales of player i . The state equation (4.33) describes a market where buyers are perfectly mobile and switch instantaneously to the firm which has the largest rate of advertising expenditure (i.e., advertises in excess of the others). In the model, the market share increases linearly according to the amount of excess advertising. The performance indices are given by

$$K_i = \int_0^T (\varphi_i x_i - u_i) \exp\{-r_i t\} dt, \quad i = 1, 2,\tag{4.34}$$

where $x_2 = M - x_1$. Assume that $r_i > 0$, $i = 1, 2$. The open-loop Nash equilibrium in explicit form was constructed in [222]:

$$u_i^{initial, NE} = \frac{k\varphi_i}{r_i} [1 - \exp\{-r_i(T - t)\}].\tag{4.35}$$

In the case $r_1 = r_2$, the optimal trajectories are given by

$$\begin{aligned}x_1(t) &= (k \log \frac{\varphi_1}{\varphi_2})t + x_1(0), \\ x_2(t) &= M - (k \log \frac{\varphi_1}{\varphi_2})t - x_1(0).\end{aligned}\tag{4.36}$$

If $r_1 \neq r_2$, then the trajectory x_1 is the solution of

$$\dot{x}_1 = k \log \frac{\varphi_1 r_2 [1 - \exp\{-r_1(T - t)\}]}{\varphi_2 r_1 [1 - \exp\{-r_2(T - t)\}]}.\tag{4.37}$$

The solution of equation (4.37) is given by

$$x_1(t) = x_1^0 + k \log \frac{\varphi_1 r_2}{\varphi_2 r_1} t + k \int_0^t \log \frac{1 - \exp\{-r_1(T - s)\}}{1 - \exp\{-r_2(T - s)\}} ds.$$

4.2.3.1.2 Game Model with Continuous Updating

Now consider this model as a game with continuous updating. It is assumed that information about the motion equations and payoff functions is updated continuously in time. At every time instant $t \in [0, +\infty)$, the players have information only on an interval $[t, t + \bar{T}]$.

Therefore, for every time instant t , we can obtain the payoff function of player i for the interval $[t, t + \bar{T}]$:

$$K_i^t = \int_t^{t+\bar{T}} (\varphi_i x_i^t - u_i^t) \exp(-r_i s) ds, \quad i = 1, 2.$$

To simplify the problem, it is possible to use the change of variables $\tau = \frac{s-t}{\bar{T}}$:

$$\begin{aligned} \dot{x}_1^t(\tau) &= \bar{T}k \log \frac{u_1^t(\tau, x)}{u_2^t(\tau, x)} = \bar{T}k(\log u_1^t(\tau) - \log u_2^t(\tau)), \quad \tau \in [0, 1], \\ \dot{x}_2^t(\tau) &= -\dot{x}_1^t(\tau), \\ x_1^t(0) &= x_1, \quad x_2^t(0) = x_2, \\ K_i^t &= \int_0^1 \bar{T}(\varphi_i x_i^t(\tau) - u_i^t(\tau, x)) \exp\{-r_i(\bar{T}\tau + t)\} d\tau, \quad i = 1, 2. \end{aligned} \tag{4.38}$$

Then the Hamiltonian functions are given by

$$\begin{aligned} H_1^t(t, \tau, x, u^t, \lambda^t) &= (\varphi_1 x_1^t - u_1^t) \bar{T} + \lambda_1^t(\tau, x) \bar{T}k(\log u_1^t - \log u_2^t), \\ H_2^t(t, \tau, x, u^t, \lambda^t) &= (\varphi_2 x_2^t - u_2^t) \bar{T} - \lambda_2^t(\tau, x) \bar{T}k(\log u_1^t - \log u_2^t). \end{aligned}$$

Note that the current-value Hamiltonian is simply $\exp(r_i(\bar{T}\tau + t))$ times the conventional Hamiltonian. Necessary conditions for maximizing H_i^t over $u_i^t \in (0, +\infty)$ are given by

$$\begin{aligned} \frac{\partial H_1^t}{\partial u_1^t} &= -\bar{T} + \lambda_1^t(\tau, x) \bar{T}k \frac{1}{u_1^t} = 0, \\ \frac{\partial H_2^t}{\partial u_2^t} &= -\bar{T} + \lambda_2^t(\tau, x) \bar{T}k \frac{1}{u_2^t} = 0. \end{aligned}$$

Therefore, the optimal control u_i^t is given by

$$u_1^t(\tau, x) = \lambda_1^t(\tau, x)k, \quad u_2^t(\tau, x) = \lambda_2^t(\tau, x)k. \quad (4.39)$$

The adjoint variables $\lambda_i^t(\tau)$ should satisfy the following equations:

$$\begin{aligned} \dot{\lambda}_1^t(\tau, x) &= -\frac{\partial H_1^t}{\partial x_1^t} + \bar{T}r_1\lambda_1^t(\tau, x) = -\varphi_1\bar{T} + \bar{T}r_1\lambda_1^t(\tau, x), \\ \dot{\lambda}_2^t(\tau, x) &= -\frac{\partial H_2^t}{\partial x_2^t} + \bar{T}r_2\lambda_2^t(\tau, x) = -\varphi_2\bar{T} + \bar{T}r_2\lambda_2^t(\tau, x). \end{aligned} \quad (4.40)$$

Note that these equations are uncoupled, and the transversality conditions are

$$\lambda_i^t(1, x) = 0, \quad i = 1, 2.$$

Solving the differential equations (4.40), we obtain the solutions

$$\begin{aligned} \lambda_1^t(\tau, x) &= \frac{\varphi_1}{r_1}[1 - \exp\{\bar{T}r_1(\tau - 1)\}], \\ \lambda_2^t(\tau, x) &= \frac{\varphi_2}{r_2}[1 - \exp\{\bar{T}r_2(\tau - 1)\}]. \end{aligned} \quad (4.41)$$

Substituting (4.41) into (4.39) yields

$$u_i^{tNE}(\tau, x) = \frac{k\varphi_i}{r_i}[1 - \exp\{\bar{T}r_i(\tau - 1)\}]. \quad (4.42)$$

Note that x is the initial state in the subgame $\Gamma(x, t, t + \bar{T})$. In this example, the open-loop strategies $u_i^{tNE}(\tau, x)$ do not depend on the initial state x .

Let us show that the obtained solution satisfies sufficiency conditions. Since $\frac{\partial^2 H_i^t}{\partial x^t \partial x^t} = 0$, $\frac{\partial^2 H_i^t}{\partial x^t \partial u_i^t} = 0$, and $\frac{\partial^2 H_i^t}{\partial u_i^t \partial u_i^t} = -\lambda_i^t(\tau)\bar{T}k\frac{1}{(u_i^t)^2} \leq 0$, the strategy profile $u^{tNE}(\tau, x)$ is indeed a Nash equilibrium in the subgame $\Gamma(x, t, t + \bar{T})$ according to [223].

Finally, we convert τ to t, s . Then the generalized open-loop Nash equilibrium

strategies become

$$\begin{aligned}\tilde{u}_1^{NE}(t, s, x) &= \frac{k\varphi_1}{r_1}[1 - \exp\{r_1(s - t - \bar{T})\}], \\ \tilde{u}_2^{NE}(t, s, x) &= \frac{k\varphi_2}{r_2}[1 - \exp\{r_2(s - t - \bar{T})\}].\end{aligned}\tag{4.43}$$

According to Definition 4.2.2, we construct the open-loop-based Nash equilibrium with continuous updating

$$u_i^{NE}(t, x) = \tilde{u}_i^{NE}(t, s, x)|_{s=t} = \frac{k\varphi_i}{r_i}[1 - \exp\{-r_i\bar{T}\}] \quad i = 1, 2.\tag{4.44}$$

Note that in the example under consideration, the strategies u_i^{tNE} are independent of the initial values of the state variables in the subgame $\Gamma(x, t, t + \bar{T})$, so the strategies $u_i^{NE}(t, x)$ actually do not depend on x .

Consider the difference between the optimal strategies in the original game and the game with continuous updating:

$$u_i^{initial, NE} - u_i^{NE} = \frac{k\varphi_i}{r_i} \exp\{-r_i\bar{T}\} [1 - \exp\{-r_i(T - t - \bar{T})\}].$$

We can see that the players' advertising expenditure is less in the game with continuous updating for $t < T - \bar{T}$, which can be explained as follows. In the continuous updating scenario, the players do not behave rationally, taking into account the information available on the entire time interval.

The corresponding Nash equilibrium trajectories $x_1^{NE}(t)$, $x_2^{NE}(t)$ in the game with continuous updating are the solutions of

$$\begin{aligned}\dot{x}_1(t) &= k \log \left(\frac{\varphi_1 r_2 [1 - \exp\{-r_1\bar{T}\}]}{r_1 \varphi_2 [1 - \exp\{-r_2\bar{T}\}]} \right), \\ \dot{x}_2(t) &= -\dot{x}_1(t), \\ x_1(0) &= x_1^0, \\ x_2(0) &= x_2^0,\end{aligned}\tag{4.45}$$

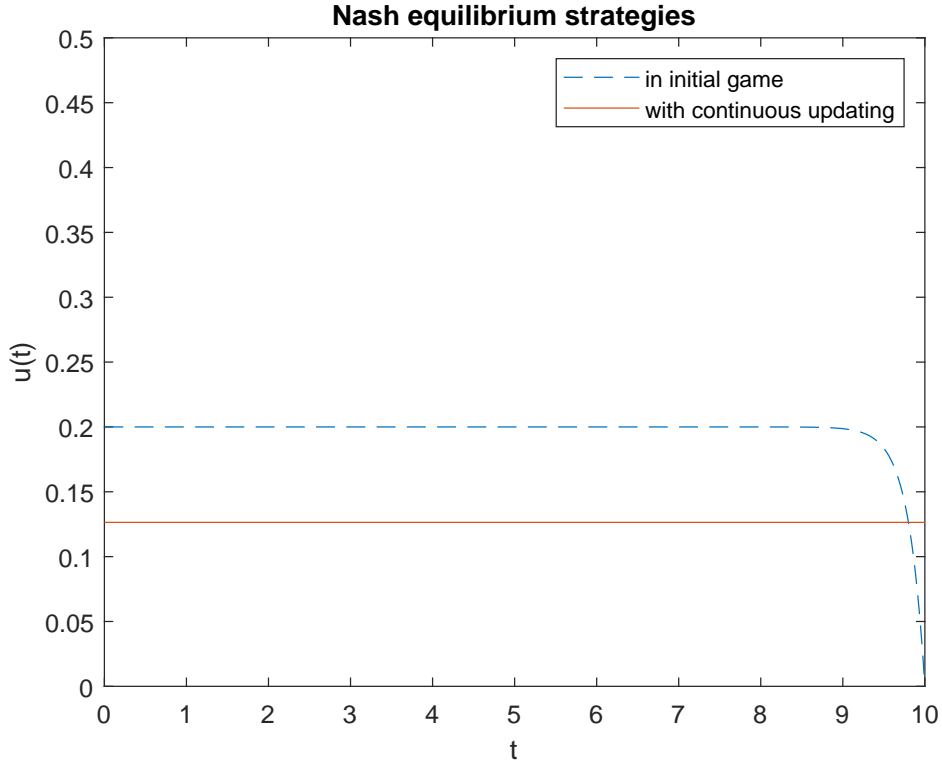


Figure 4.3: Comparison of Nash equilibrium strategies in the original model and game with continuous updating.

where $r_i > 0, i = 1, 2$. Therefore, the state dynamics of the system are

$$\begin{aligned} x_1^{NE}(t) &= x_1^0 + k \log\left(\frac{\varphi_1 r_2 [1 - \exp\{-r_1 \bar{T}\}]}{r_1 \varphi_2 [1 - \exp\{-r_2 \bar{T}\}]}\right)t, \\ x_2^{NE}(t) &= M - x_1^0 - k \log\left(\frac{\varphi_2 r_1 [1 - \exp\{-r_2 \bar{T}\}]}{r_2 \varphi_1 [1 - \exp\{-r_1 \bar{T}\}]}\right)t. \end{aligned} \quad (4.46)$$

Note that if $r_1 = r_2$, the optimal trajectories in the original game and the game with continuous updating are the same. Figures 4.3 and 4.4 represent a comparison of the results obtained in the original model and the model with continuous updating for the following parameter values: $\frac{\varphi_1}{r_1} = 0.1$, $\frac{\varphi_2}{r_2} = 0.5$, $k = 1$, $T = 10$, $\bar{T} = 0.2$, $r_1 = 5$, $r_2 = 3$, $x_1^0 = 8$, $x_2^0 = 10$.

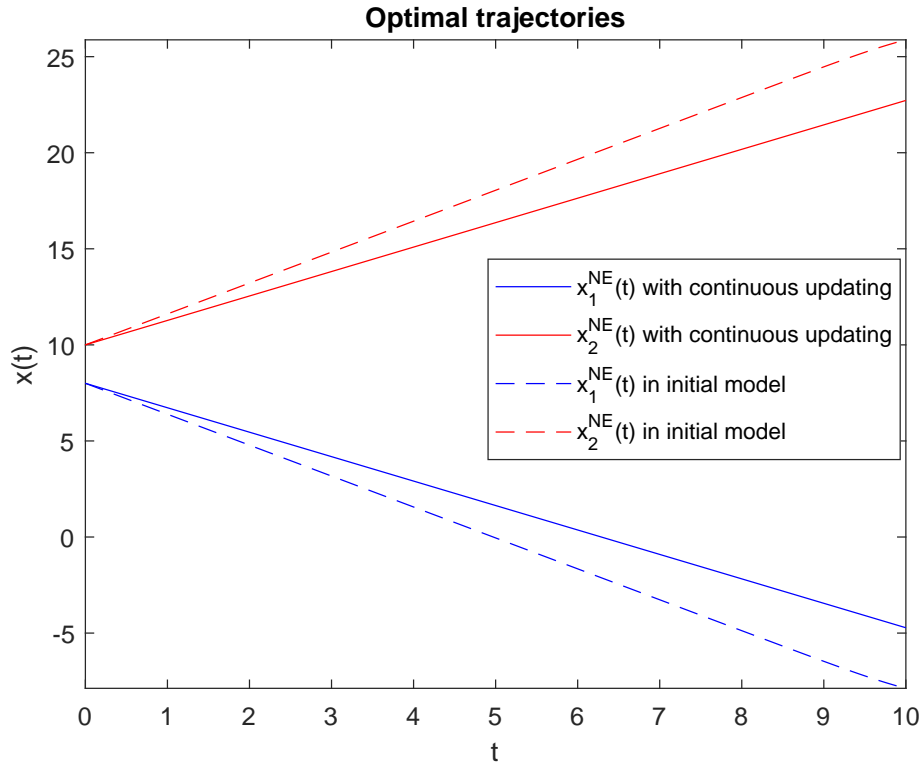


Figure 4.4: Comparison of optimal trajectories in original model and game with continuous updating.

We see that the rate of sales for player 1 in the game with continuous updating is less than in the original model. This occurs because in the continuous updating scenario, the players are not rational in the sense of the information available for the entire time interval.

4.2.4 Autonomous Linear-Quadratic Case with Continuous Updating

In this section, we will only consider the feedback-based Nash equilibrium with continuous updating $u_{fb}^{NE}(t, x)$. Therefore, we will use the simplified notation $u^{NE}(t, x)$. Below, a special class of differential games with continuous updating is considered, where the motion equations are linear and the payoff function of each player is a quadratic functional.

4.2.4.1 Original Linear-Quadratic Autonomous Game Model

Consider an n -player ($|N| = n$) linear-quadratic differential game $\Gamma(x_0, T - t_0)$ defined on a time interval $[t_0, T]$.

The motion equations have the form

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B_1 u_1(t, x) + \dots + B_n u_n(t, x), \\ x(t_0) &= x_0, \\ x &\in \mathbb{R}^l, \quad u = (u_1, \dots, u_n), \quad u_i = u_i(t, x) \in U_i \subset \text{comp} \mathbb{R}^k, \quad t \in [t_0, T]. \end{aligned} \quad (4.47)$$

The payoff function of player $i \in N$ is defined as

$$K_i(x_0, t_0, T; u) = \int_{t_0}^T \left(x'(t) Q_i x(t) + \sum_{j=1}^n u_j'(t, x) R_{ij} u_j(t, x) \right) dt, \quad i \in N, \quad (4.48)$$

where Q_i and R_{ij} are assumed to be symmetric, R_{ii} is positive definite, and $(\cdot)'$ denotes transpose.

4.2.4.2 Linear-Quadratic Autonomous Game Model with Continuous Updating

Consider an n -player differential game $\Gamma(x, t, \bar{T})$, $t \in [t_0, +\infty)$, defined on an interval $[t, t + \bar{T}]$, where $0 < \bar{T} < +\infty$.

The motion equations of $\Gamma(x, t, \bar{T})$ have the form

$$\begin{aligned} \dot{x}^t(s) &= Ax^t(s) + B_1 u_1^t(s, x^t) + \dots + B_n u_n^t(s, x^t), \\ x^t(t) &= x, \\ x^t &\in \mathbb{R}^l, \quad u^t = (u_1^t, \dots, u_n^t), \quad u_i^t = u_i^t(s, x^t) \in U_i \subset \text{comp} \mathbb{R}^k, \quad t \in [t_0, +\infty). \end{aligned} \quad (4.49)$$

The payoff function of player $i \in N$ in the game $\Gamma(x, t, \bar{T})$ is defined as

$$K_i^t(x^t, t, \bar{T}; u^t) = \int_t^{t+\bar{T}} \left((x^t(s))' Q_i x^t(s) + \sum_{j=1}^n (u_j^t(s, x^t))' R_{ij} u_j^t(s, x^t) \right) ds, \quad (4.50)$$

where $x^t(s)$ and $u^t(s, x)$ are the trajectory and strategies in the game $\Gamma(x, t, \bar{T})$. The matrices Q_i, R_{ij} satisfy the same conditions as in the original linear-quadratic differential game model.

4.2.4.3 Theorem on Nash Equilibrium with Continuous Updating for LQ Differential Games

We present an explicit form of the Nash equilibrium with continuous updating for a two-player differential game.

Theorem 4.2.3 *For any initial state, the two-player ($N = \{1, 2\}$) linear-quadratic differential game $\Gamma(x_0, t_0, \bar{T})$ has a linear feedback Nash equilibrium with continuous updating if and only if the following set of coupled Riccati differential equations has a set of symmetric solutions K_1, K_2 on the interval $[0, 1]$:*

$$\begin{aligned} \dot{K}_i(\tau) = & -(A\bar{T} - S_j K_j(\tau))' K_i(\tau) - K_i(\tau)(A\bar{T} - S_j K_j(\tau)) + \\ & + K_i(\tau) S_i K_i(\tau) - Q_i - K_j(\tau) S_{ji} K_j(\tau), \\ & K_i(1) = 0, \quad i \neq j \in N, \end{aligned} \quad (4.51)$$

where

$$S_i = \bar{T}^2 B_i R_{ii}^{-1} B_i', \quad S_{ij} = \bar{T}^2 B_i R_{ii}^{-1} R_{ji} R_{ii}^{-1} B_i', \quad i \neq j \in N. \quad (4.52)$$

In this case, there is a unique feedback Nash equilibrium with continuous updating, which has the form:

$$u_i^{NE}(t, x) = -R_{ii}^{-1} B_i' K_i(0) \bar{T} x, \quad i \in N. \quad (4.53)$$

Proof To prove the theorem, we introduce the following change of variables:

$$\begin{aligned} s &= t + \bar{T}\tau, \\ y(\tau) &= x(t + \bar{T}\tau), \\ v_i(\tau, y) &= u_i(t + \bar{T}\tau, x), \quad i \in N. \end{aligned} \quad (4.54)$$

Substituting (4.54) into the motion equations (4.5) and the payoff function (4.50),

we obtain

$$\dot{y}(\tau) = \bar{T}Ay(s) + \bar{T}B_1v_1(\tau, y) + \bar{T}B_2v_2(\tau, y) \quad (4.55)$$

and

$$K_i(y, \tau; v) = \int_0^1 y'(s)Q_iy(s) + \sum_{j=1}^2 (v_j(s, y))'R_{ij}v_j(s, y)ds, \quad i \in N. \quad (4.56)$$

As is known [224], a criterion for the existence of feedback Nash equilibrium is the existence of a symmetric solution for the system of differential equations (4.51). According to [224], the feedback Nash equilibrium strategies have the form

$$v_i^{NE}(\tau, y) = -R_{ii}^{-1}B_i'K_i(\tau)\bar{T}y. \quad (4.57)$$

From (4.54) we have

$$\tau = \frac{s-t}{\bar{T}}.$$

Returning to the original variables, we obtain the strategies

$$u_i^t(s, x) = -R_{ii}^{-1}B_i'K_i\left(\frac{s-t}{\bar{T}}\right)\bar{T}x.$$

These strategies are a feedback Nash equilibrium in the subgame $\Gamma(x, t, \bar{T})$ by construction.

The game (4.55), (4.56) and the solution (4.57) have the same form for all values t in the original game with continuous updating. The generalized Nash equilibrium with continuous updating has the form

$$\tilde{u}_i^{NE}(t, s, x) = -R_{ii}^{-1}B_i'K_i\left(\frac{s-t}{\bar{T}}\right)\bar{T}x. \quad (4.58)$$

We apply the procedure (4.11) to determine the Nash equilibrium with continuous updating using the generalized Nash equilibrium (5.16), $s = t$:

$$u_i^{NE}(t, x) = -R_{ii}^{-1}B_i'K_i(0)\bar{T}x, \quad t \in [t_0, +\infty), \quad i \in N. \quad (4.59)$$

This proves the theorem. □

4.2.4.4 LQ Game Model with Dynamic Updating

In this section, we first define the game model with dynamic updating and then demonstrate the convergence of the Nash equilibrium strategies and the corresponding trajectories for the case of dynamic and continuous updating.

In the papers [76; 77; 108; 110; 111; 112; 116], the method for constructing the differential game model with dynamic updating was described. It was assumed that the players have information about the game structure only on a truncated interval and make their decision accordingly. To model the behavior of players under dynamic updating, consider the case when the information is updated every $\Delta t > 0$, and the behavior of players on each interval $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, $j = 0, 1, 2, \dots$, is described by a truncated subgame.

Definition 4.2.3 *Let $j = 0, 1, 2, \dots$. A truncated subgame $\bar{\Gamma}_j(x_0^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ is the game defined on the interval $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ in the following way. On the interval $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, the payoff functions and motion equations in the truncated subgame and original game model $\Gamma(x_0, T - t_0)$ coincide:*

$$\begin{aligned} \dot{x}^j(s) &= Ax^j(s) + B_1 u_1^j(s, x^j) + \dots + B_n u_n^j(s, x^j), \\ x^j(t_0 + j\Delta t) &= x_0^j, \\ x^j \in \mathbb{R}^n, \quad u^j &= (u_1^j, \dots, u_n^j), \quad u_i^j = u_i^j(s, x^j) \in U_i \subset \text{comp} \mathbb{R}^k, \quad t \in [t_0, +\infty). \end{aligned} \quad (4.60)$$

$$\begin{aligned} K_i^j(x^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; u^j) &= \int_{t_0 + j\Delta t}^{t_0 + j\Delta t + \bar{T}} (x^j(s))' Q_i x^j(s) \\ &+ \sum_{k=1}^n (u^k(s, x^j))' R_{ik} u^k(s, x^j) ds, \quad i \in N. \end{aligned} \quad (4.61)$$

At any time instant $t = t_0 + j\Delta t$, information about the game structure is updated, and the players adapt accordingly. As a solution concept in the differential game model with dynamic updating, we will use the feedback Nash equilibrium with dynamic updating. Like in Section 4.2.4.3, we will need to define a special form of this equilibrium. Earlier in the Section 3 the notion of Nash equilibrium has not been directly introduced because mainly the cooperative case of the game with dynamic

updating has been considered. To determine the solution of the games with dynamic updating, the notion of resulting feedback Nash equilibrium is introduced:

Definition 4.2.4 *The resulting feedback Nash equilibrium*

$$\hat{u}^{NE}(t, x) = (\hat{u}_1^{NE}(t, x), \dots, \hat{u}_n^{NE}(t, x))$$

of players in the game model with dynamic updating has the form

$$\{\hat{u}^{NE}(t, x)\}_{t=t_0}^{\infty} = \begin{cases} u_0^{NE}(t, x), & t \in [t_0, t_0 + \Delta t], \\ \dots \\ u_j^{NE}(t, x), & t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t], \\ \dots \end{cases} \quad (4.62)$$

where $u_j^{NE}(t, x) = (u_1^{j,NE}(t, x), \dots, u_n^{j,NE}(t, x))$ is some fixed feedback Nash equilibrium in the truncated subgame $\bar{\Gamma}_j(x_0^{j,NE}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, $j = 0, 1, 2, \dots$, starting along the equilibrium trajectory of the previous truncated subgame: $x_0^{j,NE} = x^{j-1,NE}(t_0 + j\Delta t)$.

We denote by $\hat{x}^{NE}(t)$ the trajectory obtained using the original linear-quadratic motion equation (4.47) and the resulting feedback Nash equilibrium $\hat{u}^{NE}(t, x) = (\hat{u}_1^{NE}(t, x), \dots, \hat{u}_n^{NE}(t, x))$ and call it the resulting equilibrium trajectory.

4.2.4.5 Resulting Feedback Nash Equilibrium with Dynamic Updating

The following theorem presents necessary and sufficient conditions for Nash equilibrium with continuous updating in the linear-quadratic differential game with continuous updating.

Theorem 4.2.4 *For any initial state, the two-player linear-quadratic differential game $\bar{\Gamma}_j(x_0^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ has a linear feedback Nash equilibrium if and only if the following set of coupled Riccati differential equations has a set of symmetric solutions K_1, K_2 on*

the interval $[0, 1]$:

$$\begin{aligned} \dot{K}_i(\tau) = & -(A\bar{T} - S_j K_j(\tau))' K_i(\tau) - K_i(\tau)(A\bar{T} - S_j K_j(\tau)) + \\ & + K_i(\tau) S_i K_i(\tau) - Q_i - K_j(\tau) S_{ji} K_j(\tau), \\ & K_i(1) = 0, \quad i \neq j, \end{aligned} \quad (4.63)$$

where

$$S_i = \bar{T}^2 B_i R_{ii}^{-1} B_i', \quad S_{ij} = \bar{T}^2 B_i R_{ii}^{-1} R_{ji} R_{ii}^{-1} B_i', \quad i \neq j \in N. \quad (4.64)$$

The equilibrium strategies are unique and have the form

$$u_i^{j,NE}(t, x) = -R_{ii}^{-1} B_i' K_i \left(\frac{t - (t_0 + j\Delta t)}{\bar{T}} \right) \bar{T} x. \quad (4.65)$$

Proof To prove this theorem, for each truncated subgame we use a similar change of variables as in (4.54):

$$\tau = \frac{t - (t_0 + j\Delta t)}{\bar{T}}. \quad (4.66)$$

The remainder of the proof is straightforward. \square

According to (4.62), the Nash equilibrium with dynamic updating $\hat{u}_i^{NE}(t, x)$ can be constructed using the Nash equilibrium defined in each truncated subgame $u_i^{j,NE}(t, x)$. The corresponding trajectory $\hat{x}^{NE}(t)$ is constructed using $\hat{u}_i^{NE}(t, x)$ and (4.91).

4.2.4.6 Convergence of Resulting Nash Equilibrium Strategies and Trajectory

Theorem 4.2.5 For $x \in X$, where X is a bounded set, the resulting feedback Nash equilibrium strategies $\hat{u}_i^{NE}(t, x)$ in the game with dynamic updating uniformly converge to the feedback Nash equilibrium with continuous updating $\tilde{u}_i^{NE}(t, x)$ as $\Delta t \rightarrow 0$:

$$\hat{u}_i^{NE}(t, x) \underset{[t_0, +\infty)}{\rightrightarrows} \tilde{u}_i^{NE}(t, x), \quad i \in N. \quad (4.67)$$

Proof We introduce the notation $t_j \stackrel{\text{def}}{=} t_0 + j\Delta t$, and let $t \in [t_j, t_{j+1}]$ for some j . According to the definition of $\hat{u}_i^{NE}(t, x)$ (4.62), we will need to show that $\|\hat{u}_i^{NE}(t, x) -$

$u_i^{j,NE}(t, x) \parallel \rightarrow 0$ as $\Delta t \rightarrow 0$. Consider the expressions for \tilde{u}_i^{NE} and $u_i^{j,NE}$:

$$\begin{aligned}\tilde{u}_i^{NE}(t, x) &= -R_{ii}^{-1}B_i'K_i(0)\bar{T}x, \\ u_i^{j,NE}(t, x) &= -R_{ii}^{-1}B_i'K_i\left(\frac{t-t_j}{\bar{T}}\right)\bar{T}x.\end{aligned}$$

Form the inequality for the norm of the difference:

$$\|\tilde{u}_i^{NE}(t, x) - u_i^{j,NE}(t, x)\| \leq \|R_{ii}^{-1}B_i'\| \|x\| \left(\|\dot{K}(0)\| \frac{\Delta t}{\bar{T}} + o(\Delta t) \right). \quad (4.68)$$

As $\Delta t \rightarrow 0$, the right-hand side of (4.68) tends to zero, and consequently, the left-hand side of (4.68) does so. The proof is complete. \square

Theorem 4.2.6 *The equilibrium trajectory in the game with dynamic updating $\hat{x}^{NE}(t)$ pointwise converges to the equilibrium trajectory $\tilde{x}^{NE}(t)$ in the game with continuous updating $\tilde{x}^{NE}(t)$ as $\Delta t \rightarrow 0$:*

$$\hat{x}^{NE}(t) \xrightarrow{[t_0, +\infty)} \tilde{x}^{NE}(t). \quad (4.69)$$

Proof Let $t \in [t_j, t_{j+1}]$ for some j . According to the definition of $\hat{x}^{NE}(t)$, we will need to show that $\|\tilde{x}^{NE}(t) - x_j^{NE}(t)\| \rightarrow 0$ as $\Delta t \rightarrow 0$.

The trajectories $\tilde{x}^{NE}(t)$ and $x_j^{NE}(t)$ satisfy the following differential equations:

$$\begin{aligned}\dot{\tilde{x}}(t) &= (A - B_1R_{11}^{-1}B_1'K_1(0)\bar{T} - B_2R_{22}^{-1}B_2'K_2(0)\bar{T})\tilde{x}(t), \\ \dot{x}_j(t) &= \left(A - B_1R_{11}^{-1}B_1'K_1\left(\frac{t-t_j}{\bar{T}}\right)\bar{T} - B_2R_{22}^{-1}B_2'K_2\left(\frac{t-t_j}{\bar{T}}\right)\bar{T} \right) x_j(t).\end{aligned}$$

Note that

$$K_i(0)\tilde{x} - K_i\left(\frac{t-t_j}{\bar{T}}\right)x_j = K_i(0)(\tilde{x} - x_j) + \left(K_i(0) - K_i\left(\frac{t-t_j}{\bar{T}}\right) \right) x_j.$$

Let $y_j^{NE}(t) = \tilde{x}^{NE}(t) - x_j^{NE}(t)$, $\tilde{A} = A - B_1 R_{11}^{-1} B_1' K_1(0) \bar{T} - B_2 R_{22}^{-1} B_2' K_2(0) \bar{T}$, and

$$f_j(t) = -B_1 R_{11}^{-1} B_1' \left[K_1(0) - K_1 \left(\frac{t - t_j}{\bar{T}} \right) \right] \bar{T} x_j(t) - \\ -B_2 R_{22}^{-1} B_2' \left[K_2(0) - K_2 \left(\frac{t - t_j}{\bar{T}} \right) \right] \bar{T} x_j(t).$$

Then $y_j^{NE}(t)$ satisfies the differential equation

$$\dot{y}_j(t) = \tilde{A} y_j(t) + f_j(t).$$

Consider

$$y(t) = \begin{cases} y_0(t), & t \in [t_0, t_0 + \Delta t], \\ \dots \\ y_j(t), & t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t], \\ \dots \end{cases} \quad (4.70)$$

and

$$f(t) = \begin{cases} f_0(t), & t \in [t_0, t_0 + \Delta t], \\ \dots \\ f_j(t), & t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t]. \\ \dots \end{cases}$$

Then (4.70) satisfies the differential equation

$$\dot{y}(t) = \tilde{A} y(t) + f(t)$$

with the initial state $y(t_0) = 0$ since $\hat{x}^{NE}(t_0) = \tilde{x}^{NE}(t_0) = x_0$. By the Cauchy formula, we obtain for any $t \geq t_0$:

$$y(t) = \int_{t_0}^t e^{\tilde{A}(t-s)} f(s) ds.$$

Taking this expression into account, we have for fixed t :

$$\lim_{\Delta t \rightarrow 0} \|y(t)\| \leq \lim_{\Delta t \rightarrow 0} \|e^{\tilde{A}(t-t_0)}\| (t - t_0) \beta \left(\frac{\Delta t}{\bar{T}} + o(\Delta t) \right) = 0, \quad (4.71)$$

where

$$\beta = \left(\|B_1 R_{11}^{-1} B_1'\| \|\dot{K}_1(0)\| + \|B_2 R_{22}^{-1} B_2'\| \|\dot{K}_2(0)\| \right) \bar{T} M(t),$$

$$M(t) = \max_{\tau \in [t_0, t]} \|\hat{x}^{NE}(\tau)\|.$$

According to (4.71), $y(t) \xrightarrow{[t_0, +\infty)} 0$ as $\Delta t \rightarrow 0$. This proves the theorem. \square

4.2.4.7 Differential Game Model of Investment in Public Stock of Knowledge with Continuous Updating

4.2.4.7.1 Common Description

Consider a model with two individuals investing in a public stock of knowledge; also, see Dockner et al. [50]. Let $x(t)$ be the stock of knowledge at a time instant t , and $u_i(t)$ be the investment of player i in public knowledge at the time t . Assume that the stock of knowledge evolves according to the accumulation equation

$$\dot{x}(t) = -\beta x(t) + u_1(t, x) + u_2(t, x), \quad x(0) = x_0, \quad (4.72)$$

where β is the depreciation rate. Assume that each player obtains a quadratic utility from consuming the stock of knowledge, and the cost of investment increases quadratically with the investment effort. That is, the cost function of both players is given by

$$K_i(x_0, t_0, T; u) = \int_0^T \left(-q_i x^2(t) + r_i u_i^2(t, x) \right) dt, \quad i = 1, 2. \quad (4.73)$$

Consider the original game (4.72), (4.73) in the terms of LQ game theory [224]. To find a feedback Nash equilibrium, we need to solve the following set of coupled Riccati differential equations:

$$\begin{cases} \dot{k}_1(t) = -2\left(-\beta - \frac{1}{r_2} k_2(t)\right) k_1(t) + \frac{1}{r_1} k_1^2(t) + q_1, \\ \dot{k}_2(t) = -2\left(-\beta - \frac{1}{r_1} k_1(t)\right) k_2(t) + \frac{1}{r_2} k_2^2(t) + q_2, \\ k_1(T) = 0, \\ k_2(T) = 0. \end{cases} \quad (4.74)$$

As an example, consider the symmetric case in which $r_1 = r_2 = r$ and $q_1 = q_2 = q$. Let $k(t) = k_1(t) = k_2(t)$. We obtain the following differential equation:

$$\begin{cases} \dot{k}(t) = 2\beta k(t) + \frac{3k^2(t)}{r} + q, \\ k(T) = 0. \end{cases} \quad (4.75)$$

The solution of the Cauchy problem (4.75) is

$$k(t) = \frac{\beta r + v}{3} \left(\frac{2v}{(v - \beta r)e^{\frac{2v}{r}(t-T)} + v + \beta r} - 1 \right),$$

where $v = \sqrt{\beta^2 r^2 - 3qr}$. According to [224], the feedback Nash equilibrium in the original game model has the form

$$u_i^{NE}(t, x) = -\frac{k(t)x}{r}, \quad i = 1, 2. \quad (4.76)$$

Substituting the expression for $k(t)$ into (4.76), we obtain

$$u_i^{NE}(t, x) = \frac{\beta r + v}{3r} \left(1 - \frac{2v}{(v - \beta r)e^{\frac{2v}{r}(t-T)} + v + \beta r} \right) x(t).$$

4.2.4.7.2 Game Model with Continuous Updating

Now consider the case of continuous updating. Suppose that at each time instant $t \in [t_0, +\infty)$, two individuals use information about the motion equations and payoff functions on the interval $[t, t + \bar{T}]$. As the current time t evolves, the interval defining the information shifts as well. The motion equations for the game model with continuous updating have the form

$$\dot{x}^t(s) = -\beta x^t(s) + u_1^t(s, x) + u_2^t(s, x), \quad x^t(t) = x, \quad t \in [t_0, +\infty). \quad (4.77)$$

The payoff function of player $i \in N$ in the game model with continuous updating is defined as

$$K_i^t(x^t, t, \bar{T}; u^t) = \int_t^{t+\bar{T}} \left(-(x^t(s))^2 q_i + (u_i^t(s, x))^2 r_i \right) ds, \quad i = 1, 2. \quad (4.78)$$

According to Theorem 4.2.4 defining the feedback Nash equilibrium with continuous updating, in the first step, we need to solve the following differential equation:

$$\begin{cases} \dot{k}(\tau) = 2\beta\bar{T}k(\tau) + \frac{3\bar{T}k^2(\tau)}{r} + \bar{T}q, \\ k(1) = 0. \end{cases} \quad (4.79)$$

The solution of (4.79) is

$$k(\tau) = \frac{\beta r + v}{3} \left(\frac{2v}{(v - \beta r)e^{\frac{2v\bar{T}}{r}(\tau-1)} + v + \beta r} - 1 \right), \quad (4.80)$$

where $v = \sqrt{\beta^2 r^2 - 3qr}$. According to (4.97), the feedback Nash equilibrium with continuous updating has the form

$$\tilde{u}_i^{NE}(t, x) = -\frac{k(0)x\bar{T}}{r}. \quad (4.81)$$

Substituting (4.80) into (4.81), we obtain

$$\tilde{u}_i^{NE}(t, x) = \frac{\beta r + v}{3r} \left(1 - \frac{2v}{(v - \beta r)e^{-\frac{2v\bar{T}}{r}} + v + \beta r} \right) \bar{T}x. \quad (4.82)$$

Substituting (4.82) into (4.72), we obtain $\tilde{x}^{NE}(t)$ as the solution of the equation

$$\dot{\tilde{x}}^{NE}(t) = -\beta\tilde{x}^{NE}(t) + \tilde{u}_1^{NE}(t, x) + \tilde{u}_2^{NE}(t, x), \quad \tilde{x}^{NE}(0) = x_0. \quad (4.83)$$

4.2.4.7.3 Game Model with Dynamic Updating

We perform similar calculations for the resulting Nash equilibrium in the game with dynamic updating based on the calculations for the original game and the approach described in Section 4.2.5.4.1. As a result, we arrive in

$$\tilde{u}_i^{NE}(t, x) = -\frac{k\left(\frac{t-t_i}{\bar{T}}\right)x\bar{T}}{r}, \quad t \in [t_i, t_{i+1}]. \quad (4.84)$$

Substituting (4.80) into (4.84), we obtain

$$\hat{u}_i^{NE}(t, x) = \frac{\beta r + v}{3r} \left(1 - \frac{2v}{(v - \beta r)e^{\frac{2v(t-t_i-\bar{T})}{r}} + v + \beta r} \right) \bar{T}x, \quad t \in [t_i, t_{i+1}]. \quad (4.85)$$

Substituting (4.85) into (4.72), we obtain $\hat{x}^{NE}(t)$ as the solution of the equation

$$\dot{\hat{x}}^{NE}(t) = -\beta \hat{x}^{NE}(t) + \hat{u}_1^{NE}(t, x) + \hat{u}_2^{NE}(t, x), \quad \hat{x}^{NE}(0) = x_0. \quad (4.86)$$

4.2.4.7.4 Game Model on Infinite Interval

Consider the classical approach to Nash equilibrium in the game on an infinite interval $[0, +\infty)$. The motion equations have the form

$$\dot{x}(t) = -\beta x(t) + u_1(t, x) + u_2(t, x), \quad x(0) = x_0. \quad (4.87)$$

The payoff function of player $i \in N$ is defined as

$$K_i(x_0; u) = \lim_{T \rightarrow \infty} \int_0^T (-q_i x^2(t) + r_i u_i^2(t, x)) dt, \quad i = 1, 2. \quad (4.88)$$

According to [224], the feedback Nash equilibrium strategies in the symmetric case ($r_1 = r_2 = r$ and $q_1 = q_2 = q$) have the form

$$u^{NE}(t, x) = -\frac{kx}{r}, \quad (4.89)$$

where k is solution of

$$\frac{3k^2}{r} + 2\beta k + q = 0.$$

Substituting (4.89) into (4.87), we obtain $x^{NE}(t)$ as the solution of the equation

$$\dot{x}^{NE}(t) = \left(-\beta - \frac{2k}{r} \right) x^{NE}(t), \quad x^{NE}(0) = x_0. \quad (4.90)$$

4.2.4.7.5 Numerical Simulation

Consider the numerical simulation results for the game model presented above on the interval $[0, 8]$, i.e., $t_0 = 0$ and $T = 8$. At the initial instant $t_0 = 0$, the stock of knowledge is 100, i.e., $x_0 = 100$. The other parameters of the model are $\beta = 0.9$, $r = 6$, and $q = 1$. Suppose that under dynamic updating (blue solid and dotted lines, Fig.4.5-4.6), the intervals between the updating instants are $\Delta t = 2$; therefore, $l = 4$. In Figure 4.5, a comparison of the resulting Nash equilibrium in the game with dynamic updating (blue line) and the Nash equilibrium with continuous updating (red lines) is presented. In Figure 4.6, similar results are presented for the strategies.

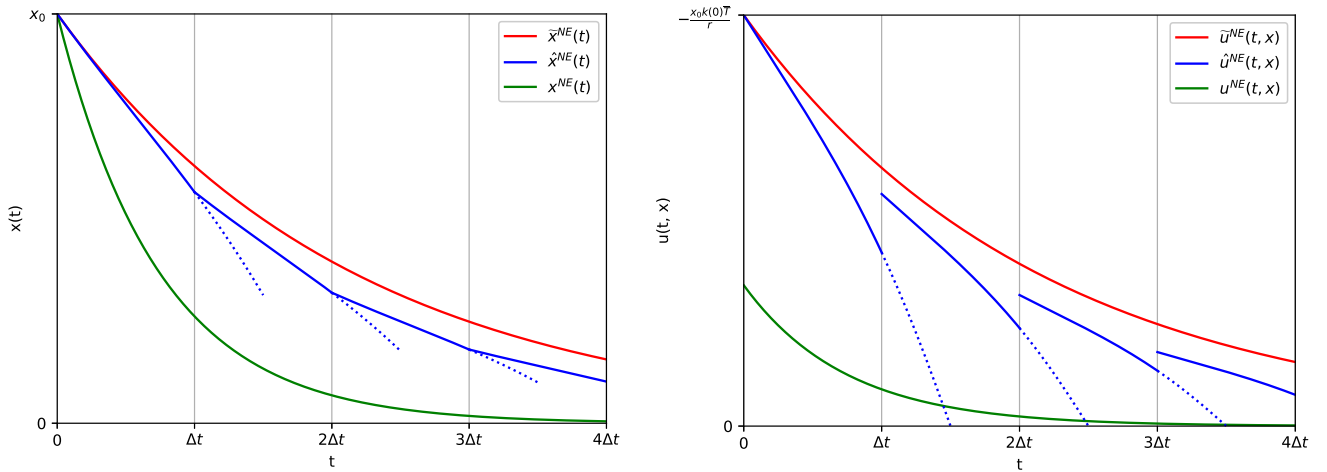


Figure 4.5: $\tilde{x}^{NE}(t)$ (4.83) — red upper line, $\hat{x}^{NE}(t)$ (4.86) — blue broken line, and $x^{NE}(t)$ (4.90) — green lower line. Figure 4.6: $\tilde{u}^{NE}(t)$ (4.82) — red upper line, $\hat{u}^{NE}(t)$ (4.85) — blue broken line, and $u^{NE}(t)$ (4.89) — green lower line.

To demonstrate the results of Theorems 4.2.5 and 4.2.6 on the convergence of the resulting equilibrium strategies and corresponding trajectory to the equilibrium strategies and trajectory with continuous updating, consider the simulation results for a case of frequent updating, namely, $l = 20$. Figures 4.7-4.8 represent the same solutions as in Figures 4.5 - 4.6, but for the case $\Delta t = 0.4$. Therefore, the convergence results are confirmed by the numerical experiments presented below.

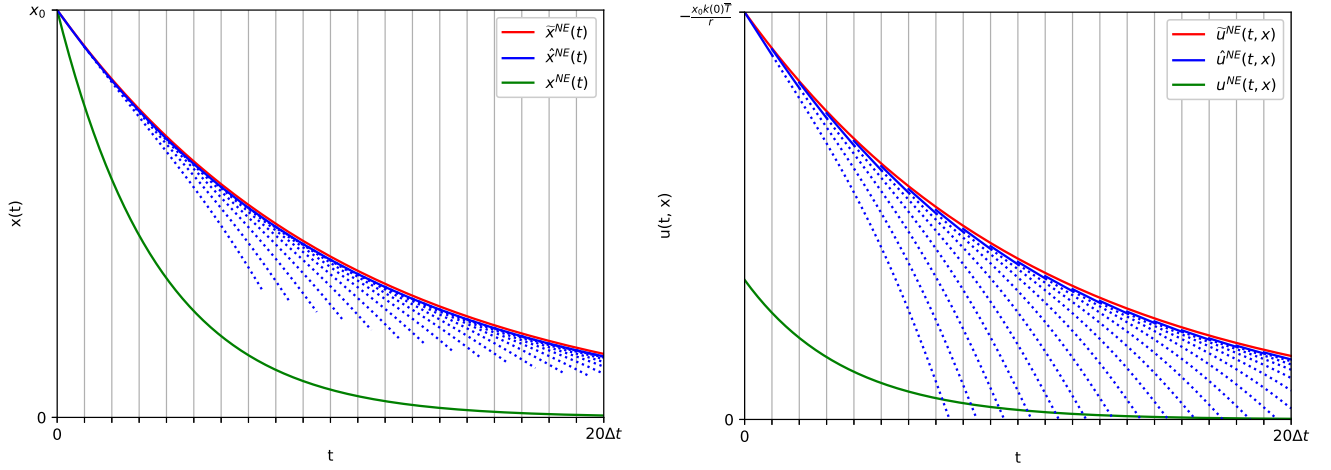


Figure 4.7: $\tilde{x}^{NE}(t)$ (4.83) — red upper line, $\hat{x}^{NE}(t)$ (4.86) — blue broken line, and $x^{NE}(t)$ (4.90) — green lower line. Figure 4.8: $\tilde{u}^{NE}(t, x)$ (4.82) — red upper line, $\hat{u}^{NE}(t, x)$ (4.85) — blue broken line, and $u^{NE}(t, x)$ (4.89) — green lower line.

4.2.5 Non-autonomous Linear-Quadratic Case with Continuous Updating

Here, a special type of the non-autonomous linear-quadratic differential game model with continuous updating is considered. It is an important class for the continuous updating setting because it models the updating of the motion equations and payoff functions depending on the current time t .

4.2.5.1 Original Linear-Quadratic Non-autonomous Game Model

Consider an n -player ($|N| = n$) linear quadratic non-autonomous differential game $\Gamma(x_0, t_0, T)$ defined on an interval $[t_0, T]$.

The motion equations have the form

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B_1(t)u_1(t, x) + \dots + B_n(t)u_n(t, x), \\ x(t_0) &= x_0, \\ x &\in \mathbb{R}^l, u = (u_1, \dots, u_n), u_i = u_i(t, x) \in U_i \subset \text{comp} \mathbb{R}^k, t \in [t_0, T]. \end{aligned} \tag{4.91}$$

The payoff function of player $i \in N$ is defined as

$$K_i(x_0, t_0, T; u) = \int_{t_0}^T \left(x'(t)Q_i(t)x(t) + \sum_{j=1}^n u'_j(t, x)R_{ij}(t)u_j(t, x) \right) dt, \quad i \in N, \quad (4.92)$$

where $Q_i(t)$ and $R_{ij}(t)$ are assumed to be symmetric for $t \in [t_0, T]$, $R_{ii}(t)$ is positive definite for $t \in [t_0, T]$, and $(\cdot)'$ denotes transpose.

4.2.5.2 Linear-Quadratic Non-autonomous Game Model with Continuous Updating

For the class of games with continuous updating, the difference between the non-autonomous and autonomous models is that the right-hand side of the motion equations and the integrand in the payoff function explicitly depend on the current time t . For the linear-quadratic model, the dependence on the current time t in (4.91), (4.92) is introduced through the following matrices:

$$A(t), B_i(t), Q_i(t), R_{ij}(t), \quad i, j = 1, \dots, n. \quad (4.93)$$

There is a special meaning of this dependence. At the beginning of the entire game at the time instant t_0 , the players have only information about the motion equations and payoff functions for the interval $[t_0, t_0 + \bar{T}]$. Thus, they know the matrices (4.93) for the interval $[t_0, t_0 + \bar{T}]$ because the form of (4.93) can be different for the interval $[t_0 + \bar{T}, +\infty]$. In the autonomous case, the matrices (4.93) are constant. Hence, they are known for the entire interval on which the game is defined.

Consider an n -player differential game $\Gamma(x, t, t + \bar{T})$, $t \in [t_0, +\infty)$ defined on the interval $[t, t + \bar{T}]$, where $0 < \bar{T} < +\infty$.

The motion equations of $\Gamma(x, t, t + \bar{T})$ have the form

$$\begin{aligned} \dot{x}^t(s) &= A(s)x^t(s) + B_1(s)u_1^t(s, x^t) + \dots + B_n(s)u_n^t(s, x^t), \\ x^t(t) &= x, \\ x^t \in \mathbb{R}^l, \quad u^t &= (u_1^t, \dots, u_n^t), \quad u_i^t = u_i^t(s, x^t) \in U_i \subset \text{comp} \mathbb{R}^k, \quad t \in [t_0, +\infty). \end{aligned} \quad (4.94)$$

The payoff function of player $i \in N$ in the game $\Gamma(x, t, t + \bar{T})$ is defined as

$$K_i^t(x, t, \bar{T}; u^t) = \int_t^{t+\bar{T}} \left((x^t(s))' Q_i(s) x^t(s) + \sum_{j=1}^n (u_j^t(s, x^t))' R_{ij}(s) u_j^t(s, x^t) \right) ds, \quad (4.95)$$

where $x^t(s)$ and $u^t(s, x^t)$ are the trajectory and strategies in the game $\Gamma(x, t, t + \bar{T})$.

4.2.5.3 Theorems on Nash Equilibrium with Continuous Updating for LQ Differential Games

One of the main results of this section is the establishment of sufficient conditions for the existence of a feedback-based Nash equilibrium with continuous updating in the non-autonomous case. Related results can be found in the papers [79] and [117], where the sufficient conditions for a Nash equilibrium with continuous updating in the autonomous case were presented.

Theorem 4.2.7 *For an N -person linear-quadratic differential game $\Gamma(x_0, t_0, t_0 + \bar{T})$ with continuous updating with $Q_i(\cdot) \geq 0$ and $R_{ij}(\cdot) \geq 0$ ($i, j \in N, i \neq j$), let the system of N coupled matrix Riccati differential equations*

$$\begin{aligned} & \frac{dZ_i^t(\tau)}{d\tau} + Z_i^t(\tau) F^t(\tau) + (F^t(\tau))' Z_i^t(\tau) + Q_i(t + \bar{T}\tau) + \\ & + \bar{T}^2 \sum_{j \in N} Z_j^t(\tau) B_j(t + \bar{T}\tau) R_{jj}^{-1}(t + \bar{T}\tau) R_{ij}(t + \bar{T}\tau) \times \\ & \times R_{jj}^{-1}(t + \bar{T}\tau) B_j'(t + \bar{T}\tau) Z_j^t(\tau) = 0, \quad \tau \in [0, 1], \\ & Z_i^t(1) = 0, \quad i \in N, \end{aligned} \quad (4.96)$$

where

$$F^t(\tau) = \bar{T}A(t + \bar{T}\tau) - \bar{T}^2 \sum_{i \in N} B_i(t + \bar{T}\tau) R_{ii}^{-1}(t + \bar{T}\tau) B_i'(t + \bar{T}\tau) Z_i^t(\tau),$$

have a solution $Z_i^t(\cdot) \geq 0$, $i \in N$ for $t \geq t_0$. Then the linear-quadratic differential game with continuous updating has a linear feedback-based Nash equilibrium with

continuous updating

$$u_{fb,i}^{NE}(t, x) = -R_{ii}^{-1}(t)B'_i(t)Z_i^t(0)\bar{T}x, \quad i \in N. \quad (4.97)$$

Proof To prove the theorem, we introduce the change of variables

$$\begin{aligned} s &= t + \bar{T}\tau, \\ y^t(\tau) &= x_{fb}^t(t + \bar{T}\tau), \\ v_i^t(\tau, y) &= u_{fb,i}(t + \bar{T}\tau, x), \quad i \in N. \end{aligned} \quad (4.98)$$

Substituting (4.98) into the motion equations (4.94) and payoff function (4.95), we obtain

$$\dot{y}^t(\tau) = \bar{T}A(t + \bar{T}\tau)y^t(\tau) + \sum_{i=1}^N \bar{T}B_i(t + \bar{T}\tau)v_i^t(\tau, y) \quad (4.99)$$

and

$$\begin{aligned} K_i^t(y^t, \tau; v^t) &= \int_0^1 (y^t(\tau))' Q_i(t + \bar{T}\tau)y^t(\tau) \\ &+ \sum_{j=1}^N (v_j^t(\tau, y))' R_{ij}(t + \bar{T}\tau)v_j^t(\tau, y)d\tau, \quad i \in N. \end{aligned} \quad (4.100)$$

Corollary 6.5 from [224] (sufficient conditions for the existence of Nash equilibrium in the affine-quadratic game) and the existence of a solution $Z_i^t(\tau) \geq 0$ for the system of differential equations (4.96) lead to the feedback-based Nash equilibrium strategies in the subgame $\Gamma(x, t, t + \bar{T})$ that have the form

$$v_i^{t,NE}(\tau, y) = -R_{ii}^{-1}(t + \bar{T}\tau)B'_i(t + \bar{T}\tau)Z_i^t(\tau)\bar{T}y. \quad (4.101)$$

From (4.98) we have

$$\tau = \frac{s - t}{\bar{T}}.$$

Returning to the original variables, we obtain the strategies

$$u_{fb,i}^t(s, x) = -R_{ii}^{-1}(s)B'_i(s)Z_i^t\left(\frac{s - t}{\bar{T}}\right)\bar{T}x. \quad (4.102)$$

These strategies are Nash equilibrium in feedback-based strategies in the subgame $\Gamma(x, t, t + \bar{T})$ by construction.

Problem (4.99), (4.100) and the solution (4.101) have the same form for all values t in the original game with continuous updating. Then a generalized feedback-based Nash equilibrium with continuous updating has the form

$$\tilde{u}_{fb,i}^{NE}(t, s, x) = -R_{ii}^{-1}(s)B'_i(s)Z_i\left(\frac{s-t}{\bar{T}}\right)\bar{T}x. \quad (4.103)$$

We apply the procedure (4.11) to determine the Nash equilibrium with continuous updating using the generalized Nash equilibrium (4.103), $s = t$:

$$u_{fb,i}^{NE}(t, x) = -R_{ii}^{-1}(t)B'_i(t)Z_i(0)\bar{T}x, \quad t \in [t_0, +\infty), \quad i \in N. \quad (4.104)$$

This proves the theorem. □

The form of open-loop-based Nash equilibrium with continuous updating was presented in [117]. We will use these results later to study convergence in Section 4.2.5.4.

Theorem 4.2.8 *For an N -person linear-quadratic differential game with $Q_i(\cdot) \geq 0$ and $R_{ij}(\cdot) \geq 0$ ($i, j \in N, i \neq j$), let there exist a solution set $\{M_i^t, i \in N, t \geq t_0\}$ for the coupled matrix Riccati differential equations*

$$\begin{aligned} & \frac{dM_i^t(\tau)}{d\tau} + \bar{T}M_i^t(\tau)A(t + \bar{T}\tau) + \bar{T}A'(t + \bar{T}\tau)M_i^t(\tau) + Q_i(t + \bar{T}\tau) - \\ & - \bar{T}^2 M_i^t(\tau) \sum_{j \in N} B_j(t + \bar{T}\tau) (R_{jj}(t + \bar{T}\tau))^{-1} B'_j(t + \bar{T}\tau) M_j^t(\tau) = 0, \end{aligned} \quad (4.105)$$

$$M_i^t(1) = 0, \quad \tau \in [0, 1], \quad i \in N.$$

Then the differential game with continuous updating admits an open-loop-based Nash equilibrium with continuous updating given by

$$u_{ol,i}^{NE}(t, x) = -R_{ii}^{-1}(t)B'_i(t)M_i^t(0)\bar{T}x, \quad i \in N.$$

Remark 2 *Note that the open-loop-based solution with continuous updating has a feedback form; i. e., the open-loop-based Nash equilibrium with continuous updating explicitly depends on the current state. This happens because of the way the solution*

is constructed: at each current time t , the players reconsider their decisions under the continuously updated information.

Example 4.2.1 Consider the following autonomous differential game model with two players to compare the open-loop-based and feedback-based Nash equilibrium with continuous updating. Let the motion equations in the subgame have the form

$$\begin{aligned} \dot{x}^t(s) &= -\beta x^t(s) + u_1^t(s, x) + u_2^t(s, x), \\ x^t(t) &= x, \quad x^t \in \mathbb{R}^1 \end{aligned} \quad (4.106)$$

and the payoff function of player $i \in \{1, 2\}$ be defined as

$$K_i(x, t, \bar{T}, u^t) = \int_t^{t+\bar{T}} \left(q(x^t(s))^2 + r_1(u_i^t(s, x^t))^2 + r_2(u_j^t(s, x^t))^2 \right) ds, \quad (4.107)$$

where $i, j \in \{1, 2\}$, $i \neq j$, $q, r_1, r_2 > 0$.

Due to the players' symmetry, we need to consider only one equation. Hence, for the feedback-based Nash equilibrium with continuous updating, the Riccati differential equation has the form

$$\begin{aligned} \dot{z}(\tau) &= 2\bar{T}\beta z(\tau) + 2\bar{T}^2 z^2(\tau) \left(\frac{2}{r_1} - \frac{r_2}{r_1^2} \right) - q, \quad \tau \in [0, 1], \\ z(1) &= 0. \end{aligned}$$

For the open-loop-based Nash equilibrium with continuous updating, we obtain

$$\begin{aligned} \dot{m}(\tau) &= 2\bar{T}\beta m(\tau) + 2\bar{T}^2 \frac{m^2(\tau)}{r_1} - q, \quad \tau \in [0, 1], \\ m(1) &= 0. \end{aligned}$$

Note an important feature: since we consider the case of two players, the conditions of Theorems 4.2.7 and 4.2.8 become necessary as well. Therefore, the feedback-based and open-loop-based Nash equilibria with continuous updating are unique for this model. We have simulated the Nash equilibrium obtained using the Riccati equations above for the parameter values $\beta = 0.01$, $r_1 = 1$, $r_2 = 0.1$, $q = 5$, $\bar{T} = 1$, $t \in [0, 2]$, and $x(0) = 100$ and plotted them (Fig. 4.9) and the corresponding

equilibrium trajectories (Fig. 4.10) for the open-loop-based and feedback-based cases. The solutions are obviously different, even though both of them are of the feedback-based form (depend on the state x).

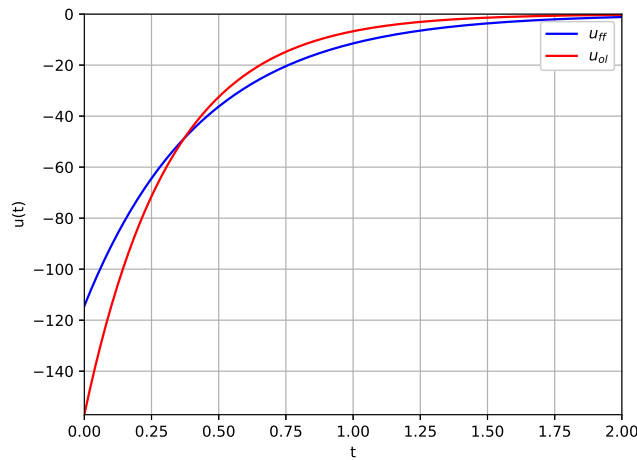


Figure 4.9: $u_{fb}^{NE}(t)$ — feedback-based Nash equilibrium with continuous updating for (4.106), (4.107), $u_{ol}^{NE}(t)$ — open-loop-based Nash equilibrium with continuous updating for (4.106), (4.107).

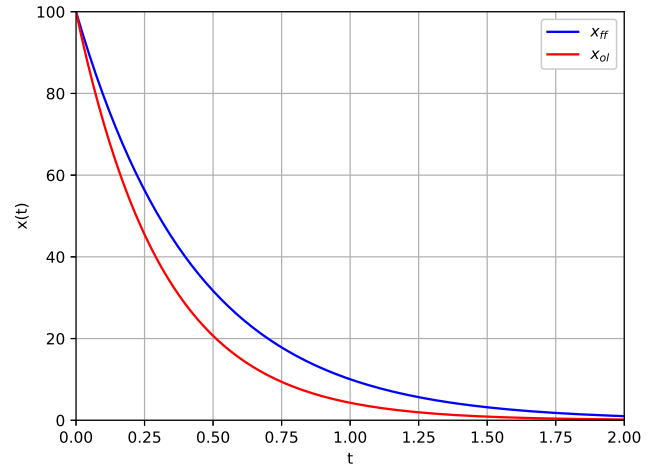


Figure 4.10: $x_{fb}^{NE}(t)$ — trajectory for feedback-based NE with continuous updating for (4.106), (4.107), $x_{ol}^{NE}(t)$ — trajectory for open-loop-based NE with continuous updating for (4.106), (4.107).

4.2.5.3.1 Existence, Uniqueness and Continuity

An important assumption of the previous section is the existence of Nash equilibrium with continuous updating. According to the definition, for the construction of Nash equilibrium with continuous updating $u^{NE}(t, x)$, it is necessary to have a Nash equilibrium $\tilde{u}^{NE}(t, x; s, x_t)$ ($\tilde{u}^{NE}(t, x; s)$) in each subgame $\Gamma(x, t, t + \bar{T})$, $t \geq t_0$.

The uniqueness of Nash equilibrium in differential games is not a trivial problem. As is known ([225]), even in the case of an autonomous linear-quadratic game, the open-loop Nash equilibrium can be non-unique even when sufficient conditions similar to the theorems presented above are satisfied.

Suppose that in the games with continuous updating, in each subgame $\Gamma(x, t, t + \bar{T})$, $t \geq t_0$, there are at least two Nash equilibria on the interval $[t, t + \bar{T}]$. Then the set of Nash equilibria with continuous updating will be uncountable because the players can switch from one Nash equilibrium to the other at every current time

instant time t .

Example 4.2.2 Consider the following game model with continuous updating defined by the motion equations with the open-loop strategies of the players:

$$\dot{x}^t(s) = \begin{bmatrix} u_1(s) & 0 \\ 0 & u_2(s) \end{bmatrix},$$

$$x^t(t) = x, \quad x^t \in \mathbb{R}^2.$$

The payoff function of player $i \in \{1, 2\}$ is defined as

$$K_i^t \left(x, t, \frac{\pi}{2}, u^t \right) = \frac{1}{2} \int_t^{t+\frac{\pi}{2}} \left((x^t(s))' Q_i x^t(s) + (u_i^t(s))^2 \right) ds,$$

where

$$Q_1 = \begin{bmatrix} \frac{1}{2} & -1 \\ -1 & 4 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 15 & -2 \\ -2 & \frac{1}{3} \end{bmatrix}.$$

If $x = (0, 0)'$, then (see Example 4.1 in [225]) this differential subgame has the trivial solution $u_1^t = u_2^t \equiv 0$ and the set of nontrivial solutions $u_1^t = \frac{2}{3}u_2^t \equiv \alpha$, $\alpha \in \mathbb{R}$. Thus, the set of open-loop-based Nash equilibria with continuous updating is formed by the piecewise continuous functions

$$u_{ol,i}(t) = \begin{cases} 0, & t \in [t_0, t_1), \\ u_i^*(t), & t \in [t_1, \infty), \end{cases}$$

where t_1 is a switching point from the trivial to nontrivial solution, and $u_i^*(t)$ is some nontrivial strategy of player i after this switching point. As we can see, the set of switching points is uncountable. Therefore, the set of open-loop-based Nash equilibria with continuous updating is uncountable as well.

Moreover, an open-loop-based Nash equilibrium with continuous updating may have continuity or may have a point of discontinuity (jump) at some time instant t_1 , where the players switch their strategies from the trivial to nontrivial ones. Thus, if there are several Nash equilibria in the subgames, then in the game with continuous updating, the Nash equilibrium with continuous updating may be discontinuous.

Lemma 4.2.1 *Let the differential game be autonomous, and the Nash equilibrium in each subgame $\Gamma(x, t, t + \bar{T})$ be unique and continuous. Then the Nash equilibrium with continuous updating is unique and continuous.*

Proof The uniqueness of a Nash equilibrium with continuous updating follows from its definition and the uniqueness of the Nash equilibrium in each subgame $\Gamma(x, t, t + \bar{T})$.

In the autonomous case, the players consider the same subgame $\Gamma(x, t, t + \bar{T})$ at every time instant t . Without loss of generality, let $u^*(s, x)$ be the Nash equilibrium in this common subgame $\Gamma(x, t_0, t_0 + \bar{T})$. Then, according to Definition 4.2.2, the Nash equilibrium with continuous updating has the form $u^{NE}(t, x) = u^*(t_0, x)$. Thus, the Nash equilibrium with continuous updating does not explicitly depend on time, and its continuity follows from the continuity of the Nash equilibrium in the subgame $\Gamma(x, t, t + \bar{T})$. \square

4.2.5.4 Convergence Results for Strategies and Trajectories

Clearly, it is impossible to continuously and instantly process new information about the game in real-life problems. At least, it takes time to obtain this new information, which by this time will have lost its relevance due to the continuous updating of the game.

In this section, the convergence results of Nash equilibrium strategies and trajectories with continuously updating are presented. For this purpose, the concept of dynamic updating is presented, which models the noncooperative behavior of players when information is updated at discrete time instants. The convergence of strategies and trajectories with dynamic updating and the strategies with continuous updating is proved both for the open-loop-based and feedback-based cases.

4.2.5.4.1 LQ Game Model with Dynamic Updating

In the papers [76; 77; 108; 110; 111; 112; 116], the method for constructing a differential game model with dynamic updating was described. It was assumed

that the players have information about the game structure only over a truncated interval and make their decisions accordingly. To model the behavior of players under dynamic updating, consider the case when information is updated every $\Delta t > 0$, and the behavior of players on each interval $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, $j = 0, 1, 2, \dots$, is described by a truncated subgame.

In this section, we will consider both feedback and open-loop based strategies for dynamic updating case. Therefore, we will use the notations $v_{fb}^{NE}(t, x)$ and $v_{ol}^{NE}(t, x)$, respectively.

Definition 4.2.5 *Let $j = 0, 1, 2, \dots$. A truncated subgame $\bar{\Gamma}_j(x_0^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ is a game defined for the interval $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ in the following way. On the interval $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, the payoff functions and motion equations in the truncated subgame and original game model $\Gamma(x_0, t_0, T)$ coincide:*

$$\begin{aligned} \dot{x}^j(s) &= A(s)x^j(s) + B_1(s)v_1^j(s, x^j) + \dots + B_n(s)v_n^j(s, x^j), \\ x^j(t_0 + j\Delta t) &= x_0^j, \\ x^j \in \mathbb{R}^n, \quad v^j &= (v_1^j, \dots, v_n^j), \quad v_i^j = v_i^j(s, x^j) \in U_i \subset \text{comp}\mathbb{R}^k, \quad t \in [t_0, +\infty). \end{aligned} \tag{4.108}$$

$$\begin{aligned} K_i^j(x^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; v^j) &= \int_{t_0 + j\Delta t}^{t_0 + j\Delta t + \bar{T}} (x^j(s))' Q_i(s) x^j(s) \\ &+ \sum_{k=1}^n (v^k(s, x^j))' R_{ik}(s) v^k(s, x^j) ds, \quad i \in N. \end{aligned} \tag{4.109}$$

At any instant $t = t_0 + j\Delta t$, information about the game structure is updated, and the players adapt accordingly. This class of game models is called differential games with dynamic updating.

Like in Section 4.2.5.3, we will need to define a special form of the Nash equilibrium. According to the approach described above, at any time instant $t \in [t_0, +\infty)$, the players have or use truncated information about the game structure. Therefore, the classical approaches for determining optimal strategies cannot be directly applied. To determine the solution for games with dynamic

updating, the notion of *feedback-based (open-loop-based) Nash equilibrium with dynamic updating* is introduced:

Definition 4.2.6 *A feedback-based (open-loop-based) Nash equilibrium with dynamic updating*

$$\begin{aligned} v_{fb}^{NE}(t, x) &= (v_{fb,1}^{NE}(t, x), \dots, v_{fb,n}^{NE}(t, x)) \\ (v_{ol}^{NE}(t, x) &= (v_{ol,1}^{NE}(t, x), \dots, v_{ol,n}^{NE}(t, x))) \end{aligned}$$

of players in the game model with dynamic updating has the form

$$\begin{aligned} \{v_{fb}^{NE}(t, x)\}_{t=t_0}^{\infty} &= \tilde{v}_{fb,j}^{NE}(t, x), \quad t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t], \quad j = 0, 1, 2, \dots \\ (\{v_{ol}^{NE}(t, x)\}_{t=t_0}^{\infty} &= \tilde{v}_{ol,j}^{NE}(t, x), \quad t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t], \quad j = 0, 1, 2, \dots) \end{aligned} \quad (4.110)$$

where $\tilde{v}_{fb,j}^{NE}(t, x) = (\tilde{v}_{fb,1}^{j,NE}(t, x), \dots, \tilde{v}_{fb,n}^{j,NE}(t, x))$ ($\tilde{v}_{ol,j}^{NE}(t, x) = (\tilde{v}_{ol,1}^{j,NE}(t, x), \dots, \tilde{v}_{ol,n}^{j,NE}(t, x))$) is some fixed feedback-based (open-loop) Nash equilibrium in the truncated subgame $\bar{\Gamma}_j(x_0^{j,NE}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, $j = 0, 1, 2, \dots$, starting along the equilibrium trajectory of the previous truncated subgame: $x_0^{j,NE} = x^{j-1,NE}(t_0 + j\Delta t)$.

Note that the Nash equilibrium with dynamic updating $v^{NE}(t, x)$ is not the Nash equilibrium in the classical sense, but can be used as a solution concept related to Nash equilibrium for the class of games with dynamic updating. The corresponding trajectory $\hat{x}_{fb}^{NE}(t)$ ($\hat{x}_{ol}^{NE}(t)$) is obtained using the motion equation (4.91) and the feedback-based (open-loop-based) Nash equilibrium with dynamic updating $v^{NE}(t, x) = (v_1^{NE}(t, x), \dots, v_n^{NE}(t, x))$.

4.2.5.4.2 Resulting Nash Equilibrium with Dynamic Updating

Another important set of results concerns sufficient conditions for the existence of a feedback-based and open-loop-based Nash equilibrium with dynamic updating in the non-autonomous case.

Theorem 4.2.9 *For an N -person linear-quadratic differential game $\bar{\Gamma}_j(x_0^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ with dynamic updating with $Q_i(\cdot) \geq 0$ and $R_{ik}(\cdot) \geq 0$*

($i, k \in N, i \neq k$), let the system of N coupled matrix Riccati differential equations

$$\begin{aligned} & \frac{dZ_i^{t_j}(\tau)}{d\tau} + Z_i^{t_j}(\tau)F^{t_j}(\tau) + (F^{t_j}(\tau))' Z_i^{t_j}(\tau) + Q_i(t_j + \bar{T}\tau) + \\ & + \bar{T}^2 \sum_{k \in N} Z_k^{t_j}(\tau) B_k(t_j + \bar{T}\tau) R_{kk}^{-1}(t_j + \bar{T}\tau) R_{ik}(t_j + \bar{T}\tau) \times \\ & \quad \times R_{kk}^{-1}(t_j + \bar{T}\tau) B_k'(t_j + \bar{T}\tau) Z_k^{t_j}(\tau) = 0, \\ & \quad Z_i^{t_j}(1) = 0, \quad \tau \in [0, 1], \quad i \in N, \end{aligned} \quad (4.111)$$

where

$$F^{t_j}(\tau) = \bar{T}A(t_j + \bar{T}\tau) - \bar{T}^2 \sum_{i \in N} B_i(t_j + \bar{T}\tau) R_{ii}^{-1}(t_j + \bar{T}\tau) B_i'(t_j + \bar{T}\tau) Z_i^{t_j}(\tau),$$

have a solution $Z_i^{t_j}(\cdot) \geq 0, i \in N$, for $t \geq t_0 + j\Delta t$. Then the linear-quadratic differential game with dynamic updating has a feedback-based Nash equilibrium with dynamic updating given by, $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$:

$$v_{fb,i}^{NE}(t, x) = -R_{ii}^{-1}(t) B_i'(t) Z_i^{t_j} \left(\frac{t - (t_0 + j\Delta t)}{\bar{T}} \right) \bar{T}x, \quad j = 0, 1, 2, \dots, \quad i \in N. \quad (4.112)$$

Proof The proof of this theorem is similar to the proof of Theorem 4.2.7. The essential difference is the time interval on which the solutions of the entire game and the subgame match. On one hand, the Nash equilibrium with continuous updating coincides with the Nash equilibrium in the subgame $\Gamma(x, t, t + \bar{T})$ at only one instant t . On the other hand, the Nash equilibrium with dynamic updating coincides with the Nash equilibrium in the subgame $\bar{\Gamma}_j(x_0^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ on the time interval $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$. Thus, the proof of this theorem repeats the proof of Theorem 4.2.7 down to (4.102).

In the game with dynamic updating, we update information only at discrete time instants $t_j = t_0 + j\Delta t, j \in \mathbb{N}$, and specify (4.102) for the game with dynamic updating

$$v_{fb,i}^{t_j}(s, x^t) = -R_{ii}^{-1}(s) B_i'(s) Z_i^{t_j} \left(\frac{s - t_j}{\bar{T}} \right) \bar{T}x^t(s),$$

which is equal to (4.112). □

Theorem 4.2.10 For an N -person linear-quadratic differential game $\bar{\Gamma}_j(x_0^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ with dynamic updating with $Q_i(\cdot) \geq 0$ and $R_{ik}(\cdot) \geq 0$ ($i, k \in N, i \neq k$), let there exist a solution set $\{M_i^{t_j}, i \in N, t_j \geq t_0\}$ for the coupled matrix Riccati differential equations

$$\begin{aligned} \frac{dM_i^{t_j}(\tau)}{d\tau} + \bar{T}M_i^{t_j}(\tau)A(t_j + \bar{T}\tau) + \bar{T}A'(t_j + \bar{T}\tau)M_i^{t_j}(\tau) + Q_i(t_j + \bar{T}\tau) - \\ - \bar{T}^2 M_i^{t_j}(\tau) \sum_{k \in N} B_k(t_j + \bar{T}\tau) (R_{kk}(t_j + \bar{T}\tau))^{-1} B_k'(t_j + \bar{T}\tau) M_k^{t_j}(\tau) = 0, \end{aligned} \quad (4.113)$$

$$M_i^{t_j}(1) = 0, \quad \tau \in [0, 1], \quad i \in N.$$

Then the differential game with dynamic updating admits an open-loop-based Nash equilibrium with dynamic updating given by

$$\begin{aligned} v_{ol,i}^{NE}(t, x) = -R_{ii}^{-1}(t)B_i'(t)M_i^{t_j} \left(\frac{t - (t_0 + j\Delta t)}{\bar{T}} \right) \times \\ \times \Phi^{t_j} \left(\frac{t - (t_0 + j\Delta t)}{\bar{T}} \right) \bar{T}x(t_0 + j\Delta t), \quad j = 0, 1, 2, \dots, \quad i \in N, \end{aligned}$$

where $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, Φ^{t_j} is the solution of the equation

$$\frac{d\Phi^t}{d\tau} = \left(A(t + \bar{T}\tau) - \sum_{i \in N} B_i(t + \bar{T}\tau)R_{ii}^{-1}(t + \bar{T}\tau)B_i'(t + \bar{T}\tau) \right) \Phi^t(\tau),$$

$$\Phi^t(0) = E.$$

Proof The proof of this theorem can be obtained using Theorem 4.2.8 like the proof of Theorem 4.2.9 using Theorem 4.2.7. \square

4.2.5.4.3 Convergence of Resulting Nash Equilibrium Strategies and Trajectory

Now we show the convergence of the Nash equilibrium and corresponding equilibrium trajectory in the differential game with dynamic updating to the corresponding equilibria in the differential game with continuous updating.

Lemma 4.2.2 *Let some matrix function $U(t)$ be uniformly bounded for $t \geq t_0$, some parametric matrix function $P^t(\tau)$ be continuously differentiable with respect to both t and τ , and $\frac{dP^t(\tau)}{dt}$ and $\frac{dP^t(\tau)}{d\tau}$ be uniformly bounded for $t \geq t_0$, $\tau \in [0, 1]$.*

For $x \in X$, where X is a bounded set, and $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $\Delta t \leq \bar{T}$, the following convergence holds as $\Delta t \rightarrow 0$:

$$U(t)P^t(0)x \underset{[t_0, +\infty)}{\rightrightarrows} U(t)P^{t_0+j\Delta t} \left(\frac{t - t_0 + j\Delta t}{\bar{T}} \right) x.$$

Proof We introduce the notation $t_j \stackrel{\text{def}}{=} t_0 + j\Delta t$. Then $t \in [t_j, t_{j+1}]$. We need to show that $\left\| U(t)P^t(0)x - U(t)P^{t_j} \left(\frac{t-t_j}{\bar{T}} \right) x \right\| \rightarrow 0$ as $\Delta t \rightarrow 0$.

From the Taylor expansion for $P^t(\tau)$ at the point $t = t_j$ we obtain

$$P^t(\tau) = P^{t_j}(\tau) + \left. \frac{dP^t(\tau)}{dt} \right|_{t=t_j} (t - t_j) + o(t - t_j).$$

From the Taylor expansion for $P^{t_j}(\tau)$ at the point $\tau = 0$ we obtain

$$P^{t_j}(\tau) = P^{t_j}(0) + \left. \frac{dP^{t_j}(\tau)}{d\tau} \right|_{\tau=0} \tau + o(\tau).$$

As the result, we have the upper bound:

$$\begin{aligned} & \left\| U(t)P^t(0)x - U(t)P^{t_j} \left(\frac{t - t_j}{\bar{T}} \right) x \right\| \leq \\ & \leq \|U(t)\| \|x\| \left(\left\| \left. \frac{dP^t(\tau)}{dt} \right|_{t=t_j} \right\| \Delta t + \left\| \left. \frac{dP^{t_j}(\tau)}{d\tau} \right|_{\tau=0} \right\| \Delta t + o(\Delta t) \right), \end{aligned} \quad (4.114)$$

where $\tau = \frac{t-t_j}{\bar{T}}$.

As $\Delta t \rightarrow 0$, the right-hand side of (4.114) tends to zero, and consequently, the left-hand side of (4.114) does so. The proof is complete. \square

Theorem 4.2.11 *Let the conditions of Theorem 4.2.7 be satisfied, the Nash equilibrium be unique in the game $\Gamma(x, t, t + \bar{T}) \forall t \geq t_0$, $R_{ii}^{-1}(t)B_i'(t)$ be uniformly*

bounded for $t \geq t_0$, the solution $Z_i^t(\tau)$ of the Riccati equation (4.96) be continuously differentiable with respect to t and τ , and $\frac{dZ_i^t(\tau)}{dt}$ and $\frac{dZ_i^t(\tau)}{d\tau}$ be uniformly bounded for $t \geq t_0$, $\tau \in [0, 1]$.

For $x \in X$, where X is a bounded set, the feedback-based Nash equilibrium with dynamic updating $v_{fb,i}^{NE}(t, x)$ uniformly converges to the feedback-based Nash equilibrium with continuous updating $u_{fb,i}^{NE}(t, x)$ as $\Delta t \rightarrow 0$:

$$v_{fb,i}^{NE}(t, x) \underset{[t_0, +\infty)}{\rightrightarrows} u_{fb,i}^{NE}(t, x), \quad i \in N. \quad (4.115)$$

Proof Denoting $t_j \stackrel{\text{def}}{=} t_0 + j\Delta t$, for $t \in [t_j, t_{j+1}]$ we consider the expressions for $u_{fb,i}^{NE}$ and $v_{fb,i}^{NE}$:

$$\begin{aligned} u_{fb,i}^{NE}(t, x) &= -R_{ii}^{-1}(t)B_i'(t)Z_i^t(0)\bar{T}x, \\ v_{fb,i}^{NE}(t, x) &= -R_{ii}^{-1}(t)B_i'(t)Z_i^{t_j}\left(\frac{t-t_j}{T}\right)\bar{T}x, \quad t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t], \end{aligned}$$

where $Z_i^t(\tau)$ is the solution of (4.96). Let $U(t) = -\bar{T}R_{ii}^{-1}(t)B_i'(t)$ and $P^t(\tau) = Z_i^t(\tau)$. Then application of Lemma 4.2.2 completes the proof. \square

Theorem 4.2.12 *Let the conditions of Theorem 4.2.8 be satisfied, the Nash equilibrium be unique in the game $\Gamma(x, t, t + \bar{T}) \forall t \geq t_0$, $R_{ii}^{-1}(t)B_i'(t)$ be uniformly bounded for $t \geq t_0$, the product of the solution $M_i^t(\tau)$ of the Riccati equation (4.111) and $\Phi^t(\tau)$ be continuously differentiable with respect to both t and τ , and $\frac{d(M_i^t(\tau)\Phi^t(\tau))}{dt}$ and $\frac{d(M_i^t(\tau)\Phi^t(\tau))}{d\tau}$ be uniformly bounded for $t \geq t_0$, $\tau \in [0, 1]$, where*

$$\frac{d\Phi^t}{d\tau} = \left(A(t + \bar{T}\tau) - \sum_{i \in N} B_i(t + \bar{T}\tau)R_{ii}^{-1}(t + \bar{T}\tau)B_i'(t + \bar{T}\tau) \right) \Phi^t(\tau),$$

$$\Phi^t(0) = E.$$

For $x \in X$, where X is a bounded set, the open-loop-based Nash equilibrium with dynamic updating $v_{ol,i}^{NE}(t, x)$ uniformly converges to the open-loop-based Nash equilibrium with continuous updating $u_{ol,i}^{NE}(t, x)$ as $\Delta t \rightarrow 0$:

$$v_{ol,i}^{NE}(t, x) \underset{[t_0, +\infty)}{\rightrightarrows} u_{ol,i}^{NE}(t, x), \quad i \in N. \quad (4.116)$$

Proof Denoting $t_j \stackrel{\text{def}}{=} t_0 + j\Delta t$, for $t \in [t_j, t_{j+1}]$ we consider the expressions for $u_{ol,i}^{NE}$ and $v_{ol,i}^{NE}$:

$$\begin{aligned} u_{ol,i}^{NE}(t, x) &= -R_{ii}^{-1}(t)B'_i(t)M_i^t(0)\bar{T}x, \\ v_{ol,i}^{NE}(t, x) &= -R_{ii}^{-1}(t)B'_i(t)M_i^{t_j} \left(\frac{t - (t_0 + j\Delta t)}{\bar{T}} \right) \times \\ &\quad \times \Phi^{t_j} \left(\frac{t - (t_0 + j\Delta t)}{\bar{T}} \right) \bar{T}x, \quad t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t], \end{aligned}$$

where $M_i^t(\tau)$ is the solution of (4.111).

Let $U(t) = -\bar{T}R_{ii}^{-1}(t)B'_i(t)$ and $P^t(\tau) = M_i^t(\tau)\Phi^t(\tau)$. Then application of Lemma 4.2.2 completes the proof. \square

Lemma 4.2.3 *Let some matrix functions $V_1(t)$ and $V_2(t)$ be uniformly bounded for $t \geq t_0$, some parametric matrix function $P^t(\tau)$ be continuously differentiable with respect to both t and τ , and $\frac{dP^t(\tau)}{t}$ and $\frac{dP^t(\tau)}{\tau}$ be uniformly bounded for $t \geq t_0$ and $\tau \in [0, 1]$.*

Consider a differential equation for $y(t)$, $t \in [t_0, +\infty]$, of the form

$$\frac{dy(t)}{dt} = (V_1(t) + V_2(t)P^t(0))y, \quad (4.117)$$

where $y(t_0) = y_0$.

Consider a differential equation for $z_j(t)$, $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, of the form

$$\frac{dz_j(t)}{dt} = \left(V_1(t) + V_2(t)P^{t_0+j\Delta t} \left(\frac{t - (t_0 + j\Delta t)}{\bar{T}} \right) \right) z_j, \quad (4.118)$$

where $z_j(t_0 + j\Delta t) = z_{j-1}(t_0 + j\Delta t)$ for $j \geq 1$, $z_0(t_0) = y_0$.

Let $y^(t)$ be the solution of (4.117) and*

$$z^*(t) = \begin{cases} z_0(t), & t \in [t_0, t_0 + \Delta t], \\ \dots \\ z_j(t), & t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t], \\ \dots \end{cases}$$

where $z_j(t)$ satisfies (4.118). Let $y^*(t)$ and $z^*(t)$ exist for $t \geq t_0$. Then $z^*(t)$ pointwise converges to $y^*(t)$ as $\Delta t \rightarrow 0$:

$$z^*(t) \xrightarrow{\Delta t \rightarrow 0} y^*(t).$$

Proof We introduce the notation $t_j \stackrel{\text{def}}{=} t_0 + j\Delta t$. Then $t \in [t_j, t_{j+1}]$. We need to show that $\|z^*(t) - y^*(t)\| \rightarrow 0$ as $\Delta t \rightarrow 0$.

The trajectories $z^*(t)$ and $y^*(t)$ satisfy the differential equations (4.118), (4.117) for $t \in [t_j, t_{j+1}]$. Note that

$$P^t(0)y - P^{t_j} \left(\frac{t - t_j}{T} \right) z = P^t(0)(y - z) + \left(P^t(0) - P^{t_j} \left(\frac{t - t_j}{T} \right) \right) z.$$

Let $w_j(t) = y(t) - z(t)$ for $t \in [t_j, t_{j+1}]$, $V_3(t) = V_1(t) + V_2(t)P^t(0)$ and

$$f_j(t) = V_2(t) \left(P^t(0) - P^{t_j} \left(\frac{t - t_j}{T} \right) \right) z(t).$$

Then $w(t)$ satisfies the differential equation

$$\frac{dw_j(t)}{dt} = V_3(t)w_j(t) + f_j(t).$$

Consider

$$w(t) = \begin{cases} w_0(t), & t \in [t_0, t_0 + \Delta t], \\ \dots \\ w_j(t), & t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t], \\ \dots \end{cases} \quad (4.119)$$

and

$$f(t) = \begin{cases} f_0(t), & t \in [t_0, t_0 + \Delta t], \\ \dots \\ f_j(t), & t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t], \\ \dots \end{cases}$$

Then (4.119) satisfies the differential equation

$$\dot{w}(t) = V_2(t)w(t) + f(t).$$

with the initial state $w(t_0) = 0$ since $z(t_0) = y(t_0)$.

By the Cauchy formula, we have for any $t \geq t_0$:

$$w(t) = Y(t) \int_{t_0}^t Y^{-1}(\xi) f(\xi) d\xi,$$

where $Y(t)$ is the fundamental matrix of $\frac{dw}{dt} = V_2(t)w(t)$. Taking this into account, we have for fixed t :

$$\lim_{\Delta t \rightarrow 0} \|w(t)\| \leq \lim_{\Delta t \rightarrow 0} [\|Y(t)\| \hat{w}(t) \beta \Delta t (t - t_0) + o(\Delta t)] = 0, \quad (4.120)$$

where

$$\begin{aligned} \hat{w}(t) &= \max_{\tau \in [t_0, t]} \|Y^{-1}(\tau)\| \\ \beta &= \sum_{i=1}^N \|V_2(t)\| \left(\left\| \frac{dP^t(\tau)}{dt} \Big|_{t=t_j} \right\| + \left\| \frac{dP^{t_j}(\tau)}{d\tau} \Big|_{\tau=0} \right\| \right) M(t), \\ M(t) &= \max_{\tau \in [t_0, t]} \|z(\tau)\|. \end{aligned}$$

According to (4.120), $w(t) \xrightarrow{[t_0, +\infty)} 0$ as $\Delta t \rightarrow 0$. This proves the lemma. \square

Theorem 4.2.13 *Let the conditions of Theorem 4.2.7 be satisfied, the Nash equilibrium be unique in the game $\Gamma(x, t, t + \bar{T}) \forall t \geq t_0$, the matrices $A(t)$ and $B_i(t)R_{ii}^{-1}(t)B_i'(t)$ be uniformly bounded for $i \in N$ and $t \geq t_0$, the solution of the Riccati equation (4.96) $Z_i^t(\tau)$ be continuously differentiable with respect to both t and τ , and $\frac{dZ_i^t(\tau)}{dt}$ and $\frac{dZ_i^t(\tau)}{d\tau}$ be uniformly bounded for $t \geq t_0, \tau \in [0, 1]$.*

Then the feedback-based equilibrium trajectory in the game with dynamic updating

$\hat{x}_{fb}^{NE}(t)$ pointwise converges to the feedback-based equilibrium trajectory $\tilde{x}_{fb}^{NE}(t)$ in

the game with continuous updating as $\Delta t \rightarrow 0$:

$$\hat{x}_{fb}^{NE}(t) \xrightarrow{[t_0, +\infty)} \tilde{x}_{fb}^{NE}(t). \quad (4.121)$$

Proof Let

$$\begin{aligned} V_1(t) &= A(t), \\ V_2(t) &= [-\bar{T}B_1(t)R_{11}^{-1}(t)B'_1(t), \dots, -\bar{T}B_N(t)R_{NN}^{-1}(t)B'_N(t)], \\ P^t(\tau) &= \begin{bmatrix} Z_1^t \\ \vdots \\ Z_N^t \end{bmatrix}. \end{aligned}$$

Then application of Lemma 4.2.3 completes the proof. \square

Theorem 4.2.14 *Let the conditions of Theorem 4.2.8 be satisfied, the Nash equilibrium be unique in the game $\Gamma(x, t, t + \bar{T}) \forall t \geq t_0$, the matrices $A(t)$ and $B_i(t)R_{ii}^{-1}(t)B'_i(t)$ be uniformly bounded for $i \in N$ and $t \geq t_0$, the product $M_i^t(\tau)\Phi^t(\tau)$ be continuously differentiable with respect to both t and τ , and $\frac{dM_i^t(\tau)\Phi^t(\tau)}{dt}$ and $\frac{dM_i^t(\tau)\Phi^t(\tau)}{d\tau}$ be uniformly bounded for $t \geq t_0, \tau \in [0, 1]$, where $M_i^t(\tau)$ is the solution of the Riccati equation (4.111) and $\Phi^t(\tau)$ is the solution of*

$$\begin{aligned} \frac{d\Phi^t}{d\tau} &= \left(A(t + \bar{T}\tau) - \sum_{i \in N} B_i(t + \bar{T}\tau)R_{ii}^{-1}(t + \bar{T}\tau)B'_i(t + \bar{T}\tau) \right) \Phi^t(\tau), \\ \Phi^t(0) &= E. \end{aligned}$$

Then the open-loop-based equilibrium trajectory in the game with dynamic updating $\hat{x}_{ol}^{NE}(t)$ pointwise converges to the open-loop-based equilibrium trajectory in the game with continuous updating $\tilde{x}_{ol}^{NE}(t)$ as $\Delta t \rightarrow 0$:

$$\hat{x}_{ol}^{NE}(t) \xrightarrow{\Delta t \rightarrow 0} \tilde{x}_{ol}^{NE}(t). \quad (4.122)$$

Proof Let

$$\begin{aligned} V_1(t) &= A(t), \\ V_2(t) &= [-\bar{T}B_1(t)R_{11}^{-1}(t)B'_1(t), \dots, -\bar{T}B_N(t)R_{NN}^{-1}(t)B'_N(t)], \end{aligned}$$

$$P^t(\tau) = \begin{bmatrix} M_1^t(\tau)\Phi^t(\tau) \\ \vdots \\ M_N^t(\tau)\Phi^t(\tau). \end{bmatrix}$$

Then applying Lemma 4.2.3 completes the proof. \square

4.2.5.5 Differential Game Model of Government Debt Stabilization with Continuous Updating

4.2.5.5.1 Common Description

Consider the following differential game based on government debt stabilization [226]. A dynamic government budget constraint shows the relation between the primary fiscal deficits, $f(t)$, the monetization or seignorage, $m(t)$, the interest payments on government debt, $rd(t)$, and the government debt accumulation $\dot{d}(t)$:

$$\begin{aligned} \dot{d}(t) &= rd(t) + f(t) - m(t), \\ d(t_0) &= d_0, \quad x^t \in \mathbb{R}^1. \end{aligned}$$

Here $d(t)$, $f(t)$, and $m(t)$ are expressed as fractions of GDP, and $r > 0$ represents the rate of interest on outstanding government debt minus the growth rate of output. Assume that fiscal and monetary policies are controlled by different institutions with different objectives (payoff functions). The objective of the fiscal authority is to minimize the sum of time profiles of the primary fiscal deficit, base-money growth, and government debt:

$$J_1 = \int_{t_0}^{\infty} e^{-\delta t} (f^2(t) + \eta m^2(t) + \lambda d^2(t)) dt.$$

The parameters η and λ express the relative priority assigned to base-money growth and government debt by the fiscal authority. The monetary authorities are assumed to choose the growth of base money for minimizing the sum of time profiles of

base-money growth and government debt:

$$J_2 = \int_{t_0}^{\infty} e^{-\delta t} (m^2(t) + \kappa d^2(t)) dt.$$

Here $1/\kappa$ can be interpreted as a measure of the central bank's conservatism with respect to money growth. Furthermore, all variables are normalized so that their targets are zero, and all parameters are positive.

4.2.5.5.2 Game Model with Continuous Updating

Now consider the non-autonomous case with continuous updating. Suppose that at each time instant $t \in [t_0, +\infty)$, both institutions use information about the motion equations and payoff functions on the interval $[t, t + \bar{T}]$. As the current time t evolves, the interval defining the information shifts as well. The motion equations in the game model with continuous updating have the form

$$\dot{d}^t(s) = rd^t(s) + f^t(s, d) - m^t(s, d), \quad d^t(t) = d. \quad (4.123)$$

The cost functions of players in the game model with continuous updating are defined as

$$\begin{aligned} J_1^t &= \int_t^{t+\bar{T}} e^{-\delta s} \left((f^t(s))^2 + \eta (m^t(s, d))^2 + \lambda (d^t(s, d))^2 \right) ds, \\ J_2^t &= \int_t^{t+\bar{T}} e^{-\delta s} \left((m^t(s, d))^2 + \kappa (d^t(s, d))^2 \right) ds. \end{aligned} \quad (4.124)$$

Note that the non-autonomous nature of this game model is expressed through the additional term $e^{-\delta s}$, which represents the dependence of the payoff function on the current time s .

For the sake of analytical solvability, let $\eta = 0$, $\lambda = \kappa = q$, and $\delta = 2r$. Then the players are symmetric. When defining the form of feedback-based Nash equilibrium with continuous updating (see Theorem 4.2.7), in the first step we need to solve the

following differential equation:

$$\begin{cases} \dot{k}^t(\tau) = 3\bar{T}^2 e^{2r(t+\bar{T}\tau)} (k^2(\tau))^2 - 2\bar{T}rk(\tau) - qe^{-2r(t+\bar{T}\tau)}, \\ k^t(1) = 0. \end{cases} \quad (4.125)$$

According to [227], the solution of (4.125) is

$$k^t(\tau) = \gamma e^{-2r(t+\bar{T}\tau)} \tanh\left(3\bar{T}^2\gamma(1-\tau)\right), \quad (4.126)$$

where $\gamma = \sqrt{\frac{q}{3\bar{T}^2}}$. According to (4.97), the feedback-based Nash equilibrium with continuous updating has the form

$$\begin{aligned} \tilde{f}^{NE}(t, d) &= -\bar{T}k^t(0)d, \\ \tilde{m}^{NE}(t, d) &= \bar{T}k^t(0)d. \end{aligned} \quad (4.127)$$

Substituting (4.126) into (4.127), we obtain

$$\begin{aligned} \tilde{f}^{NE}(t, d) &= -\gamma\bar{T}de^{-2rt} \tanh\left(3\bar{T}^2\gamma\right), \\ \tilde{m}^{NE}(t, d) &= \gamma\bar{T}de^{-2rt} \tanh\left(3\bar{T}^2\gamma\right). \end{aligned} \quad (4.128)$$

Substituting (4.128) into (4.123), we obtain $\tilde{d}^{NE}(t)$ as the solution of the equation

$$\begin{aligned} \dot{d}(t) &= \left(r - 2\gamma\bar{T}e^{-2rt} \tanh\left(3\bar{T}^2\gamma\right)\right) d(t), \\ d(t_0) &= d_0, \end{aligned} \quad (4.129)$$

where $\gamma = \sqrt{\frac{q}{3\bar{T}^2}}$, $t \in [t_0, +\infty)$.

4.2.5.5.3 Game Model with Dynamic Updating

Similar calculations can be performed for the Nash equilibrium with dynamic updating based on the calculations for the original game and the approach described

in Section 4.2.5.4.1. As the result, we obtain:

$$\begin{aligned}\hat{f}^{NE}(t, d) &= -\bar{T}k^{t_i} \left(\frac{t - t_i}{\bar{T}} \right) d, \\ \hat{m}^{NE}(t, d) &= \bar{T}k^{t_i} \left(\frac{t - t_i}{\bar{T}} \right) d,\end{aligned}\tag{4.130}$$

where $t \in [t_i, t_{i+1}]$. Substituting (4.126) into (4.130), we obtain:

$$\begin{aligned}\hat{f}^{NE}(t, d) &= -\gamma\bar{T}de^{-2rt} \tanh \left(3\bar{T}^2\gamma \left(1 - \frac{t - t_i}{\bar{T}} \right) \right), \\ \hat{m}^{NE}(t, d) &= \gamma\bar{T}de^{-2rt} \tanh \left(3\bar{T}^2\gamma \left(1 - \frac{t - t_i}{\bar{T}} \right) \right).\end{aligned}\tag{4.131}$$

Substituting (4.131) into (4.123), we obtain $\hat{d}^{NE}(t)$ as the solution of the equation

$$\begin{aligned}\dot{d}(t) &= \left(r - 2\gamma\bar{T}e^{-2rt} \tanh \left[3\bar{T}^2\gamma \left(1 - \frac{t - t_i}{\bar{T}} \right) \right] \right) d(t), \\ d(t_0) &= d_0,\end{aligned}\tag{4.132}$$

where $\gamma = \sqrt{\frac{q}{3\bar{T}^2}}$.

4.2.5.5.4 Autonomous and Non-autonomous Cases

Suppose that information does not change in time. That is, consider the autonomous case with continuous updating. For this purpose, we fix $t = t_0$ in (4.123) and (4.124). Then we find the following strategies:

$$\begin{aligned}\tilde{f}_{t_0}^{NE}(t, d) &= -\gamma\bar{T}de^{-2rt_0} \tanh \left(3\bar{T}^2\gamma \right), \\ \tilde{m}_{t_0}^{NE}(t, d) &= \gamma\bar{T}de^{-2rt_0} \tanh \left(3\bar{T}^2\gamma \right).\end{aligned}\tag{4.133}$$

Substituting (4.133) into (4.123), we obtain $\tilde{d}_{t_0}^{NE}(t)$ as the solution of the equation

$$\begin{aligned}\dot{d}(t) &= \left(r - 2\gamma\bar{T}e^{-2rt_0} \tanh \left(3\bar{T}^2\gamma \right) \right) d(t), \\ d(t_0) &= d_0,\end{aligned}\tag{4.134}$$

where $\gamma = \sqrt{\frac{q}{3T^2}}$, $t \in [t_0, +\infty)$.

4.2.5.5.5 Solution without Updating

Consider the following change of variables:

$$\bar{d}(t) = e^{-\frac{1}{2}\delta t}d(t), \quad \bar{m}(t) = e^{-\frac{1}{2}\delta t}m(t), \quad \bar{f}(t) = e^{-\frac{1}{2}\delta t}f(t).$$

It reduces the example model under consideration to the classical game model from [224]. Then the system dynamics are described by the system of differential equations

$$\begin{aligned} \dot{\bar{d}}(t) &= \left(r - \frac{1}{2}\delta\right) \bar{d} + \bar{f} - \bar{m}, \\ \bar{d}(t_0) &= e^{-\frac{1}{2}\delta t_0}d_0, \end{aligned}$$

and the cost functions of players are defined as

$$\begin{aligned} \bar{J}_1 &= \int_{t_0}^{\infty} (\bar{f}^2(t) + \eta\bar{m}^2(t) + \lambda\bar{d}^2(t)) dt, \\ \bar{J}_2 &= \int_{t_0}^{\infty} (\bar{m}^2(t) + \kappa\bar{d}^2(t)) dt. \end{aligned}$$

Assume that $\eta = 0$, $\lambda = \kappa = q$, and $\delta = 2r$ for correct comparison with the continuous updating setting. According to [224], the feedback-based Nash equilibrium is

$$\bar{f}^{NE}(t, \bar{d}) = -\sqrt{\frac{q}{2}}\bar{d}, \quad \bar{m}^{NE}(t, \bar{d}) = \sqrt{\frac{q}{2}}\bar{d}.$$

Then we obtain the following feedback-based Nash equilibrium in the original variables:

$$\begin{aligned} f^{NE}(t, d) &= -e^{-2rt}\sqrt{\frac{q}{2}}d, \\ m^{NE}(t, d) &= e^{-2rt}\sqrt{\frac{q}{2}}d. \end{aligned} \tag{4.135}$$

Substituting (4.135) into (4.123), we obtain $d^{NE}(t)$ as the solution of the equation

$$\begin{aligned} \dot{d}(t) &= \left(r - e^{-2rt} \sqrt{2q} \right) d(t), \\ d(t_0) &= d_0. \end{aligned} \quad (4.136)$$

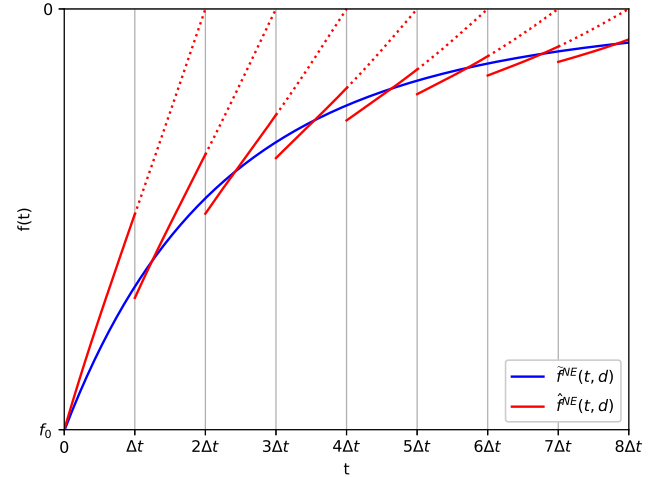
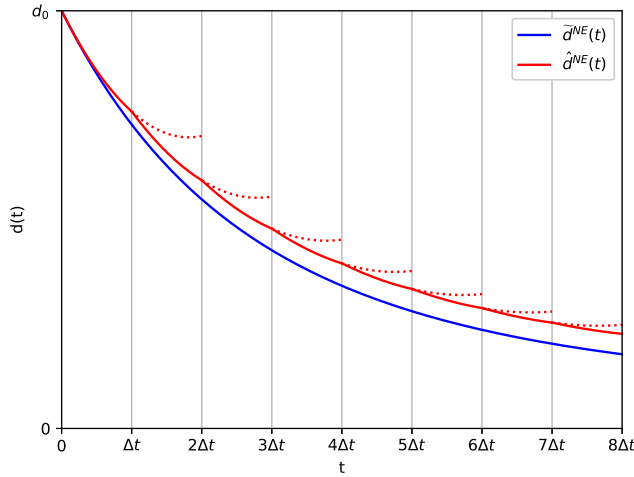


Figure 4.11: $\tilde{d}^{NE}(t)$ (4.129) — blue line, $\hat{d}^{NE}(t)$ (4.132) — red broken line.

Figure 4.12: $\tilde{f}^{NE}(t)$ (4.128) — blue line, $\hat{f}^{NE}(t)$ (4.131) — red broken line.

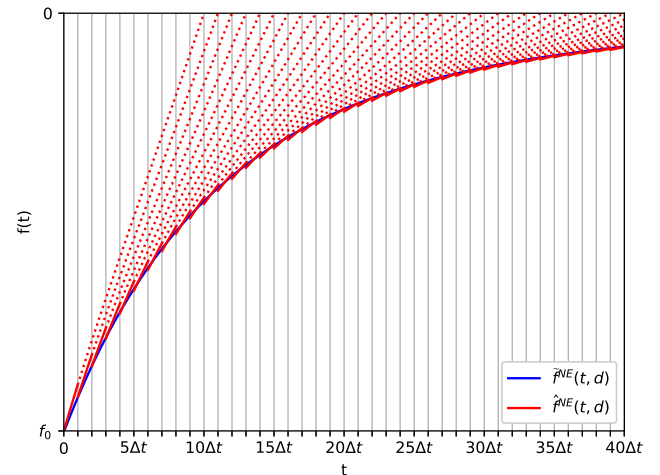
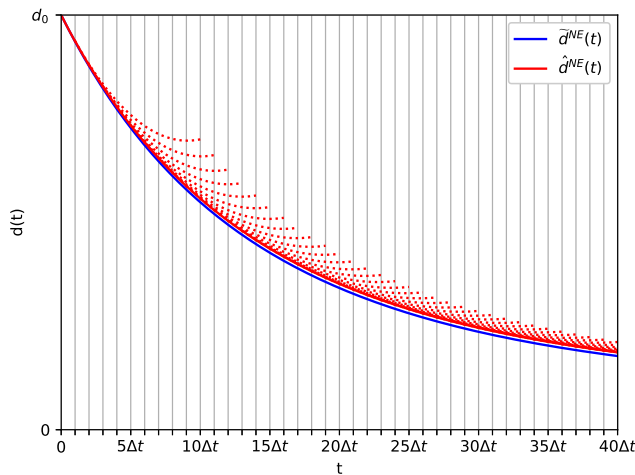


Figure 4.13: $\tilde{d}^{NE}(t)$ (4.129) — blue line, $\hat{d}^{NE}(t)$ (4.132) — red broken line.

Figure 4.14: $\tilde{f}^{NE}(t)$ (4.128) — blue line, $\hat{f}^{NE}(t)$ (4.131) — red broken line.

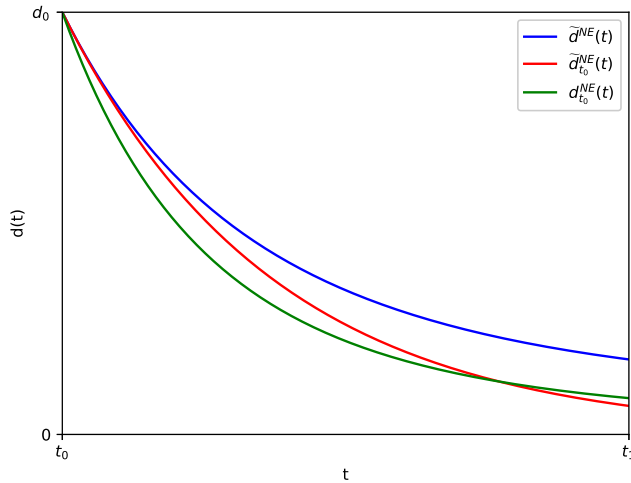


Figure 4.15: $\tilde{d}^{NE}(t)$ (4.129) — blue line, $\tilde{d}_{t_0}^{NE}(t)$ (4.134) — red line, $d^{NE}(t)$ (4.136) — green line.

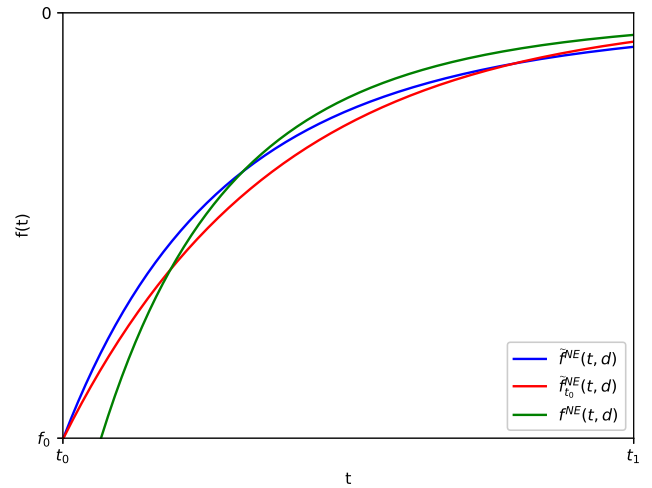


Figure 4.16: $\tilde{f}^{NE}(t, x)$ (4.128) — blue line, $\tilde{f}_{t_0}^{NE}(t, x)$ (4.133) — red line, and $f^{NE}(t, x)$ (4.135) — green line.

4.2.5.5.6 Numerical Simulation

Consider the numerical simulation results for the game model presented above on the interval $[1, 5]$, i.e., $t_0 = 1$ and $T = 5$. At the initial instant $t_0 = 1$, the stock of knowledge is 1, i.e., $x_0 = 1$. The other parameters of the models are $r = 0.1$, $q = 0.8$, and $\bar{T} = 1$. Suppose that under dynamic updating (red solid and dotted lines in Figures 4.25 - 4.26), the intervals between the updating instants are $\Delta t = 0.5$. Therefore, the number of intervals in the game with dynamic updating is $l = \frac{T-t_0}{\Delta t} = 8$. In Figure 4.25, the comparison of the Nash equilibrium with dynamic updating (red line) and Nash equilibrium with continuous updating (blue line) is presented. In Figure 4.26, similar results are presented for the strategies.

To illustrate Theorems 4.2.11 and 4.2.13 on the convergence of the equilibrium strategies with dynamic updating and the corresponding trajectory to the equilibrium strategies and trajectory with continuous updating, consider the case of frequent updating, namely, $l = \frac{T-t_0}{\Delta t} = 40$. Figures 4.27-4.28 represent the same solutions as in Figures 4.25 - 4.26, but for the case $\Delta t = 0.1$. Therefore, the convergence results are confirmed by the numerical experiments presented below.

To compare the autonomous and non-autonomous cases with continuous updating, consider the simulation results in $t \in [1, 5]$, with the parameters $r = 0.1$,

$q = 0.8$, $\bar{T} = 1$, and $x_0 = 1$. The strategies $\tilde{f}^{NE}(t, x)$ and $\tilde{f}_{t_0}^{NE}(t, x)$ obtained in the non-autonomous and autonomous cases, respectively, are presented in Figure 4.16. The additional obtained trajectories $\tilde{d}^{NE}(t)$ and $\tilde{d}_{t_0}^{NE}(t)$ are presented in Figure 4.15. Clearly, there is a difference between the strategies with continuous updating and corresponding strategies in the initial differential game model without updating $f^{NE}(t)$ and the corresponding trajectory $d^{NE}(t)$.

4.3 Cooperative Differential Game Model with Continuous Updating

In the cooperative differential game model with transferable utility, there are also two problems: defining the cooperative behavior using cooperative strategies and defining the cooperative solutions. Both problems should be addressed using the continuous updating approach.

4.3.1 Cooperative Game Model with Continuous Updating

4.3.1.1 Cooperative Strategies in Game with Continuous Updating

Under continuously updated information, it is important to model the behavior of players. In this section, we consider the case of cooperative strategies. For the class of differential games with continuous updating, we shall give it the following form:

- for any fixed $t \in [t_0, +\infty)$, $u^*(t, x) = (u_1^*(t, x), \dots, u_n^*(t, x))$ coincides with the cooperative strategies in the game (4.5), (4.6) defined on the interval $[t, t + \bar{T}]$ at the instant t .

However, direct application of the classical approaches for defining the cooperative strategies is not possible. Consider two intervals, $[t, t + \bar{T}]$ and $[t + \epsilon, t + \bar{T} + \epsilon]$, where $\epsilon \ll \bar{T}$. Then, according to the problem statement, we have:

- $u^*(t, x)$ at the instant t coincides with the cooperative strategies in the game defined on the interval $[t, t + \bar{T}]$.

- $u^*(t + \epsilon, x)$ at the instant $t + \epsilon$ must coincide with the cooperative strategies in the game defined on the interval $[t + \epsilon, t + \epsilon + \bar{T}]$.

To construct such strategies, we consider the concept of generalized *feedback-based (open-loop-based)* cooperative strategies

$$\tilde{u}_{fb}^*(t, x; s, x_t) = (\tilde{u}_{fb,1}^*(t, x; s, x_t), \dots, \tilde{u}_{fb,n}^*(t, x; s, x_t)), \quad t \in [t_0, T], \quad s \in [t, t + \bar{T}],$$

$$(\tilde{u}_{ol}^*(t, x; s) = (\tilde{u}_{ol,1}^*(t, x; s), \dots, \tilde{u}_{ol,n}^*(t, x; s)), \quad t \in [t_0, T], \quad s \in [t, t + \bar{T}]).$$

They will be used further for constructing the strategies $u_{fb}^*(t, x)$ ($u_{ol}^*(t, x)$). For any fixed t , strategies $\tilde{u}_{fb}^*(t, x; s, x_t) = (\tilde{u}_{fb,1}^*(t, x; s, x_t), \dots, \tilde{u}_{fb,n}^*(t, x; s, x_t))$ ($\tilde{u}_{ol}^*(t, x; s) = (\tilde{u}_{ol,1}^*(t, x; s), \dots, \tilde{u}_{ol,n}^*(t, x; s))$) are defined in the following way:

$$\tilde{u}_{fb}^*(t, x; s, x_t) = (\tilde{u}_{fb,1}^*(t, x; s, x_t), \dots, \tilde{u}_{fb,n}^*(t, x; s, x_t)) = \arg \max_{\tilde{u}_{fb,1}, \dots, \tilde{u}_{fb,n}} \sum_{i=1}^n K_i^t(x, t; \tilde{u}(t)), \quad (4.137)$$

$$\left(\tilde{u}_{ol}^*(t, x; s) = (\tilde{u}_{ol,1}^*(t, x; s), \dots, \tilde{u}_{ol,n}^*(t, x; s)) = \arg \max_{\tilde{u}_{ol,1}, \dots, \tilde{u}_{ol,n}} \sum_{i=1}^n K_i^t(x, t; \tilde{u}(t)) \right). \quad (4.138)$$

For a fixed current time instant t , the optimization problem (4.137) (optimization problem (4.138)) is constructed based on the information available to the players at the instant t , i.e., over the interval $[t, t + \bar{T}]$. Suppose that the maximum in (4.137) (in (4.138)) is achieved on the set of admissible strategies $\tilde{u}_{fb}^*(t, x; s, x_t)$ ($\tilde{u}_{ol}^*(t, x; s)$) with fixed t .

Definition 4.3.1 *A strategy profile $\tilde{u}_{fb}^*(t, x; s, x_t) = (\tilde{u}_{fb,1}^*(t, x; s, x_t), \dots, \tilde{u}_{fb,n}^*(t, x; s, x_t))$ ($\tilde{u}_{ol}^*(t, x; s) = (\tilde{u}_{ol,1}^*(t, x; s), \dots, \tilde{u}_{ol,n}^*(t, x; s))$) forms the generalized feedback-based (open-loop-based) cooperative strategies in the game with continuous updating if for any fixed $t \in [t_0, +\infty)$, the components of $\tilde{u}_{fb}^*(t, x; s, x_t)$ ($\tilde{u}_{ol}^*(t, x; s)$) are the feedback (open-loop) cooperative strategies in the game $\Gamma(x, t, t + \bar{T})$.*

Note that for a fixed t , the generalized feedback-based (open-loop-based) cooperative strategies $\tilde{u}_{fb}^*(t, x; s, x_t)$ ($\tilde{u}_{ol}^*(t, x; s)$) are functions of s and x , where s is defined for the interval $[t, t + \bar{T}]$. Using the generalized cooperative strategies,

we define the cooperative strategies realized in the game model with continuous updating.

Definition 4.3.2 *A strategy profile $u_{fb}^*(t, x)$ ($u_{ol}^*(t, x)$) forms the feedback-based (open-loop-based) cooperative strategies with continuous updating if*

$$u^*(t, x) = \tilde{u}^*(t, s, x)|_{s=t} = (\tilde{u}_1^*(t, s, x)|_{s=t}, \dots, \tilde{u}_n^*(t, s, x)|_{s=t}), \quad (4.139)$$

$$t \in [t_0, +\infty),$$

where $\tilde{u}_{fb}^*(t, x; s, x_t)$ ($\tilde{u}_{ol}^*(t, x; s)$) forms the generalized feedback-based (open-loop-based) cooperative strategies in the sense of Definition 4.3.1.

Unlike the generalized cooperative strategies, $u_{fb}^*(t, x)$ ($u_{ol}^*(t, x)$) does not contain cooperative strategies for any $s \in [t, t + \bar{T}]$. The strategy profile $u_{fb}^*(t, x)$ ($u_{ol}^*(t, x)$) only accounts for the players' strategies when they follow the continuous updating procedure (see Section 4.2) with $s = t$. $u_{fb}^*(t, x)$ ($u_{ol}^*(t, x)$) will be used as the cooperative strategies in the game model with continuous updating.

4.3.1.2 Cooperative Trajectory in Game with Continuous Updating

The trajectory $x^*(t)$ corresponding to the cooperative strategies with continuous updating $u^*(t, x)$ can be obtained from the system

$$\begin{aligned} \dot{x}(t) &= f(t, x, u^*), \\ x(t_0) &= x_0, \\ x &\in \mathbb{R}^l. \end{aligned} \quad (4.140)$$

4.3.1.3 Cooperative Solution with Continuous Updating

To allocate the joint payoff among players, we need defining how the overall game evolves and how the players forecast their behavior at each current time instant $t \in [t_0, +\infty]$ for the future interval $[t, t + \bar{T}]$. Therefore, we need defining:

- 1 the game trajectory (Section 4.3.1.3.1),

- 2 how the characteristic function is calculated along the forecasted trajectory (Section 4.3.1.3.2),
- 3 how the imputation and cooperative solution are chosen along the forecasted trajectory (Section 4.3.1.3.3),

and finally how the cooperative solution with continuous updating is constructed (Section 4.3.1.3.4), including its properties (Section 4.3.1.4).

4.3.1.3.1 Cooperative Trajectories in Game Model on Interval $[t, t + \bar{T}]$

For each differential game model $\Gamma(x, s, t + \bar{T})$ defined on the interval $[t, t + \bar{T}]$, we define a cooperative trajectory realized in the game at the time instant t in the position x based on the information available at this instant. We denote by $\tilde{x}_t^*(s)$ the corresponding cooperative trajectory as the solution of the following system:

$$\begin{aligned} \dot{x}(s) &= f(s, x, \tilde{u}^*(t, s, x)), \\ x(t) &= x \quad (x(t_0) = x_0), \\ x &\in \mathbb{R}^l \end{aligned} \tag{4.141}$$

with $\tilde{u}^*(t, s, x)$ involved with the fixed current time t .

4.3.1.3.2 Characteristic Function on Interval $[t, t + \bar{T}]$

Since the game evolves continuously in time, the procedure for calculating the characteristic function should also be defined in a specific manner. But at this point, let us focus on considering the solution at the current time instant t , i.e., the solution for the game defined on the interval $[t, t + \bar{T}]$. The notion of characteristic function $\tilde{V}_t(S; x, s, t + \bar{T})$, $S \subseteq N$, is defined for each subgame $\Gamma(x, t, t + \bar{T})$ at a fixed current time instant $t \in [t_0, +\infty)$ and a time instant $s \in [t, t + \bar{T}]$. Suppose

that $\tilde{V}_t(S; x, s, t + \bar{T})$, $S \subseteq N$, is constructed as in [228]:

$$\tilde{V}_t(S; x, s, t + \bar{T}) = \begin{cases} \max_{u^t} \sum_{i \in N} K_i^t(x, s, t + \bar{T}; u^t), & S = N, \\ \max_{u_i^t, i \in S} \sum_{i \in S} K_i^t(x, s, t + \bar{T}; u_S^t, u_{N \setminus S}^{t, NE}), & S \subset N, \\ 0, & S = \emptyset, \end{cases}$$

where $\max_{u_i^t, i \in S} \sum_{i \in S} K_i^t(x, s, t + \bar{T}; u_S^t, u_{N \setminus S}^{t, NE})$ is defined as the maximum joint payoff of coalition S provided that the players from coalition $N \setminus S$ use their strategies $u_{N \setminus S}^{t, NE} = \{u_j^{t, NE}\}_{j \in N \setminus S}$ from a fixed Nash equilibrium $u^{t, NE} = (u_1^{t, NE}, \dots, u_n^{t, NE})$ in the noncooperative game $\Gamma(x, t, t + \bar{T})$.

4.3.1.3.3 Cooperative Solution on Interval $[t, t + \bar{T}]$

To construct the resulting cooperative solution in the game model with continuous updating, consider the evolution of the game in time. In this section, we construct the cooperative solution for any fixed current time instant $t \in [t_0, +\infty)$ along the local cooperative trajectory $\tilde{x}_t^*(s)$, $t \in [t, t + \bar{T}]$. $\tilde{x}_t^*(s)$ is the cooperative trajectory in the game (4.5), (4.6) defined on the interval $[t, t + \bar{T}]$.

An imputation $\tilde{\xi}_t(\tilde{x}_t^*(s), s, t + \bar{T})$ in each cooperative subgame $\Gamma(x, t, t + \bar{T})$ is defined as an arbitrary vector that satisfies the conditions:

$$\begin{aligned} \tilde{\xi}_i^t(\tilde{x}_t^*(s), s, t + \bar{T}) &\geq \tilde{V}_t(\{i\}; \tilde{x}_t^*(s), s, t + \bar{T}), \quad i \in N, \\ \sum_{i \in N} \tilde{\xi}_i^t(\tilde{x}_t^*(s), s, t + \bar{T}) &= \tilde{V}_t(N; \tilde{x}_t^*(s), s, t + \bar{T}). \end{aligned} \quad (4.142)$$

We denote by $\tilde{L}_t(\tilde{x}_t^*(s), s, t + \bar{T})$ the set of all possible imputations in each game defined on the interval $[t, t + \bar{T}]$. Suppose that in each game defined on the interval $[t, t + \bar{T}]$, a non-empty solution is defined as

$$\tilde{M}_t(\tilde{x}_t^*(s), s, t + \bar{T}) \subset \tilde{L}_t(\tilde{x}_t^*(s), s, t + \bar{T}). \quad (4.143)$$

$\tilde{M}_t(\tilde{x}_t^*(s), s, t + \bar{T})$ can be the Core, NM solution, Nucleus, or Shapley value.

4.3.1.3.4 Construction of Cooperative Solution in Game with Continuous Updating

To construct the cooperative solution in the game with continuous updating, we will use a combination of the imputations and corresponding IDPs for each current time instant $t \in [t_0, T]$.

Suppose that in each game defined on the interval $[t, t + \bar{T}]$, the solution $\widetilde{M}_t(\tilde{x}_t^*(s), s, t + \bar{T}) \neq \emptyset$ along the cooperative trajectory $\tilde{x}_t^*(s)$ is selected, $s \in [t, t + \bar{T}]$. Also, suppose that for any game $\Gamma(x, t, t + \bar{T})$ in the starting position $\tilde{x}_t^*(s)$, the imputation

$$\tilde{\xi}_t(\tilde{x}_t^*(s), s, t + \bar{T}) \in \widetilde{M}_t(\tilde{x}_t^*(s), s, t + \bar{T}), \quad s \in [t, t + \bar{T}]$$

and the corresponding IDPs are selected

$$\tilde{\beta}_t(s, \tilde{x}_t^*) = [\tilde{\beta}_1^t(s, \tilde{x}_t^*), \dots, \tilde{\beta}_n^t(s, \tilde{x}_t^*)], \quad s \in (t, t + \bar{T}],$$

to ensure the time-consistency of the selected imputation (see Definition 1.1.3):

$$\tilde{\xi}_t(x_t^*(s), s, t + \bar{T}) = \int_s^{t+\bar{T}} \tilde{\beta}_t(\tau, \tilde{x}_t^*) d\tau, \quad s \in (t, t + \bar{T}]. \quad (4.144)$$

Note that condition (4.144) is equivalent to the condition described in (1.13):

$$\tilde{\beta}_t(s, \tilde{x}_t^*) = -\frac{d}{ds} \tilde{\xi}_t(x_t^*(s), s, t + \bar{T}), \quad s \in (t, t + \bar{T}]. \quad (4.145)$$

The new solution concept in the game $\Gamma(x_0, t_0, T)$ with continuous updating consists of a combination of the solutions $\widetilde{M}_t(\tilde{x}_t^*(t), t, t + \bar{T})$ (corresponding to the IDP) in the games $\Gamma(x, t, t + \bar{T})$, $t \in [t_0, T]$.

Definition 4.3.3 *An IDP with continuous updating $\beta(s, x^*)$ is defined for each set $\tilde{\xi}_t(\tilde{x}_t^*(s), s, t + \bar{T}) \in \widetilde{M}_t(\tilde{x}_t^*(s), s, t + \bar{T})$, $t \in [t_0, T]$ using the corresponding $\tilde{\beta}_t(s, \tilde{x}_t^*)$ as follows:*

$$\beta(t, x^*) = \tilde{\beta}_t(t, \tilde{x}_t^*(t)), \quad t \in [t_0, T]. \quad (4.146)$$

Note that the trajectory with continuous updating $x^*(t)$ defined in (4.140) at the instant t coincides with $\tilde{x}_t^*(t)$. This follows from the definition of $\tilde{x}_t^*(s)$ given in Section 4.3.1.2. Using the IDP with continuous updating $\beta(t, x^*)$, we define the following vector.

Definition 4.3.4 *An imputation with continuous updating $\xi(x^*(t), t, T)$ is the vector defined using the IDP $\beta(t, x^*)$ in the following way:*

$$\xi(x^*(t), t, T) = \int_t^T \beta(\tau, x^*(\tau)) d\tau, \quad t \in [t_0, T]. \quad (4.147)$$

Using the imputation with continuous updating, we introduce the cooperative solution with continuous updating as follows.

Definition 4.3.5 *The cooperative solution with continuous updating $W(x^*(t), t, T)$ is the set of imputations $\xi(x^*(t), t, T)$, (4.147) for all possible IDPs $\beta(t, x^*)$ (4.146).*

The approach presented in this section is a natural way to model the cooperative solution in the game model with continuous updating. What is important, at each current time t the players have only the information about the game defined on the interval $[t, t + \bar{T}]$; based on this information, they can construct the cooperative solution at each current time t (4.144) and later compose the resulting imputation (4.147) using the IDP with continuous updating (4.146) as the instant imputation for each current time t . Some properties of the constructed solution have been discussed in Sections 4.3.1.4 and 4.3.1.5.

4.3.1.4 Time-consistency Property of Cooperative Solution in Game with Continuous Updating

The cooperative solution with continuous updating $W(x_0, t_0, T)$ is time-consistent by construction. But it is possible to prove that this solution also possesses the property of strong time-consistency [159].

Theorem 4.3.1 *An arbitrary cooperative solution with continuous updating $W(x_0, t_0, T)$ is strong time-consistent in the game $\Gamma(x_0, t_0, T)$ with continuous updating.*

An arbitrary cooperative solution is any cooperative solution that can be formed by choosing different $\tilde{\xi}_t(\tilde{x}_t^*(s), s, t + \bar{T}) \in \tilde{M}_t(\tilde{x}_t^*(s), s, t + \bar{T})$ in each game defined on the interval $[t, t + \bar{T}]$, i.e., at each time instant $t \in [t_0, T]$.

Proof According to the definition of strong time-consistency, we need to prove that for every $\xi(x_0, t_0, T) \in W(x_0, t_0, T)$, there exists an IDP $\beta(t, x^*)$ such that for any $\xi'(x^*(t), t, T) \in W(x^*(t), t, T)$, the following condition is satisfied:

$$\int_{t_0}^t \beta(\tau, x^*(\tau)) d\tau + \xi'(x^*(t), t, T) \in W(x_0, t_0, T). \quad (4.148)$$

For a fixed imputation $\bar{\xi}(x_0, t_0, T) \in W(x_0, t_0, T)$, we need to prove that there exists an IDP $\bar{\beta}(t, x^*)$ such that for any $\xi'(x^*(t), t, T) \in W(x^*(t), t, T)$,

$$\int_{t_0}^t \bar{\beta}(\tau, x^*(\tau)) d\tau + \xi'(x^*(t), t, T) \in W(x_0, t_0, T). \quad (4.149)$$

Let us choose $\bar{\beta}$ by the following rule (the continuous updating procedure):

$$\bar{\beta}(\tau, x^*(\tau)) = \tilde{\beta}_\tau(\tau, \tilde{x}_\tau^*(\tau)), \quad (4.150)$$

where $\tilde{\beta}_\tau(\tau, \tilde{x}_\tau^*(\tau))$ is chosen according to the time-consistency property of the imputation $\tilde{\xi}_\tau(x_\tau^*(s), s, \tau + \bar{T})$ using (4.145). Here $\xi'(x^*(t), t, T)$ is the imputation with continuous updating in the subgame starting at the time instant t at $x^*(t)$. According to Definition 4.3.4, $\xi'(x^*(t), t, T)$ is a vector defined using the IDP:

$$\xi'(x^*(t), t, T) = \int_t^T \beta'(\tau, x^*(\tau)) d\tau, \quad (4.151)$$

where $\beta'(\tau, x^*(\tau))$ is constructed by the procedure (4.146). This means that $\beta'(\tau, x^*(\tau)) = \tilde{\beta}'_\tau(\tau, \tilde{x}_\tau^*(\tau))$, where $\tilde{\beta}'_\tau(s, \tilde{x}_\tau^*(s))$ is an IDP of some $\tilde{\xi}'_t(\tilde{x}_t^*(s), s, t + \bar{T}) \in$

$\widetilde{M}_t(\widetilde{x}_t^*(s), s, t + \overline{T})$. Taking into account (4.151), we write formula (4.149) as

$$\int_{t_0}^t \bar{\beta}(\tau, x^*(\tau)) d\tau + \xi'(x^*(t), t, T) = \int_{t_0}^t \bar{\beta}(\tau, x^*(\tau)) d\tau + \int_t^T \beta'(\tau, x^*(\tau)) d\tau. \quad (4.152)$$

Consider now the following IDP:

$$\widehat{\beta}(\tau, x^*(\tau)) = \begin{cases} \bar{\beta}(\tau, x^*(\tau)), & \tau \in [t_0, t), \\ \beta'(\tau, x^*(\tau)), & \tau \in [t, T), \end{cases} \quad (4.153)$$

Then, we obtain

$$\int_{t_0}^t \bar{\beta}(\tau, x^*(\tau)) d\tau + \xi'(x^*(t), t, T) = \int_{t_0}^T \widehat{\beta}(\tau, x^*(\tau)) d\tau. \quad (4.154)$$

But $\widehat{\beta}(\tau, x^*(\tau))$ is constructed according to the continuous updating approach formula (4.146). Therefore, $\int_{t_0}^t \bar{\beta}(\tau, x^*(\tau)) d\tau + \xi'(x^*(t), t, T)$ is an imputation with continuous updating, and hence $\int_{t_0}^t \bar{\beta}(\tau, x^*(\tau)) d\tau + \xi'(x^*(t), t, T)$ belongs to $W(x_0, t_0, T)$. \square

4.3.1.5 Characteristic Function in Game Model with Continuous Updating

Suppose that the function $\widetilde{V}_t(S; \widetilde{x}_t^*(s), s, t + \overline{T})$, $S \subseteq N$, is continuously differentiable with respect to $s \in [t, t + \overline{T}]$ and integrable with respect to $t \in [t_0, +\infty)$. We define the characteristic function in the game model with continuous updating $V(S; x^*(t), t)$ in the following way:

Definition 4.3.6 *A function $V(S; x^*(t), t)$, $t \in [t_0, +\infty)$, $S \subseteq N$, is the characteristic function with continuous updating if*

$$V(S; x^*(t), t) = \int_t^{+\infty} -\frac{d}{ds} \widetilde{V}_\tau(S; \widetilde{x}_\tau^*(s), s, \tau + \overline{T})|_{s=\tau} d\tau, \quad t \in [t_0, +\infty), \quad S \subseteq N, \quad (4.155)$$

where $\widetilde{V}_\tau(S; \widetilde{x}_\tau^*(s), s, \tau + \overline{T})$, $s \in [\tau, \tau + \overline{T}]$, $\tau \in [t, +\infty)$, $S \subseteq N$, is the characteristic function in the game $\Gamma(\widetilde{x}_\tau^*(s), s, \tau + \overline{T})$ defined on the interval $[s, \tau + \overline{T}]$.

Note that the integral in (4.155) can be infinite. Therefore, it is necessary to define the cooperative game model for a finite time interval, i.e., $t \in [t_0, T]$:

$$V(S; x^*(t), t, T) = \int_t^T -\frac{d}{ds} \tilde{V}_\tau(S; \tilde{x}_\tau^*(s), s, \tau + \bar{T})|_{s=\tau} d\tau, \quad t \in [t_0, T], \quad S \subseteq N. \quad (4.156)$$

The imputation with continuous updating $\xi(x_0, t_0, T)$ (4.147) constructed above can be treated as an imputation in the game model with continuous updating. To establish this fact, we show that it is an imputation for some specific characteristic function. The characteristic function with continuous updating $V(S; x_0, t_0, T)$ (4.156) will be used for this purpose. This approach will justify using the imputation with continuous updating and the characteristic function with continuous updating simultaneously.

Theorem 4.3.2 *An imputation with continuous updating $\xi(x_0, t_0, T)$ is an imputation in the game $\Gamma(x_0, t_0, T)$ with continuous updating and the characteristic function defined as $V(S; x^*(t), t, T)$ (4.156) if the following condition is satisfied:*

$$\frac{d}{ds} \tilde{\xi}_i^t(x_t^*(s), s, t + \bar{T})|_{s=t} \leq \frac{d}{ds} \tilde{V}_t(\{i\}; x_t^*(s), s, t + \bar{T})|_{s=t}. \quad (4.157)$$

Proof First, let us show that $\forall t \in [t_0, T]$ the following conditions are satisfied:

$$\sum_{i=1}^n \xi_i(x^*(t), t, T) = V(N; x^*(t), t, T), \quad (4.158)$$

$$\xi_i(x^*(t), t, T) \geq V(\{i\}; x^*(t), t, T), \quad i \in N. \quad (4.159)$$

According to the definition of $\xi(x^*(t), t, T)$ and $V(S; x^*(t), t, T)$, the left-hand side of (4.158) can be written as

$$\begin{aligned} \sum_{i=1}^n \xi_i(x^*(t), t, T) &= \sum_{i=1}^n \int_t^T \beta_i(s, x^*(s)) ds = \\ &= \sum_{i=1}^n \int_t^T -\frac{d}{dl} \tilde{\xi}_i^s(x_s^*(l), l, s + \bar{T})|_{l=s} ds. \end{aligned} \quad (4.160)$$

According to the definition of the characteristic function, the right-hand side of

(4.158) can be written as

$$V(N; x^*(t), t, T) = \int_t^T -\frac{d}{dl} \tilde{V}_s(N; x_s^*(l), l, s + \bar{T}) \Big|_{l=s} ds. \quad (4.161)$$

Due to

$$\sum_{i=1}^n \tilde{\xi}_i^s(x_s^*(l), l, s + \bar{T}) = \tilde{V}_s(N; x_s^*(l), l, s + \bar{T}), \quad (4.162)$$

it follows that the right-hand sides of (4.160) and (4.161) are equivalent, and therefore condition (4.158) holds.

Now let us derive formula (4.159) by substituting $\xi(x^*(t), t, T)$ and $V(\{i\}; x^*(t), t, T)$ into the left-hand side of (4.159):

$$\xi_i(x^*(t), t, T) = \int_t^T \beta_i(s, x^*(s)) ds = \int_t^T -\frac{d}{dl} \tilde{\xi}_i^s(x_s^*(l), l, s + \bar{T}) \Big|_{l=s} ds. \quad (4.163)$$

Taking into account formula (4.163) for the imputation, formula (4.161) for the characteristic function in the case $N = \{i\}$, and condition (4.157), we arrive in formula (4.159). The proof is complete. \square

In this section, the notion of the characteristic function with continuous updating $V(S; x_0, t_0, T)$ has been introduced. It has been shown that the imputation with continuous updating $\xi(x_0, t_0, T)$ is indeed an imputation in the classical sense when the additional conditions (4.157) are satisfied. Note that condition (4.157) does not always hold. We can hardly say in general when condition (4.157) is satisfied: it depends on both the type of imputation (Shapley value, proportional imputation, etc.) and characteristic function ($\alpha, \beta, \gamma, \sigma$, etc.). The main idea of this result is to show that in the game model with continuous updating, it is possible (under conditions (4.157)) to use the characteristic function with continuous updating (4.156) for constructing the imputation with continuous updating without directly using the approach presented in Section 4.3.1.3.4.

4.3.1.6 Resulting Solution as Function of Resulting Characteristic Function

This section shows that if players choose the imputation $\tilde{\xi}_t(\tilde{x}_t^*(s), s, t + \bar{T}) \in \tilde{M}_t(\tilde{x}_t^*(s), s, t + \bar{T})$ based on $\tilde{V}_t(S; \tilde{x}_t^*(s), s, t + \bar{T})$, $s \in [t, t + \bar{T}]$, at every time instant $t \in [t_0, T]$ by the same rule, then the corresponding imputation with continuous updating $\xi(x^*(t), t, T)$ corresponds to the imputation chosen by the same rule using the characteristic function with continuous updating $V(S; x^*(t), t, T)$. Further we prove this for several optimality principles.

Let us prove it for all types $\tilde{\xi}_t(\tilde{x}_t^*(s), s, t + \bar{T})$ that are finitely additive functions of $\tilde{V}_t(S; \tilde{x}_t^*(s), s, t + \bar{T})$, $S \subseteq N$, $t \in [t_0, T]$. Indeed, most classical imputations are finitely additive functions of the values of the characteristic functions. In particular, the Shapley value is an additive function of the characteristic function $V(S)$, $S \subseteq N$, (see the first axiom of linearity):

$$Sh_i(V) = \sum_{S \subseteq N, i \in S} \frac{(s-1)!(n-s)!}{n!} [V(S) - V(S \setminus i)], \quad i \in N. \quad (4.164)$$

We denote by $\xi(\tilde{V}_t; \tilde{x}_t^*(s), s, T) = \tilde{\xi}_t(\tilde{x}_t^*(s), s, t + \bar{T})$ the imputation calculated using the values of the characteristic function $\tilde{V}_t(S; \tilde{x}_t^*(s), s, t + \bar{T})$, $S \subseteq N$, and by $\xi(V; x^*(t), t, T)$ the same imputation type (Shapley value, etc.) calculated using the values of the characteristic function with continuous updating $V(S; x^*(t), t, T)$, $S \subseteq N$.

The theorem presented below for linear additive imputations (e.g., the Shapley value as a special case) shows that to calculate the cooperative solution in the game with continuous updating, there is no need to calculate the imputation for each current time instant $t \in [t_0, T]$. It is enough to calculate the values of the characteristic function with continuous updating $V(S; x^*(t), t, T)$, $S \subseteq N$, and use them for calculating the corresponding cooperative solution as a characteristic function.

Theorem 4.3.3 *Suppose that in every game $\Gamma(x^*(t), t, t + \bar{T})$, as an imputation any finitely additive function $\xi(\tilde{V}_t; \tilde{x}_t^*(s), s, t + \bar{T}) = \tilde{\xi}_t(\tilde{x}_t^*(s), s, t + \bar{T})$ is used. Then the corresponding imputation $\xi(x^*(t), t, T)$ (4.147) coincides with the same*

imputation $\xi(V; x^*(t), t, T)$ calculated using the values of the characteristic function with continuous updating.

Proof We need to prove that

$$\xi(x^*(t), t, T) = \xi(V; x^*(t), t, T). \quad (4.165)$$

We write the left-hand side of formula (4.165) as

$$\xi(x^*(t), t, T) = \int_t^T \beta(x^*(s), s) ds = \int_t^T -\frac{d}{dl} \tilde{\xi}_s(\tilde{x}_s^*(l), l, s + \bar{T}) |_{l=s} ds \quad (4.166)$$

Next, we write the formula for V in the right-hand side of (4.165) as

$$\begin{aligned} V(S; x^*(t), t, T) &= \int_t^T -\frac{d}{dl} \tilde{V}_s(S; \tilde{x}_s^*(l), l, s + \bar{T}) |_{l=s} ds = \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n \left(-\frac{d}{dl} \tilde{V}_{y_i}(S; \tilde{x}_{y_i}^*(l), l, y_i + \bar{T}) |_{l=y_i} \right) (s_{i+1} - s_i), \end{aligned} \quad (4.167)$$

where $t = s_0 < s_1 < \dots < s_i < s_{i+1} < \dots < s_n = T$ and $y_i \in (s_i, s_{i+1})$.

Taking into account the property of an additive function, we obtain

$$\begin{aligned} \xi(V; x^*(t), t, T) &= \\ &= \xi \left(\lim_{n \rightarrow \infty} \sum_{i=1}^n \left(-\frac{d}{dl} \tilde{V}_{y_i}(S; \tilde{x}_{y_i}^*(l), l, y_i + \bar{T}) |_{l=y_i} \right) (s_{i+1} - s_i); x^*(t), t, T \right) = \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n \left(-\frac{d}{dl} \xi(\tilde{V}_{y_i}; \tilde{x}_{y_i}^*(l), l, y_i + \bar{T}; x^*(t), t, T) |_{l=y_i} \right) (s_{i+1} - s_i) = \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n \left(-\frac{d}{dl} \tilde{\xi}_{y_i}(\tilde{x}_{y_i}^*(l), l, y_i + \bar{T}) |_{l=y_i} \right) (s_{i+1} - s_i). \end{aligned} \quad (4.168)$$

The proof is complete. □

4.3.2 Hamilton-Jacobi-Bellman Equations for Cooperative Game Model with Continuous Updating

This section will only consider the feedback-based cooperative strategies with continuous updating $u_{fb}^*(t, x)$. Therefore, we will use the simplified notation $u^*(t, x)$.

To construct the Hamilton-Jacobi-Bellman equation for the cooperative strategies with continuous updating, we need to determine the generalized cooperative strategies in feedback form $\tilde{u}^*(t, x; s, x_t)$ in the game with continuous updating. Within the continuous updating approach, the Bellman function $W(t; s, x)$ is defined as the cooperative payoff of all players in the feedback cooperative strategies $\tilde{u}^*(t, x; s, x_t)$ in the subgame starting at the instant s from the state x in the game defined on the interval $[t, t + \bar{T}]$:

$$W(t; s, x) = \sum_{i \in N} \int_s^{t+\bar{T}} g^i[\tau, x_t(\tau), \tilde{u}^*(t, x; s, x_t(\tau))] d\tau \quad (4.169)$$

subject to

$$\begin{aligned} \dot{x}_t(\tau) &= f(\tau, x_t, \tilde{u}^*), \\ x_t(s) &= x. \end{aligned} \quad (4.170)$$

The following theorem holds.

Theorem 4.3.4 $\tilde{u}^*(t, x; s, x_t)$ forms the generalized cooperative strategies in feedback form in the differential game with continuous updating $\Gamma(x_0, t_0, \bar{T})$ if there exists a function $W(t; s, x) : [t_0, +\infty) \times [t, t + \bar{T}] \times R \rightarrow R$, $i \in N$, continuously differentiable with respect to s and x , satisfying the system of partial differential equations

$$\begin{aligned} -W_s(t; s, x) &= \max_{\phi_i} \left\{ \sum_{i \in N} g^i(s, x, u) + W_x(t; s, x) f(s, x, u) \right\} = \\ &= \sum_{i \in N} g^i(s, x, \tilde{u}^*) + W_x(t; s, x) f(s, x, \tilde{u}^*), \end{aligned} \quad (4.171)$$

$$W(t; t + \bar{T}, x) = 0.$$

Proof By the definition of the generalized cooperative strategies, $\tilde{u}^*(t, x; s, x_t)$ should be the cooperative strategies for any fixed t .

Fixing t in the statement of Theorem 4.3.4 and, in particular, in (4.171), we obtain the classical sufficient conditions for feedback Nash equilibrium in the differential game with prescribed duration $[t, t + \bar{T}]$ presented in [80]. Therefore, for any fixed t , the conditions for the definition of generalized cooperative strategies are satisfied. The theorem is proved. \square

4.3.2.1 Model of Non-renewable Resource Extraction

As an illustrative example, consider a differential game model with continuous updating for extracting a nonrenewable resource (see [50]). Assume that n players exploit a natural common-property resource, which does not regenerate over time, such as natural gas or earth minerals.

4.3.2.1.1 Original Game Model

We denote by $x(t)$ the state vector indicating the resource stock at a time instant t available for extraction by the players. Let $u_i(t)$ be the extraction rate of player i . Assume that $u_i(t) \geq 0$, and if $x(t) = 0$, then the only feasible rate of extraction is $u_i(t) = 0$.

The stock dynamics are given by the following equation with initial condition:

$$\dot{x}(t) = - \sum_{i=1}^n a_i u_i(t, x), \quad x(t_0) = x_0, \quad (4.172)$$

where $a_i > 0$ for all $i = 1, \dots, n$, and $x_0 > 0$. Let the players have the logarithmic utility functions

$$h_i(x(t), u_i(t)) = \ln u_i(t, x), \quad i = 1, \dots, n. \quad (4.173)$$

The payoff function of player i is

$$K_i(x_0, T - t_0) = \int_{t_0}^T \ln u_i(s, x) ds, \quad i = 1, \dots, n. \quad (4.174)$$

Assume that information about the motion equations and payoff functions is updated continuously in time. At every instant $t \in [0, T]$, the players have information only on the interval $[t, t + \bar{T}]$. This means that at each time instant, they can count for the stability of this process only over the period \bar{T} .

Suppose that the game is played in the cooperative scenario: the players have the opportunity to cooperate for achieving the maximum total payoff:

$$\max_{u_1, \dots, u_n} \sum_{i=1}^n K_i(x_0, T - t_0) = \max_{u_1, \dots, u_n} \sum_{i=1}^n \int_{t_0}^T \ln u_i(s, x) ds. \quad (4.175)$$

4.3.2.2 Optimal Cooperative Strategies with Continuous Updating

According to Section 4.3, to determine the optimal feedback cooperative strategies in the game with continuous updating, we consider the family of auxiliary subgames $\Gamma(x, t, t + \bar{T})$ with duration \bar{T} , starting at the instant t from the state x .

The dynamic constraints in the game $\Gamma(x, t, t + \bar{T})$ are given by

$$\begin{aligned} \dot{x}_t(s) &= - \sum_{i=1}^n a_i u_i^t(s, x_t), \\ x_t(t) &= x, \\ x_t, x &\in \mathbb{R}^l, \quad u^t = (u_1^t, \dots, u_n^t), \quad u_i^t = u_i^t(s, x_t) \in U_i \subset \text{comp} \mathbb{R}^k, \quad s \in [t, t + \bar{T}]. \end{aligned} \quad (4.176)$$

The payoff function of player i in $\Gamma(x, t, t + \bar{T})$ is

$$K_i^t(x, t; u^t) = \int_t^{t+\bar{T}} \ln u_i^t(s, x_t) ds, \quad i \in N. \quad (4.177)$$

To define the feedback cooperative strategies $\tilde{u}^*(t, x; s, x_t)$ in the auxiliary subgame

$\Gamma(x, t, t + \bar{T})$, we use the dynamic programming technique. We denote by $W(t; s, x)$ the Bellman function in a the subgame starting at an instant s :

$$W(t; s, x) = \max_{u_1^t, \dots, u_n^t} \sum_{i=1}^n K_i^t(x, s; u^t). \quad (4.178)$$

The Hamilton-Jacobi-Bellman (HJB) equation has the following form:

$$-\frac{\partial W(t; s, x)}{\partial s} = \max_{u_1^t, \dots, u_n^t} \left\{ \sum_{i=1}^n \ln u_i^t(s, x) - \frac{\partial W(t; s, x)}{\partial x} \sum_{i=1}^n a_i u_i^t(s, x) \right\}. \quad (4.179)$$

$$\lim_{s \rightarrow t + \bar{T}} W(t; s, x) = 0.$$

The solution of (4.179) will be found in the form $W(t; s, x) = A(t, s) \ln x + B(t, s)$. The partial derivatives are given by

$$\frac{\partial W(t; s, x)}{\partial s} = \dot{A}(t, s) \ln x + \dot{B}(t, s); \quad \frac{\partial W(t; s, x)}{\partial x} = \frac{A(t, s)}{x}. \quad (4.180)$$

Maximizing the expression in the right-hand side of (4.179) and substituting it into (4.180), we get

$$\tilde{u}^*(t, x; s, x_t) = \frac{x}{a_i A(t, s)}. \quad (4.181)$$

Substituting (4.180) and (4.181) into (4.179), we obtain the following system of differential equations:

$$\begin{aligned} \dot{A}(t, s) \ln x + \dot{B}(t, s) &= n \ln A(t, s) - n \ln x + \ln a^N + n, \\ \lim_{s \rightarrow t + \bar{T}} A(t, s) \ln x + B(t, s) &= 0, \end{aligned} \quad (4.182)$$

where $a^N = \prod_{i=1}^n a_i$. The solution of (4.182) is

$$\begin{aligned} A(s, t) &= n(t + \bar{T} - s), \\ B(s, t) &= -(t + \bar{T} - s)(\ln a^N + n \ln n(t + \bar{T} - s)), \\ t &\in [t_0, T), \quad s \in (t, t + \bar{T}). \end{aligned} \quad (4.183)$$

Finally we get the feedback optimal cooperative strategies in the auxiliary subgame

$\Gamma(x, t, t + \bar{T})$:

$$u_i^*(t, x; s, x_t) = \frac{1}{W_x} = \frac{x}{a_i n(t + \bar{T} - s)}, \quad s \in [t, t + \bar{T}]. \quad (4.184)$$

and

$$W(t; s, x) = n(t + \bar{T} - s) \ln \frac{x}{n(t + \bar{T} - s)} - (t + \bar{T} - s) \ln a^N, \quad s \in [t, t + \bar{T}]. \quad (4.185)$$

The optimal cooperative trajectory in the auxiliary subgame $\Gamma(x, t, t + \bar{T})$ has the form

$$\tilde{x}_t^*(s) = x \frac{(t + \bar{T} - s)}{\bar{T}}. \quad (4.186)$$

Following the procedure (4.139), we construct the feedback optimal cooperative strategies with continuous updating

$$u^*(t, x) = \tilde{u}^*(t, x; s, x_t)|_{s=t} = \frac{x}{a_i n \bar{T}}, \quad s \in [t, t + \bar{T}] \quad (4.187)$$

The optimal cooperative trajectory $x^*(t)$ with continuous updating is

$$x^*(t) = x_0 e^{\frac{t_0 - t}{\bar{T}}}.$$

The results are illustrated in Figures 4.17, 4.18.

4.3.2.2.1 Characteristic Function on Interval $[t, t + \bar{T}]$

We will construct the characteristic function in the subgames $\Gamma(x, t, t + \bar{T})$ using the approach proposed in [228]:

$$\tilde{V}_t(S; x, s, t + \bar{T}) = \begin{cases} \sum_{i=1}^n K_i^t(x, s; \tilde{u}^*(t, x; s, x_t)), & S = N, \\ \max_{u_i, i \in S} \sum_{i \in S} K_i^t(x, s; \tilde{u}^{NE \setminus u^S}), & S \subset N, \\ 0, & S = \emptyset, \end{cases} \quad (4.188)$$

where $\max_{u_i, i \in S} \sum_{i \in S} K_i^t(x, s; \tilde{u}^{NE} / \tilde{u}^S)$ is the maximum joint payoff of coalition S when the players from coalition $N \setminus S$ use the strategies from some fixed Nash equilibrium

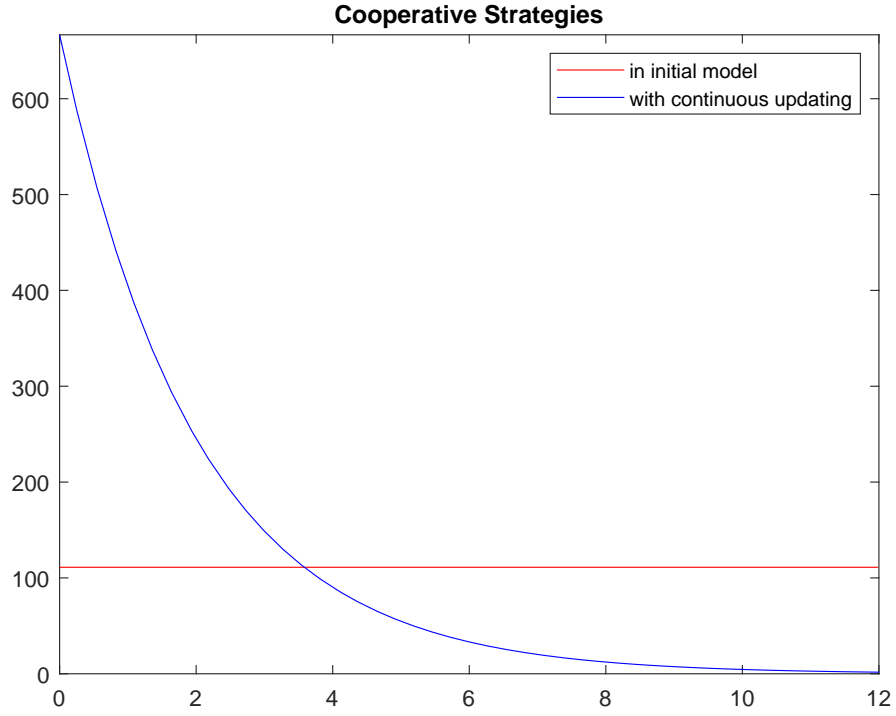


Figure 4.17: Cooperative strategies in original game (red line) and cooperative strategies with continuous updating $u^*(t, x)$ (blue line).

$\tilde{u}^{NE} = (\tilde{u}_1^{t,NE}, \dots, \tilde{u}_n^{t,NE})$ in the subgame $\Gamma(x, t, t + \bar{T})$.

In general, such a characteristic function is non-superadditive. According to (4.185), the payoff of the maximum coalition is

$$\tilde{V}_t(N; \tilde{x}_t^*(s), s, t + \bar{T}) = n(t + \bar{T} - s) \ln \frac{\tilde{x}_t^*(s)}{n(t + \bar{T} - s)} - (t + \bar{T} - s) \ln a^N.$$

The characteristic function $\tilde{V}_t(S; x, s, t + \bar{T})$ of coalition S is constructed in two steps. First, we calculate the Nash equilibrium strategies for all players. Second, we fix the strategies for the players from $N \setminus S$, assuming that the players from S seek to maximize their joint payoff.

To calculate the feedback Nash equilibrium strategies in the subgame $\Gamma(x, t, t + \bar{T})$, we use the dynamic programming technique. We denote by $W^i(t; s, x)$ the Bellman function as the payoff of player i in the Nash equilibrium situation in the subgame of $\Gamma(x, t, t + \bar{T})$ starting at the instant s .

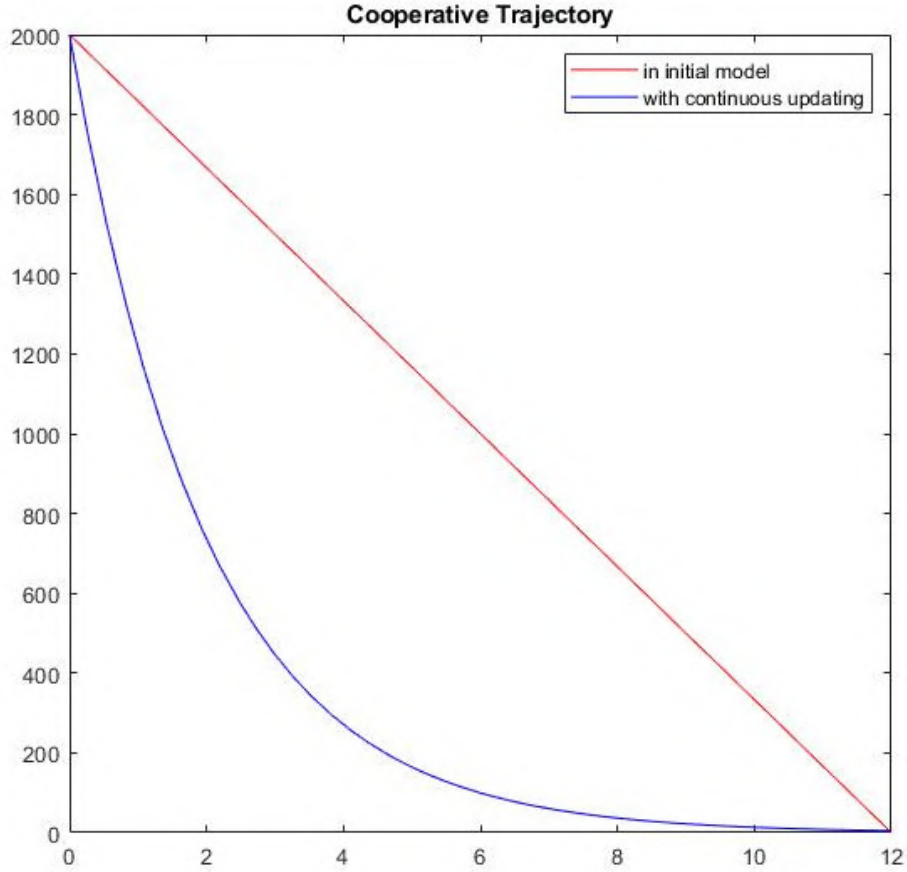


Figure 4.18: Cooperative trajectory in original game (red line) and cooperative trajectory with continuous updating $x^*(t)$ (blue line).

The system of Hamilton-Jacobi-Bellman (HJB) equations has the following form:

$$-\frac{\partial W^i(t; s, x)}{\partial s} = \max_{u_i} (\ln u_i^t(s, x) - \frac{\partial W^i(t; s, x)}{\partial x} (a_i u_i + \sum_{j \neq i} a_j \tilde{u}_j^{NE})), \quad i \neq j, \quad (4.189)$$

$$\lim_{s \rightarrow t+\bar{T}} W^i(t; s, x) = 0, \quad i = 1, \dots, n.$$

The Bellman functions $W_i(t; s, x)$ will be found in the form $W_i(t; s, x) = A_i(t, s)x + B_i(t, s)$. Maximizing the expression in the right-hand side of (4.189) and solving the corresponding differential equations in $A_i(t, s)$ and $B_i(t, s)$, we obtain:

$$A_i(t, s) = \bar{T} + t - s,$$

$$B_i(t, s) = (\bar{T} + t - s)(-\ln(\bar{T} + t - s) - \ln a_i - n - 1),$$

$$\tilde{u}_i^{NE}(t, x; s, x_t) = \frac{x}{a_i(\bar{T} + t - s)}.$$

Note that the characteristic function of coalition i coincides with $W^i(t; s, x)$:

$$V_t(\{i\}; \tilde{x}_t^*(s), s, t + \bar{T}) = W^i(t; s, x) = (\bar{T} + t - s) \left(\ln \frac{\tilde{x}_t^*(s)}{\bar{T} + t - s} - \ln a_i - n - 1 \right).$$

Now let us construct $\tilde{V}_t^S(S; x, s, \bar{T})$. We denote by $W^S(t; s, x)$ the Bellman function as the maximum payoff of coalition S in the subgame starting at the time instant s when the players from coalition $N \setminus S$ use the strategies $\tilde{u}_i^{NE}(t, x; s, x_t)$. We find the characteristic function in the form $W^S(t, s, x) = A^S(t, s) \ln x + B^S(t, s)$ as the solution of the Hamilton-Jacobi-Bellman equation

$$\begin{aligned} -\frac{\partial W^S(t; s, x)}{\partial s} &= \max_{u_i \in S} \left(\sum_{i \in S} \ln u_i^t(s, x) - \frac{\partial W^S(t; s, x)}{\partial x} \left(\sum_{i \in S} a_i u_i + \sum_{j \in N \setminus S} a_j \tilde{u}_j^{NE} \right) \right), \\ S &\subset N, \\ \lim_{s \rightarrow t + \bar{T}} W^S(t; s, x) &= 0. \end{aligned} \tag{4.190}$$

Equations (4.191) provide the strategies for players $i \in S$ maximizing the expression in the right-hand side of (4.190):

$$\tilde{u}_i^S(t, x; s, x_t) = \frac{x}{a_i A^S(t, s)}. \tag{4.191}$$

Let $|S| = k$. Substituting (4.191) into (4.190), we get the following system of differential equations:

$$\begin{aligned} \dot{A}^S(t, s) &= -k, \\ \dot{B}^S(t, s) &= k \ln A^S(t, s) + \ln a^S + k + \frac{A^S(t, s)(n-k)}{\bar{T} + t - s}, \\ \lim_{s \rightarrow t + \bar{T}} (A^S(t, s) \ln x + B^S(t, s)) &= 0, \end{aligned} \tag{4.192}$$

where $a^S = \prod_{i \in S} a_i$. The solution of (4.192) is

$$\begin{aligned} A^S(s, t) &= k(t + \bar{T} - s), \\ B^S(s, t) &= -k(t + \bar{T} - s)(\ln k(t + \bar{T} - s) + \frac{\ln a^S}{k} + n - k), \\ t &\in [t_0, T), \quad s \in (t, t + \bar{T}). \end{aligned} \tag{4.193}$$

The optimal strategies of players from coalition S (who maximize their joint payoff)

have the following form:

$$u_i^S(t, x; s, x_t) = \frac{x}{a_i k(t + \bar{T} - s)}, \quad s \in [t, t + \bar{T}). \quad (4.194)$$

The value of the characteristic function of coalition S is

$$\begin{aligned} \tilde{V}_t(S; \tilde{x}_t^*(s), s, t + \bar{T}) &= W^S(t; s, x) = -k(t + \bar{T} - s) \\ &(-\ln \tilde{x}_t^*(s) + \ln k(t + \bar{T} - s) + \frac{\ln a^S}{k} + n - k), \quad s \in [t, t + \bar{T}). \end{aligned} \quad (4.195)$$

Now, to demonstrate that the superadditivity property for the constructed characteristic function holds, let $S \subset N$, $P \subset N$, and $S \cap P = \emptyset$. Then

$$\begin{aligned} \tilde{V}_t(S \cup P; x, s, t + \bar{T}) - \tilde{V}_t(S; x, s, t + \bar{T}) - \tilde{V}_t(P; x, s, t + \bar{T}) &= \\ &= (t + \bar{T} - s) (k \ln k + p \ln p - k \ln(k + p) - p \ln(k + p) + 2pk) = \\ &= (t + \bar{T} - s) \left(k \left(\ln \frac{k}{k+p} + p \right) + p \left(\ln \frac{p}{k+p} + k \right) \right). \end{aligned} \quad (4.196)$$

Note that for all $k > 1$ and $p > 1$, the following inequalities hold: $\frac{k+p}{k} < e^p$, $\frac{k+p}{p} < e^k$. Thus, $(\ln \frac{k}{k+p} + p) > 0$ and $(\ln \frac{p}{k+p} + k) > 0$. Therefore, we argue that

$$\tilde{V}_t(S \cup P; x, s, t + \bar{T}) - \tilde{V}_t(S; x, s, t + \bar{T}) - \tilde{V}_t(P; x, s, t + \bar{T}) \geq 0, \quad (4.197)$$

for all $s \in [t, t + \bar{T})$, and the superadditivity of the characteristic function is proved.

4.3.2.2.2 Shapley Value with Continuous Updating

To find the cooperative solution, we can apply any of the known principles of optimality using the superadditive characteristic function. Let us construct the

Shapley value [229] for the game:

$$\begin{aligned}
\widetilde{sh}_i^t(\widetilde{x}_t^*(s), \bar{T} + t - s) &= \sum_{K \subseteq N, i \in K} \frac{(k-1)!(n-k)!}{n!} [(\widetilde{V}_t(K; \widetilde{x}_\tau^*(s), \bar{T} + t - s)) - \\
&\quad - (\widetilde{V}_t(K \setminus i; \widetilde{x}_t^*(s), \bar{T} + t - s))] = \\
&= (t + \bar{T} - s) \sum_{K \subseteq N, i \in K} \frac{(k-1)!(n-k)!}{n!} \left[\ln \frac{x}{t + \bar{T} - s} - \right. \\
&\quad \left. - k \ln k + (k-1) \ln(k-1) - \ln a_i + 2k - n - 1 \right] = \\
&= (t + \bar{T} - s) \left(\ln \frac{\widetilde{x}_t^*(s)}{t + \bar{T} - s} - \ln a_i - \ln n \right). \quad (4.198)
\end{aligned}$$

Now consider the cooperative solution of the differential game with continuous updating. According to (4.147), we construct the Shapley value with continuous updating using the Shapley value (4.198) of the auxiliary subgame:

$$\begin{aligned}
sh_i(x^*(t), T - t) &= \int_t^T - \frac{d\widetilde{sh}_i^\tau(\widetilde{x}_\tau^*(s), \bar{T} + \tau - s)}{ds} \Big|_{s=\tau} d\tau = \\
&= (T - t) \left(\ln \frac{x_0}{a_i n \bar{T}} + \frac{t_0}{\bar{T}} - \frac{t + T}{2\bar{T}} \right). \quad (4.199)
\end{aligned}$$

Note that conditions (4.157) are satisfied for the constructed the Shapley value with continuous updating. Thus, this vector is an imputation in the differential game with continuous updating.

The same result can be obtained via the characteristic function of the differential game with continuous updating constructed by the rule (4.155):

$$\begin{aligned}
V(S, x^*(t), t, T - t) &= \int_t^T - \frac{d\widetilde{V}_\tau(S; \widetilde{x}^*(s), s, \tau + \bar{T})}{ds} \Big|_{s=\tau} d\tau = \\
&= k(T - t) \left(\ln \frac{x_0}{k\bar{T}} + \frac{\ln a^S}{k} + \frac{t_0}{\bar{T}} - \frac{t + T}{2\bar{T}} - n + k \right). \quad (4.200)
\end{aligned}$$

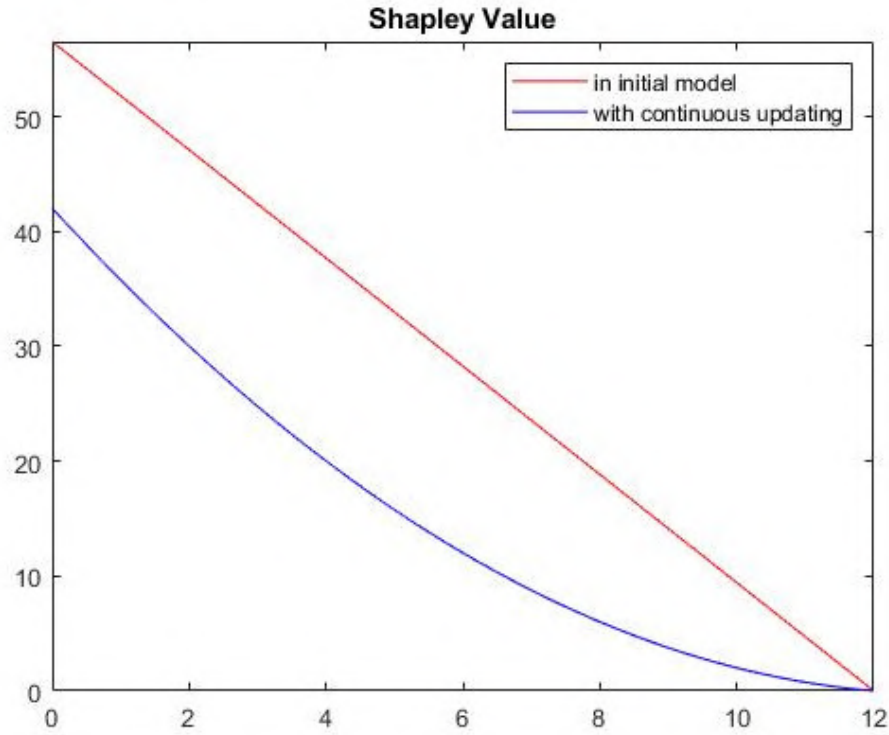


Figure 4.19: Shapley value of player i in original game (red line) and Shapley value with continuous updating $sh_i(x^*(t), T - t)$ (blue line).

4.3.3 Pontryagin's Maximum Principle for Cooperative Differential Game Model with Continuous Updating

4.3.3.1 Cooperative Differential Game with Continuous Updating

This section is devoted to the application of Pontryagin's maximum principle to the cooperative differential game model with continuous updating. In the cooperative setting, before starting the game, all players agree to behave jointly in an optimal way (cooperate). Prior to proceeding to this step, we introduce the type of the characteristic function that will be used in this section. Below, we will consider only the open-loop-based cooperative strategies with continuous updating $u_{o_i}^*(t, x)$. Therefore, we will use the simplified notation $u^*(t, x)$.

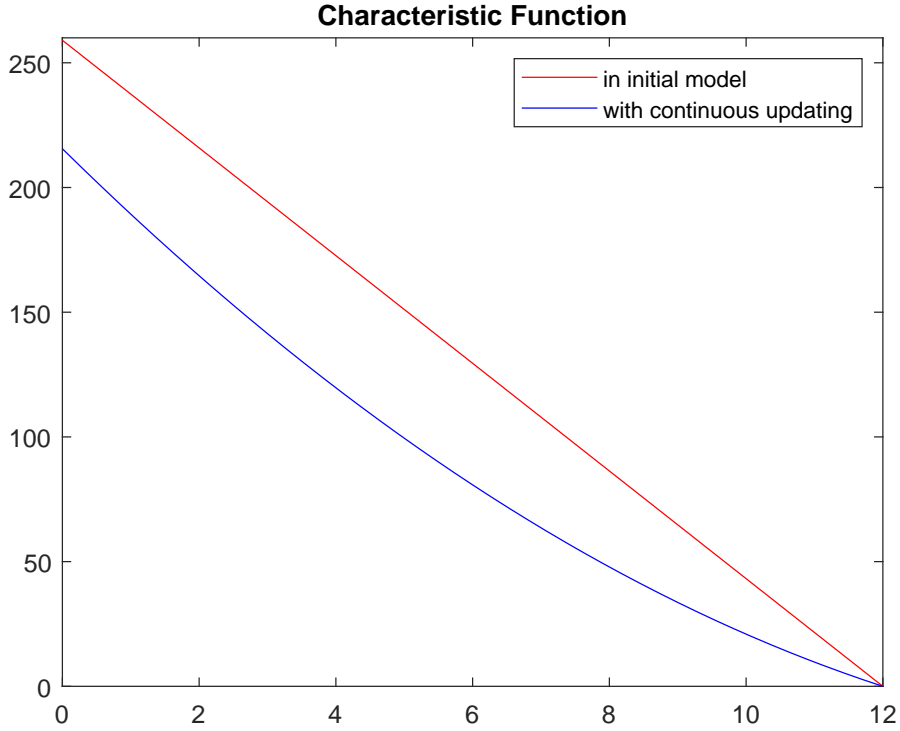


Figure 4.20: Characteristic function value for coalition N in original game (red line) and characteristic function value for coalition N with continuous updating $V(N, x^*(t), t, T - t)$ (blue line).

4.3.3.1.1 Approach to Defining Characteristic Function on Interval $[s, t + \bar{T}]$

We introduce the notion of characteristic function $\tilde{V}_t(S; x, s, t + \bar{T})$, $\forall S \subseteq N$, defined for each subgame $\Gamma(x_{t,s}, s, t + \bar{T})$, where $s \in [t, t + \bar{T}]$ and $t \in [t_0, +\infty)$. Prior to introducing the characteristic function for the subgame $\Gamma(x_{t,s}, s, t + \bar{T})$, note the following: due to equation (4.5), $x_{t,s}$ depends on the initial point x . Therefore, we can write the subgame as $\Gamma(x, s, t + \bar{T})$, and the payoff function of each player i in the subgame becomes $K_i^t(x, s, t + \bar{T}; u)$. For the characteristic function, we will adopt the constructive approach proposed by L. Petrosjan L. and G. Zaccour [230] with the δ -characteristic function. We denote by $V(S; x^*(t), t, T)$ the strength of coalition S in the subgame $\Gamma(x^*(t), t, T)$. It can be found in two stages. First, we calculate the Nash equilibrium strategies $\{u_i^{NE}\}$ for all players $i \in N$. Second, we fix (freeze) these strategies for the players from $N \setminus S$, assuming that the players from coalition S seek to maximize their joint payoff $\sum_{i \in S} K_i$ over $u_S = \{u_i\}_{i \in S}$. Therefore,

the characteristic function has the form

$$\tilde{V}_t(S; x, s, t + \bar{T}) = \begin{cases} \max_{u_1, \dots, u_n} \sum_{i=1}^n K_i^t(x, s, t + \bar{T}; u_1, \dots, u_n), & S = N, \\ \max_{u_i, i \in S} \sum_{i \in S} K_i^t(x, s, t + \bar{T}; u_S, u_{N \setminus S}^{NE}), & S \subset N, \\ 0, & S = \emptyset, \end{cases}$$

where $x = x_t(t)$, $t \in [t_0, +\infty)$, and $s \in [t, t + \bar{T}]$; see (4.5). In addition, $u_S = \{u_i\}_{i \in S}$ is the strategy profile for the players in the coalition S .

Assume that the superadditivity conditions for the characteristic function $\tilde{V}_t(S; x, s, t + \bar{T})$ are satisfied:

$$\begin{aligned} \tilde{V}_t(S_1 \cup S_2; x, s, t + \bar{T}) &\geq \tilde{V}_t(S_1; x, s, t + \bar{T}) + \tilde{V}_t(S_2; x, s, t + \bar{T}), \\ \forall S_1, S_2 \subseteq N, S_1 \cap S_2 &= \emptyset. \end{aligned} \quad (4.201)$$

4.3.3.1.2 Algorithm for Calculating Characteristic Function with Continuous Updating and Shapley Value

The first three steps are to calculate the necessary elements for defining the characteristic function. In the next step, the Shapley value is calculated.

Step 1: Optimizing total payment of grand coalition with continuous updating

We denote by $\Gamma^c(x, t, t + \bar{T})$ the cooperative differential game described above. Its duration is \bar{T} . Suppose that there is no inherent obstacle to cooperation between the players, and their payoffs are transferable. More specifically, assume that before the game actually starts, the players agree to play cooperatively.

Theorem 4.3.5 *Let:*

- (i) $f(s, \cdot, u(t, s))$ be continuously differentiable on \mathbb{R}^l , $\forall s \in [t, t + \bar{T}]$,
- (ii) $g_i(\cdot, \cdot, u(t, s))$ be continuously differentiable on $\mathbb{R} \times \mathbb{R}^l$, $\forall s \in [t, t + \bar{T}]$.

If a set of strategies $\{\tilde{u}_i^(t, s)\}_{i \in N}$ provides generalized open-loop cooperative strategies in the differential game with continuous updating, then for any fixed*

$t \in [t_0, T]$, there exists a costate variable $\psi^t(s)$, $s \in [t, t + \bar{T}]$, satisfying the following relations:

$$(1) \dot{x}_t^*(s) = f(s, x_t^*(s), \tilde{u}^*), \quad x_t^*(t) = x^*(t), \quad \text{for } \forall s \in [t, t + \bar{T}],$$

$$(2) \tilde{u}_i^*(t, s) = \arg \max_{u_i \in U_i, i \in N} H^t(s, x_t^*(s), u(t, s), \psi^t(s)), \quad \text{where } s \in [t, t + \bar{T}], \quad i \in N,$$

$$(3) \dot{\psi}^t(s) = -\frac{\partial}{\partial x_t} H^t(s, x_t(s), \tilde{u}^*(t, s), \psi^t(s)), \quad \text{where } s \in [t, t + \bar{T}],$$

$$\psi^t(t + \bar{T}) = 0.$$

Proof For a fixed $t \geq t_0$, consider the game $\Gamma^c(x, t, t + \bar{T})$. The motion equation has the form

$$\dot{x}_t(s) = f(s, x_t(s), u(t, s)), \quad x_t(t) = x^*(t), \quad s \in [t, t + \bar{T}]. \quad (4.202)$$

The payoff function of the grand coalition is given by

$$\sum_{i=1}^n K_i^t(x^*(t), t, t + \bar{T}; u_1(t, s), \dots, u_n(t, s)) = \sum_{i=1}^n \int_t^{t+\bar{T}} g_i[s, x_t(s), u(t, s)] ds. \quad (4.203)$$

For the optimization problem (4.202)-(4.203), the Hamiltonian has the form

$$H^t(s, x_t(s), u(t, s), \psi^t(s)) = \sum_{i=1}^n g_i[s, x_t(s), u(t, s)] + \psi^t(s) f(s, x_t(s), u(t, s)),$$

$$s \in [t, t + \bar{T}]. \quad (4.204)$$

If $\tilde{u}_i^*(t, s)$, $i \in N$, are the generalized open-loop cooperative strategies in the differential game with continuous updating, then by Definition 4.3.1, for every fixed $t \geq t_0$, $\tilde{u}_i^*(t, s)$, $i \in N$ is an open-loop cooperative strategy in the game $\Gamma^c(x^*(t), t, t + \bar{T})$. Therefore, for any fixed $t \geq t_0$, conditions (1)-(3) of the theorem are satisfied as necessary conditions for the cooperative open-loop strategies; see [231]. The theorem is proved.

On the other hand, if for every $t \geq t_0$, the Hamiltonian H^t is concave in $(x_t, u(t, s))$, then the conditions of the theorem are sufficient for the cooperative open-loop solution [232]. \square

Then, for any fixed $t \in [t_0, +\infty)$, we obtain the generalized open-loop cooperative strategy $\tilde{u}^*(t, s) = (\tilde{u}_1^*(t, s), \dots, \tilde{u}_n^*(t, s))$ and the corresponding state trajectory $x_t^*(s)$, $s \in [t, t + \bar{T}]$. Using *Definition 4.3.1* and *Definition 4.3.2*, we obtain the cooperative strategies with continuous updating $\{u_i^*(t)\}_{i \in N}$ and the corresponding cooperative trajectory with continuous updating $x^*(t)$, $t \in [t_0, +\infty)$.

To get the characteristic function of the grand coalition in the subgame $\Gamma(x^*(t), s, t + \bar{T})$, we substitute $\tilde{u}^*(t, \tau)$ and $x_t^*(\tau)$ into the corresponding payoff function, denoting by $\tilde{V}_t(N; x^*(t), s, t + \bar{T})$ the function of coalition N . The current-value maximized cooperative payoff $\tilde{V}_t(N; x^*(t), s, t + \bar{T})$ can be expressed as

$$\tilde{V}_t(N; x^*(t), s, t + \bar{T}) = \sum_{i=1}^n \int_s^{t+\bar{T}} g_i[\tau, x_t^*(\tau), \tilde{u}^*(t, \tau)] d\tau. \quad (4.205)$$

Step 2: Calculating generalized open-loop Nash equilibrium with continuous updating

The problem of the non-cooperative subgame along the cooperative trajectory with continuous updating $\Gamma(x^*(t), s, t + \bar{T})$ can be stated as follows:

$$\begin{aligned} & \max_{u_i \in U_i} K_i^t(x^*(t), s, t + \bar{T}; u_i(t, s), \tilde{u}_{-i}^{NE}(t, s)) = \\ & = \max_{u_i \in U_i} \int_s^{t+\bar{T}} g_i[\tau, x_t(\tau), u_i(t, \tau), \tilde{u}_{-i}^{NE}(t, \tau)] d\tau \quad s.t. \ x_t(\tau) \text{ satisfies (4.202)} \end{aligned} \quad (4.206)$$

where $\tilde{u}_{-i}^{NE}(t, \tau) = (\tilde{u}_1^{NE}(t, \tau), \dots, \tilde{u}_{i-1}^{NE}(t, \tau), \tilde{u}_{i+1}^{NE}(t, \tau), \dots, \tilde{u}_n^{NE}(t, \tau))$. In this setting, the current-value Hamiltonian function can be written as

$$H_i^t(\tau, x_t(\tau), u(t, \tau), \psi_i^t(\tau)) = g_i(\tau, x_t, u) + \psi_i^t(\tau) f(\tau, x_t, u), \tau \in [s, t + \bar{T}], \quad i \in N, \quad (4.207)$$

where $\tau \in [s, t + \bar{T}]$, $t \in [t_0, +\infty)$. Using Pontryagin's maximum principle with continuous updating [120], we get the open-loop Nash equilibrium $\{\tilde{u}_i^{NE}(t, \tau)\}_{i \in N}$ and the corresponding trajectory $x_t^{NE}(\tau)$, $\forall \tau \in [s, t + \bar{T}]$, $t \in [t_0, +\infty)$. It is then easy to derive the characteristic function of a one-player coalition as follows: for

each $i = 1, 2, \dots, n$,

$$\tilde{V}_t(\{i\}; x^*(t), s, t + \bar{T}) = \int_s^{t + \bar{T}} g_i[\tau, x_t^{NE}(\tau), \tilde{u}^{NE}(t, \tau)] d\tau, \quad \forall s \in [t, t + \bar{T}], t \in [t_0, +\infty). \quad (4.208)$$

Step 3: Calculating characteristic function with continuous updating for all remaining coalitions

We need to find only the coalitions containing more than one player, except for the grand coalition. There will be $2^n - n - 2$ subsets obtained in the following way. We apply the δ -characteristic function so that the players from S maximize their total payoff $\sum_{i \in S} K_i^t(x^*(t), s, t + \bar{T}; u_S, \tilde{u}_{N \setminus S}^{NE})$ along the cooperative strategy with continuous updating $x^*(t)$, whereas the other players (those from $N \setminus S$) use their generalized open-loop Nash-equilibrium strategies

$$\tilde{u}_{N \setminus S}^{NE} = \{\tilde{u}_j^{NE}\}_{j \in N \setminus S}.$$

Thus, we have a two-stage construction procedure for the characteristic function: (1) find the generalized open-loop Nash equilibrium strategies $\tilde{u}_i^{NE}(t, \tau)$ for all players $i \in N$, which have been obtained in *Step 2*; (2) “freeze” the Nash equilibrium strategies $\tilde{u}_j^{NE}(t, \tau)$ for the players from $N \setminus S$, and (like for the player from S) maximize their total payoff over $u_S = \{u_i\}_{i \in S}$. To demonstrate the value function of the subgame $\Gamma(x^*(t), s, t + \bar{T}), \forall t \in [t_0, +\infty), s \in [t, t + \bar{T}]$, we present the following concept.

Definition 4.3.7 *A set of strategies $\tilde{u}_S^* = \{\tilde{u}_i^*(t, \tau)\}_{i \in S}, \tau \in [s, t + \bar{T}]$, provides a generalized open-loop optimal strategy for coalition $S \subset N$ in the subgame with continuous updating $\Gamma(x^*(t), s, t + \bar{T})$ if this set is the solution obtained using the*

Pontryagin maximum principle for the following problem:

$$\begin{aligned}
& \max_{u_S \in U_S} \sum_{i \in S} K_i^t(x^*(t), s, t + \bar{T}; u_S, \tilde{u}_{N \setminus S}^{NE}) \\
& = \max_{u_S \in U_S} \sum_{i \in S} \int_s^{t+\bar{T}} g_i[\tau, x_t^S(\tau), u_S(t, \tau), \tilde{u}_{N \setminus S}^{NE}(t, \tau)] d\tau \\
& \text{s.t. } x_t^S(\tau) = f(\tau, x_t^S(\tau), u_S(t, \tau), \tilde{u}_{N \setminus S}^{NE}(t, \tau)), \quad x_t^S(s) = x_{t,s}.
\end{aligned} \tag{4.209}$$

The Hamiltonian function for problem (4.209) has the following form, $\forall S^1 \subset N$:

$$\begin{aligned}
H_S^t(\tau, x_t^S(\tau), u_S(t, \tau), \tilde{u}_{N \setminus S}^{NE}(t, \tau), \psi_S^t) & = \sum_{i \in S} g_i(\tau, x_t^S, u_S, \tilde{u}_{N \setminus S}^{NE}) \\
& + \psi_S^t(\tau) f(\tau, x_t^S, u_S, \tilde{u}_{N \setminus S}^{NE}).
\end{aligned} \tag{4.210}$$

Theorem 4.3.6 *Let:*

- (i) $f(\tau, \cdot, u(t, \tau))$ be continuously differentiable on \mathbb{R}^l , $\forall \tau \in [s, t + \bar{T}]$,
- (ii) $g_i(\cdot, \cdot, u(t, \tau))$ be continuously differentiable on $\mathbb{R} \times \mathbb{R}^l$.

A set of strategies $\tilde{u}_S^* = \{\tilde{u}_i^*(t, \tau)\}_{i \in S}$ provides generalized open-loop optimal strategies of coalition S in the subgame with continuous updating $\Gamma(x^*(t), s, t + \bar{T})$ for problem (4.209) if there exist $2^n - n - 2$ costate functions $\psi_S^t(\tau)$, where $\tau \in [s, t + \bar{T}]$ and $S \subset N$, such that the following relations are satisfied $\forall s \in [t, t + \bar{T}]$ and $t \in [t_0, +\infty)$:

$$(1) \dot{x}_t^S(\tau) = f(\tau, x_t^S, \tilde{u}_S^*(t, \tau), \tilde{u}_{N \setminus S}^{NE}), \quad x_t^S(s) = x_{t,s}, \quad \forall \tau \in [s, t + \bar{T}],$$

$$(2) \tilde{u}_S^*(t, \tau) = \arg \max_{u_S \in U_S} H_S(x_t^S(\tau), u_S(t, \tau), \tilde{u}_{N \setminus S}^{NE}(t, \tau), \psi_S^t(\tau)), \quad \forall \tau \in [s, t + \bar{T}],$$

where $\tilde{u}_S^*(t, \tau) = \{\tilde{u}_i^S(t, \tau)\}_{i \in S}$

$$(3) \dot{\psi}_S^t(\tau) = -\frac{\partial}{\partial x_t} H_S^t(\tau, x_t(\tau), \tilde{u}_S^*(t, \tau), \tilde{u}_{N \setminus S}^{NE}(t, \tau), \psi_S^t(\tau)), \quad \text{where } \tau \in [s, t + \bar{T}],$$

$S \subset N$

$$\psi_S^t(t + \bar{T}) = 0, \quad S \subset N.$$

Proof Follows from the proof of Theorem 4.3.5. □

¹The superscript "S" denotes coalition S , e.g., x_t^S , ψ_S^t , and u_S .

Hence, the players in coalition S will adopt the generalized open-loop optimal control $\tilde{u}_S^*(t, \tau)$ characterized by Theorem 4.3.6. Note that these controls are functions of the fixed time $t \in [t_0, +\infty)$ and current time $\tau \in [s, t + \bar{T}]$.

An illustration of the characteristic function for coalition $S \subseteq N$ is as follows:

$$\tilde{V}_t(S; x^*(t), s, t + \bar{T}) = \sum_{i \in S} \int_s^{t+\bar{T}} g_i[\tau, x_t^S(\tau), \tilde{u}_S^*(t, \tau), \tilde{u}_{N \setminus S}^{NE}(t, \tau)] d\tau, \quad (4.211)$$

$$\forall s \in [t, t + \bar{T}], \quad t \in [t_0, +\infty),$$

where $x_t^S(\tau)$ is the trajectory at a time instant $\tau \in [s, t + \bar{T}]$ when the players from coalition S use the generalized open-loop optimal strategies $\tilde{u}_S^*(t, \tau)$, whereas the players from $N \setminus S$ use the generalized open-loop Nash equilibrium $\tilde{u}_{N \setminus S}^{NE}(t, \tau)$ derived in Step 2.

Let us consider the characteristic function in the game model with continuous updating. First, assuming that the function $\tilde{V}_t(S; x^*(t), s, t + \bar{T})$, $\forall S \subseteq N$, is continuously differentiable with respect to $s \in [t, t + \bar{T}]$ and integrable with respect to $t \in [t_0, +\infty)$, we define the characteristic function in the game model with continuous updating $V(S; x^*(t), t, T)$ in the following way:

Definition 4.3.8 *A function $V(S; x^*(t), t, T)$, $t \in [t_0, T]$, $S \subseteq N$, is a characteristic function in the differential game with continuous updating $\Gamma(x^*(t), t, T)$ if*

$$V(S; x^*(t), t, T) = \int_t^T -\frac{d}{ds} \tilde{V}_{\tau'}(S; x^*(\tau'), s, \tau' + \bar{T})|_{s=\tau'} d\tau', \quad t \in [t_0, T], \quad S \subseteq N, \quad (4.212)$$

where $\tilde{V}_{\tau'}(S; x^*(\tau'), s, \tau' + \bar{T})$, $s \in [\tau', \tau' + \bar{T}]$, $\tau' \in [t, T]$, $S \subseteq N$, is the characteristic function in the game $\Gamma(x^*(\tau'), s, \tau' + \bar{T})$ defined on the interval $[s, \tau' + \bar{T}]$.

In (4.212), the integral is on a finite time interval: only in this case, we can claim that the values of the characteristic function with continuous updating are finite. In the model example below, we will calculate the characteristic function and Shapley value using the final interval method. We assume that the superadditivity conditions

are satisfied:

$$\begin{aligned} V(S_1 \cup S_2; x^*(t), t, T) &\geq V(S_1; x^*(t), t, T) + V(S_2; x^*(t), t, T), \\ \forall S_1, S_2 \subseteq N, S_1 \cap S_2 &= \emptyset. \end{aligned} \quad (4.213)$$

Step 4: Calculating Shapley value based on characteristic function with continuous updating

Consider again the cooperative game model $\Gamma^c(x^*(t), t, T)$ with continuous updating. Let the players form different coalitions consisting of a subset of all players $K \subseteq N$. There are k players in the subset K . An imputation set in the cooperative game $\Gamma^c(x^*(t), t, T)$ is the set $L(x^*(t), t, T) = \{\xi(x^*(t), t, T) = (\xi_1(x^*(t), t, T), \dots, \xi_n(x^*(t), t, T))\}$ satisfying the conditions

$$\begin{aligned} \xi_i(x^*(t), t, T) &\geq V(\{i\}; x^*(t), t, T), \quad i \in N; \\ \sum_{i \in N} \xi_i(x^*(t), t, T) &= V(N; x^*(t), t, T). \end{aligned} \quad (4.214)$$

A cooperative solution (optimality principle) is a non-empty subset of the imputation set $L(x^*(t), t, T)$. In particular, the Shapley value $sh(x^*(t), t, T) = (sh_1(x^*(t), t, T), \dots, sh_n(x^*(t), t, T))$ is an imputation whose components are defined as

$$sh_i(x^*(t), t, T) = \sum_{K \subseteq N, i \in K} \frac{(k-1)!(n-k)!}{n!} [V(K; x^*(t), t, T) - V(K \setminus i; x^*(t), t, T)], \quad (4.215)$$

where $K \setminus i$ is the relative complement of i in K , and the function $V(K; x^*(t), t, T)$ is given by Definition 4.3.8 (the profit of coalition K). Meanwhile, $[V(K; x^*(t), t, T) - V(K \setminus i; x^*(t), t, T)]$ is the marginal contribution of player i to coalition K .

There are many other cooperative optimality principles, e.g., the von Neumann-Morgenstern (NM) solution, Core, or Nucleus. In all cases, they involve some subsets of the game's imputation set.

4.3.3.2 Cooperative Differential Game of Pollution Control

Let us consider the following game proposed by Long [233]. Countries are numbered by $i \in N$, and $n = |N|$. Each player has an industrial production site. Assume that the production output is proportional to the pollutants u_i . Thus, the strategy of a player is to choose the amount of pollutants emitted into the atmosphere.

4.3.3.2.1 Original Game Model

Pollution accumulates over time. We denote by $x(t)$ the stock of pollution at a time instant t , supposing that the countries "contribute" to the same stock of pollution. For simplicity, the evolution of stock $x(t)$ is represented by the linear equation

$$\dot{x}(t) = \sum_{i=1}^n u_i(t) - \delta x(t), \quad x(t_0) = x_0,$$

where δ is a constant rate of decay (in other words, the absorption rate of pollution by the environment).

Let the absorption coefficient δ be equal to zero:

$$\dot{x}(t) = \sum_{i=1}^n u_i(t), \quad x(t_0) = x_0. \quad (4.216)$$

Pollution is a "public bad" due to adverse effects on health, quality of life, and productivity. Assume that these adverse effects can be represented by adding x in the variables of the instantaneous social welfare function F_i , with negative derivative:

$$F_i = F_i(x, t, u_i), \quad \frac{\partial F_i}{\partial x} < 0.$$

In each country, the aggregate social welfare is taken to be the integral of the instantaneous social welfare. Thus, the payoff of player i can be defined as follows:

$$K_i(x_0, t_0, T; u) = \int_{t_0}^T F_i(x, t, u_i) dt. \quad (4.217)$$

For tractability, the function F_i is often assumed to take the separable form:

$$F_i(x, t, u_i(t)) = R_i(u_i(t)) - D_i(x),$$

where $R_i(u_i)$ may be treated as the utility of the benefit, and $D_i(x)$ as the "disutility" caused by pollution. Following standard practice, let $R_i(u_i)$ be strictly concave and increasing in u_i , and $D_i(x)$ be convex and increasing in x . The possibility of linear D_i is not ruled out.

Assume that the environmental damage cost of player i due to the pollution stock is $D_i(x) = d_i x$, and the damage cost $D_i(x)$ increases convexly. In the environmental economics literature, typical assumptions are as follows: the production benefit function of player i is a function of emissions, i.e., $R_i(u_i(t)) = b_i u_i - \frac{1}{2} u_i^2$, satisfying $R_i(0) = 0$, where b_i and d_i are positive parameters. For this benefit function to be concavely increasing in emissions, we impose the constraint $u_i(t) \in (0, b_i)$.

Suppose that the game is played in the cooperative scenario: the players have the opportunity to cooperate for achieving the maximum total payoff:

$$\sum_{i=1}^n K_i(x_0, t_0, T; u) = \sum_{i=1}^n \int_{t_0}^T ((b_i - \frac{1}{2} u_i) u_i - d_i x) dt \rightarrow \max_{u_1, u_2, \dots, u_n} . \quad (4.218)$$

To solve the optimization problem in (4.218) and (4.216), we employ Pontryagin's maximum principle to characterize the solution as follows. Obviously, these games are linear state games.² These games have the property that their open-loop Nash equilibria are Markov perfect. The class of linear state games has a very useful property. The linearity in the state variables, together with the decoupled structure between the state variables and the control variables, implies that the open-loop equilibrium is Markov perfect, and the value functions are linear in the state variables.

As is easily verified, the optimal emissions control of player i in the original

²Games for which: (a) the system dynamics and the utility functions are polynomials of degree 1 in the state variables and (b) the interaction between the control variables and state variables has some structural property (described below). We call this class of games linear state games.

differential game model is given by

$$u_i^*(t) = b_i - \sum_{i=1}^n d_i(T - t), \quad i \in N. \quad (4.219)$$

To obtain the cooperative state trajectory in the original differential game, it suffices to insert $u_i^*(t)$ into the dynamics and solve the differential equation to get

$$x^*(t) = x_0 + \left(\sum_{i=1}^n b_i - n \sum_{i=1}^n d_i T \right) (t - t_0) + n \sum_{i=1}^n d_i \frac{t^2 - t_0^2}{2}. \quad (4.220)$$

4.3.3.2.2 Pollution Control Game Model with Continuous Updating

In the game $\Gamma(x, t, t + \bar{T})$, the dynamics of the total amount of pollution $x_t(s)$ are described by

$$\dot{x}_t(s) = \sum_{i=1}^n u_i(t, s), \quad x_t(t) = x, \quad (4.221)$$

in which the absorption coefficient corresponding to the natural purification of the atmosphere is assumed to be 0.

The instantaneous payoff of player i is defined as

$$R_i(u_i(t, s)) = b_i u_i(t, s) - \frac{1}{2} u_i^2(t, s), \quad i \in N.$$

Each player has to bear expenses due to pollution reduction. Thus, the instantaneous payoff (utility) of player i is equal to $R_i(u_i(t, s)) - d_i x_t(s)$, where $d_i > 0$.

Therefore, the payoff of player i is defined as

$$K_i^t(x, t, t + \bar{T}; u) = \int_t^{t+\bar{T}} \left((b_i - \frac{1}{2} u_i) u_i - d_i x_t \right) ds, \quad (4.222)$$

where $u_i = u_i(t, s)$ is the control of player i at a time instant $s \in [t, t + \bar{T}]$, and $x_t = x_t(s)$ is the pollution accumulation at the same time s .

Therefore, the payoff function of player i in the subgame with continuous

updating $\Gamma(x, s, t + \bar{T})$ has the form

$$K_i^t(x, s, t + \bar{T}; u) = \int_s^{t+\bar{T}} \left((b_i - \frac{1}{2}u_i(t, \tau))u_i(t, \tau) - d_i x_t(\tau) \right) d\tau, \quad i \in N, \quad (4.223)$$

where $x_t(\tau)$ and $u(t, \tau)$, $\tau \in [s, t + \bar{T}]$, are the trajectory and strategies in the game $\Gamma(x, s, t + \bar{T})$. The dynamics of the state are given by

$$\dot{x}_t(\tau) = \sum_{i=1}^n u_i(t, \tau), \quad x_t(s) = x_{t,s}. \quad (4.224)$$

Step 1: Optimizing total payment of grand coalition with continuous updating

Consider the game in the cooperative form. This means that all players join together to maximize their total payoff. We seek the optimal strategy profile $\tilde{u}^*(t, s) = (\tilde{u}_1^*(t, s), \dots, \tilde{u}_n^*(t, s))$ such that $\sum_{i=1}^n K_i^t \rightarrow \max_{u_1, u_2, \dots, u_n}$.

The optimization problem is as follows:

$$\sum_{i=1}^n K_i^t(x, t, t + \bar{T}; u) = \sum_{i=1}^n \int_t^{t+\bar{T}} \left((b_i - \frac{1}{2}u_i(t, s))u_i(t, s) - d_i x_t(s) \right) ds \rightarrow \max_{u_1, u_2, \dots, u_n} .$$

where $x_t(s)$ satisfies (4.221). (4.225)

To solve problem (4.225), we use the classical Pontryagin's maximum principle. The corresponding Hamiltonian is

$$H^t(s, x_t(s), u(t, s), \psi^t(s)) = \sum_{i=1}^n (b_i - \frac{1}{2}u_i)u_i - \sum_{i=1}^n d_i x_t + \psi^t(s)(u_1 + u_2 + \dots + u_n). \quad (4.226)$$

The first-order partial derivatives w.r.t. u_i 's are

$$\frac{\partial H^t}{\partial u_i}(s, x_t, u, \psi^t) = b_i - u_i + \psi^t = 0, \quad (4.227)$$

and the Hessian matrix $\frac{\partial^2 H^t}{\partial u^2}(s, x_t, u, \psi^t)$ is negative definite. Hence, the Hamiltonian

H^t is concave w.r.t. u_i , and we obtain the cooperative strategies

$$\tilde{u}_i^*(t, s) = b_i + \psi^t(s)$$

Consider the Pontryagin's maximum principle when dealing with the costate variable

$$\dot{\psi}^t(s) = \sum_{i=1}^n d_i, \quad \psi^t(t + \bar{T}) = 0.$$

Setting $\sum_{i=1}^n d_i = d_N$, we can get $\psi^t(s) = -d_N(t + \bar{T} - s)$. Finally, the form of the cooperative strategies is

$$\tilde{u}_i^*(t, s) = b_i - d_N(t + \bar{T} - s), \quad (4.228)$$

and from (4.221) we get the optimal (cooperative) trajectory

$$x_t^*(s) = x + b_N(s - t) - nd_N(t + \bar{T})(s - t) + nd_N \frac{s^2 - t^2}{2}, \quad (4.229)$$

where $x = x_t(t)$, $d_N = \sum_{i=1}^n d_i$, and $b_N = \sum_{i=1}^n b_i$.

According to the procedure (4.139), we construct the open-loop optimal cooperative strategies with continuous updating

$$u_i^*(t) = \tilde{u}_i^*(t, s)|_{s=t} = b_i - d_N \bar{T}. \quad (4.230)$$

Substituting $u_i^*(t)$ into the differential equation (4.216), we arrive in the optimal cooperative trajectory $x^*(t)$ with continuous updating:

$$x^*(t) = x_0 + b_N(t - t_0) - nd_N \bar{T}(t - t_0). \quad (4.231)$$

The cooperative strategies and corresponding trajectories for the original differential game model and the differential game with continuous updating are compared in Figs.4.21 and 4.22.

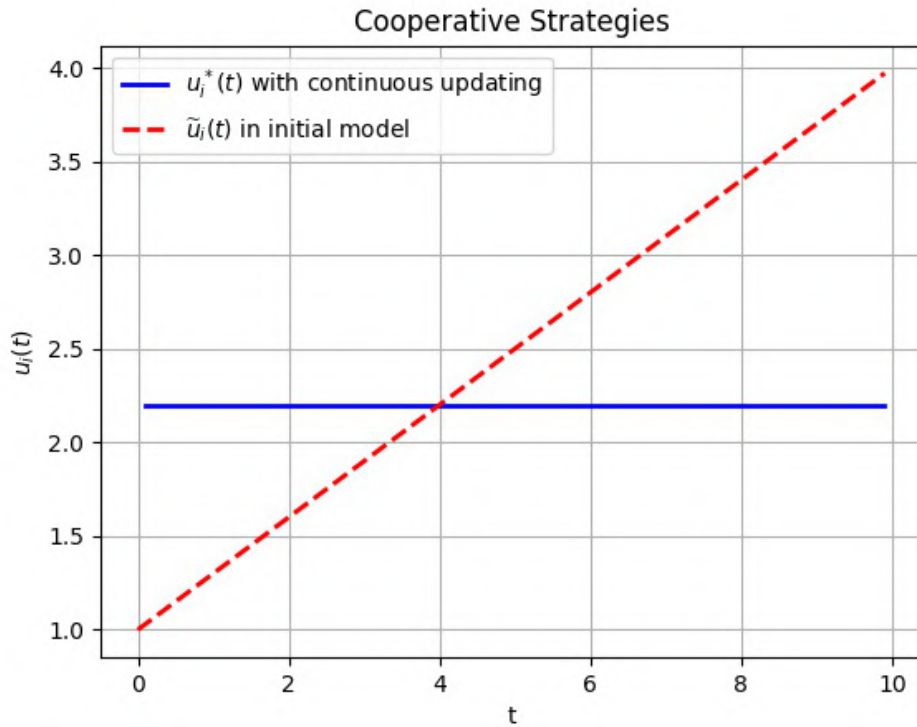


Figure 4.21: Cooperative strategies in original game $\tilde{u}(t)$ (red line) and cooperative strategies with continuous updating $u^*(t)$ (blue line).

According to Fig.4.21, the optimal control with continuous updating is more stable than the optimal control in the original game model. Also, starting from the instant $t = 4$, the optimal control in the original game exhibits violent growth. This means that the players should increase the pollution emissions into the atmosphere in the original differential game model, a harmful result. From the perspective of the environment and the stability of related activity, it is better to adopt the optimal strategies based on the continuous updating method.

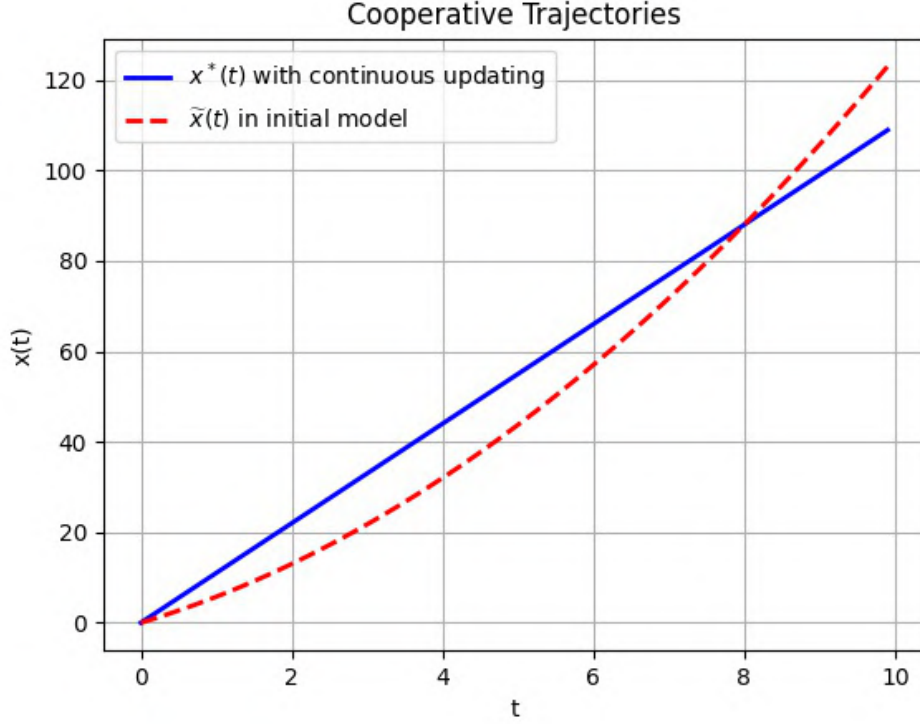


Figure 4.22: Cooperative trajectory in original game $\tilde{x}(t)$ (red line) and cooperative trajectory with continuous updating $x^*(t)$ (blue line).

According to Fig.4.22, starting from the instant $t = 8$, pollution with continuous updating is less than pollution in the original game, and the growth of pollution in the original game is sharper (hence worse for the environment). Under continuous updating, this eventually leads to a lower level of pollution. Therefore, the new method for solving this problem is better for the environment over the long term.

Next, for the subgame $\Gamma(x^*(t), s, t + \bar{T})$ of the differential game with continuous updating $\Gamma(x^*(t), t, t + \bar{T})$, the characteristic function of the grand coalition N is given by $\tilde{V}_t(N; x^*(t), s, t + \bar{T})$. It can be represented as

$$\tilde{V}_t(N; x^*(t), s, t + \bar{T}) = \sum_{i=1}^n \int_s^{t+\bar{T}} \left((b_i - \frac{1}{2} \tilde{u}_i^*(t, \tau)) \tilde{u}_i^* - d_i x_t^*(\tau) \right) d\tau,$$

where $x_t^*(\tau)$ satisfies (4.229) with $x = x^*(t)$, and $\tilde{u}_i^*(t, \tau)$ satisfies (4.228). Therefore,

we can get the value function of the grand coalition N :

$$\begin{aligned} \tilde{V}_t(N; x^*(t), s, t + \bar{T}) = & (t + \bar{T} - s) \left[\frac{1}{2} \tilde{b}_N - d_N x^*(t) + \frac{d_N b_N}{2} (t - \bar{T} - s) \right. \\ & \left. - \frac{1}{3} n d_N^2 ((t + \bar{T} - s)^2 - \frac{3}{2} \bar{T}^2) \right], \end{aligned} \quad (4.232)$$

where $d_N = \sum_{i=1}^n d_i$ and $b_N = \sum_{i=1}^n b_i$ have been defined above, and $\tilde{b}_N = \sum_{i=1}^n b_i^2$. Note that in (4.232), $x^*(t)$ represents the cooperative pollution with continuous updating at the time instant t .

For the problem under consideration, it is also possible to use dynamic programming based on the solution of the Hamilton–Jacobi–Bellman equation. As is easily checked, for the Bellman function $\tilde{V}_t = A(t, s)x_t(s) + B(t, s)$ we obtain the same result as using Pontryagin’s maximum principle.

Step 2: Calculating generalized open-loop Nash equilibrium with continuous updating

The Hamiltonian for each player $i = 1, 2, \dots, n$ is

$$H_i^t(\tau, x_t(\tau), u(t, \tau), \psi_i^t(\tau)) = (b_i - \frac{1}{2}u_i)u_i - \sum_{i=1}^n d_i x_t + \psi_i^t(\tau)(u_1 + u_2 + \dots + u_n). \quad (4.233)$$

Its first-order partial derivatives w.r.t. u_i s are

$$\frac{\partial H_i^t}{\partial u_i}(\tau, x_t(\tau), u(t, \tau), \psi_i^t(\tau)) = b_i - u_i + \psi_i^t = 0, \quad (4.234)$$

and the Hessian matrix $\frac{\partial^2 H_i^t}{\partial u_i^2}(\tau, x_t, u, \psi_i^t)$ is negative definite. Hence, the Hamiltonian H_i^t is concave in u_i . We obtain the optimal controls

$$\tilde{u}_i^{NE}(t, \tau) = b_i - d_i(t + \bar{T} - \tau), \quad i = 1, 2, \dots, n. \quad (4.235)$$

Consider the subgame starting at a time instant $s \in [t, t + \bar{T}]$. We can easily derive the corresponding trajectory (for the Nash equilibrium case) of the subgame

$\Gamma(x^*(t), s, t + \bar{T})$ along the cooperative trajectory. In other words, $x_{t,s} = x_t^*(s)$ is

$$\begin{aligned} x_t^{NE}(\tau) &= x_t^*(s) + b_N(\tau - s) + d_N\left(\frac{(t + \bar{T} - \tau)^2}{2} - \frac{(t + \bar{T} - s)^2}{2}\right) \\ &= x^*(t) + b_N(\tau - t) - nd_N(t + \bar{T})(s - t) + nd_N\frac{s^2 - t^2}{2} \\ &\quad + \frac{d_N}{2}((t + \bar{T} - \tau)^2 - (t + \bar{T} - s)^2). \end{aligned} \quad (4.236)$$

The maximum payoff of each player $i = 1, 2, \dots, n$ in the subgame starting at the time instant s from the state $x^*(t)$ takes the form

$$\begin{aligned} \tilde{V}_i(\{i\}; x^*(t), s, t + \bar{T}) &= \int_s^{t+\bar{T}} \left((b_i - \frac{1}{2}\tilde{u}_i^{NE}(t, \tau))\tilde{u}_i^{NE} - d_i x_t^{NE}(\tau) \right) d\tau \\ &= (t + \bar{T} - s) \left[\frac{b_i^2}{2} - d_i x^*(t) + \frac{nd_i d_N}{2}(s - t)(t - s + 2\bar{T}) \right. \\ &\quad \left. - \frac{1}{6}d_i^2(t + \bar{T} - s)^2 + \frac{1}{3}d_i d_N(t + \bar{T} - s)^2 - \frac{1}{2}d_i b_N(s + \bar{T} - t) \right]. \end{aligned} \quad (4.237)$$

Step 3: Calculating characteristic function for all remaining coalitions in differential games with continuous updating

We can also find the controls and corresponding value functions for different coalitions. However, their form will depend on the way of defining the corresponding optimal control problems.

Let us construct the characteristic function based on the approach of $\delta - c.f.$. The characteristic function of coalition S is calculated in two stages. The first stage is complete: the Nash equilibrium strategies for all players have been found at *Step 2*. In the second stage, assume that the left-out players $j \in N \setminus S$ implement their Nash optimal strategies $\tilde{u}_j^{NE}(t, \tau)$, whereas the players from coalition S seek to maximize their joint payoff $\sum_{i \in S} K_i$. Consider the case of S -coalition. It seems instructive to go through the calculations in detail. The corresponding Hamiltonian for coalition S is

$$H_S^t(\tau, x_t(\tau), u(t, \tau), \psi_S^t(\tau)) = \sum_{i \in S} \left((b_i - \frac{1}{2}u_i)u_i \right) - d_S x_t + \psi_S^t \left(\sum_{i \in S} u_i + \sum_{j \in N/S} \tilde{u}_j^{NE} \right),$$

where $d_S = \sum_{i \in S} d_i$. Note that $u_j, j \in N/S$, are replaced by \tilde{u}_j^{NE} found earlier.

The optimal strategies of the players from coalition S are

$$\tilde{u}_i^*(t, \tau) = b_i + \psi_S^t(\tau), \quad i \in S.$$

The differential equation for $\psi_S^t(\tau)$ is $\dot{\psi}_S^t(\tau) = d_S$, solved to ensure that $\psi_S^t(t + \bar{T}) = 0$. Finally, we get $\psi_S^t(\tau) = -d_S(t + \bar{T} - \tau)$. Substituting the expression for $\psi_S^t(\tau)$ into \tilde{u}_i^* yields

$$\tilde{u}_i^*(t, \tau) = b_i - d_S(t + \bar{T} - \tau), \quad i \in S \quad (4.238)$$

where $d_S = \sum_{i \in S} d_i$. Thus, the players from coalition S implement their optimal strategies \tilde{u}_S^* , whereas the left-out players adhere to their Nash equilibrium.

At the next step, we integrate (4.224) starting from the point $x_t^*(s)$ to get $x_t^S(\tau)$:

$$\begin{aligned} x_t^S(\tau) = & x_t^*(s) + b_N(\tau - s) + \frac{(t + \bar{T} - \tau)^2}{2}((k - 1)d_S + d_N) \\ & - \frac{(t + \bar{T} - s)^2}{2}((k - 1)d_S + d_N). \end{aligned} \quad (4.239)$$

$x_t^S(\tau)$ is the trajectory in the subgame $\Gamma(x^*(t), s, t + \bar{T})$ of the game $\Gamma(x^*(t), t, t + \bar{T})$ starting at the time instant s from $x_t^*(s)$, when the players from coalition S use the strategies $\tilde{u}_S^*(t, \tau)$, and the players from coalition $N \setminus S$ use $\tilde{u}^{NE}(t, \tau)$. Along the cooperative trajectory, we can substitute $x_t^*(s)$ (4.229) to get the state variable $x_t^S(\tau)$ depending on $x = x^*(t)$:

$$\begin{aligned} x_t^S(\tau) = & x^*(t) + b_N(\tau - t) - nd_N(t + \bar{T})(s - t) + nd_N \frac{s^2 - t^2}{2} + \\ & \frac{(t + \bar{T} - \tau)^2}{2}((k - 1)d_S + d_N) - \frac{(t + \bar{T} - s)^2}{2}((k - 1)d_S + d_N). \end{aligned} \quad (4.240)$$

The corresponding value of the characteristic function $\tilde{V}_t(S; x^*(t), s, t + \bar{T})$ is

$$\begin{aligned} \tilde{V}_t(S; x^*(t), s, t + \bar{T}) = & (t + \bar{T} - s) \left[\frac{1}{2} \tilde{b}_S - d_S x^*(t) + \frac{d_S d_N n}{2} (t + 2\bar{T} - s)(s - t) + \right. \\ & \left. (t + \bar{T} - s)^2 \left(\frac{k - 2}{6} d_S^2 + \frac{d_S d_N}{3} \right) - \frac{d_S b_N}{2} (s + \bar{T} - t) \right]. \end{aligned} \quad (4.241)$$

According to Definition 4.3.8, the characteristic function of the differential game

with continuous updating has the following form:

$$\begin{aligned}
 V(S; x^*(t), t, T) &= \int_t^T -\frac{d\tilde{V}_{\tau'}(S; x^*(\tau'), s, \tau' + \bar{T})}{ds} \Big|_{s=\tau'} d\tau' \\
 &= (T-t) \left[\frac{1}{2} \tilde{b}_S - d_S x_0 + \bar{T}^2 \left(\frac{k-2}{2} d_S^2 + (1-n) d_S d_N \right) \right. \\
 &\quad \left. - \frac{d_S (b_N - n d_N \bar{T})}{2} (T+t-2t_0) \right].
 \end{aligned} \tag{4.242}$$

We check the superadditivity condition (4.213) for the constructed characteristic function $V(S; x^*(t), t, T)$. For any $S, P \subseteq N$ and $S \cap P = \emptyset$, where $|S| = k \geq 1$ and $|P| = m \geq 1$, the following relation holds:

$$\begin{aligned}
 &V(S \cup P; x^*(t), t, T) - V(S; x^*(t), t, T) - V(P; x^*(t), t, T) \\
 &= \bar{T}^2 \left[(k+m-2) d_S d_P + \frac{k}{2} d_P^2 + \frac{m}{2} d_S^2 \right] \geq 0.
 \end{aligned} \tag{4.243}$$

Thus, the δ -characteristic function $V(S; x^*(t), t, T)$ is a superadditive function without any additional conditions imposed on the model parameters.

In the following figure, we will compare the characteristic function for the grand coalition N in the original game model and the differential game with continuous updating.

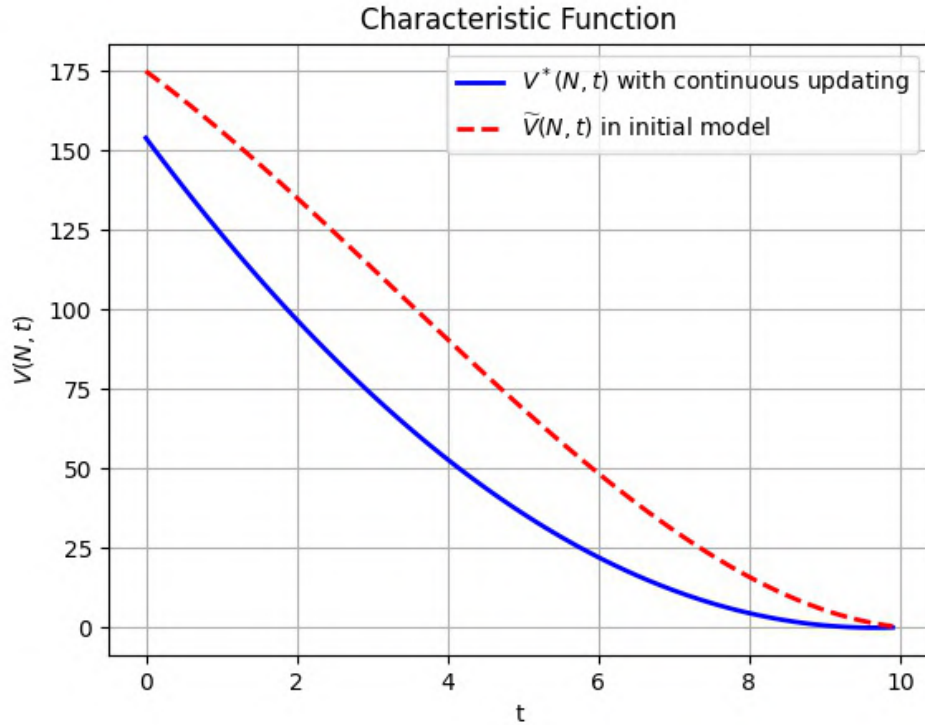


Figure 4.23: Characteristic function value for coalition N in original game (red line) and characteristic function value for coalition N with continuous updating $V^*(N, t, T)$ (blue line).

The characteristic function value in the original model is greater than in the model with continuous updating. The reason is that the complexity of the information within the continuous updating setting can reduce the effectiveness of the coalition. Note that the continuous updating case is more realistic. We can conclude that the payoff of the coalition decreases because, as time goes on, pollution accumulates in the air. The player's payoff depends on the levels of pollution, and the payoff decreases as pollution increases. Also, note that the coalition's effectiveness decreases in the original game model faster than in the game with continuous updating.

Step 4: Calculating Shapley value based on characteristic function with continuous updating

Any of the known principles of optimality can be applied to find a cooperative solution. First of all, the concept $\sum d_j d_l$ ³ represents the interaction of cost among the players. Now consider the cooperative solution of the differential game with

³Note that $d_k d_j = d_j d_k$.

continuous updating. According to the procedure (4.215), we construct the Shapley value for any $i \in N$ with continuous updating using the characteristic function with continuous updating of the auxiliary subgame and get

$$\begin{aligned}
 sh_i(x^*(t), t, T) &= \sum_{K \subseteq N, i \in K} \frac{(k-1)!(n-k)!}{n!} [V(K; x^*(t), t, T) - V(K \setminus i; x^*(t), t, T)] \\
 &= (T-t) \left[-d_i x_0 + \frac{1}{2} b_i^2 + \frac{nd_i d_N \bar{T} - d_i b_N}{2} (T+t-2t_0) \right. \\
 &\quad \left. + \bar{T}^2 \left(\frac{1-2n}{3} d_i d_N - \frac{4+n}{12} d_i^2 + \frac{1}{3} \left(\sum_{\substack{j, l \neq i \\ j \neq l \in N}} d_j d_l \right) + \frac{1}{4} \tilde{d}_N \right) \right].
 \end{aligned} \tag{4.244}$$

The Shapley value in the subgames with continuous updating and the original game model along the optimal cooperative trajectory $x^*(t)$ is shown in Figure 4.24.

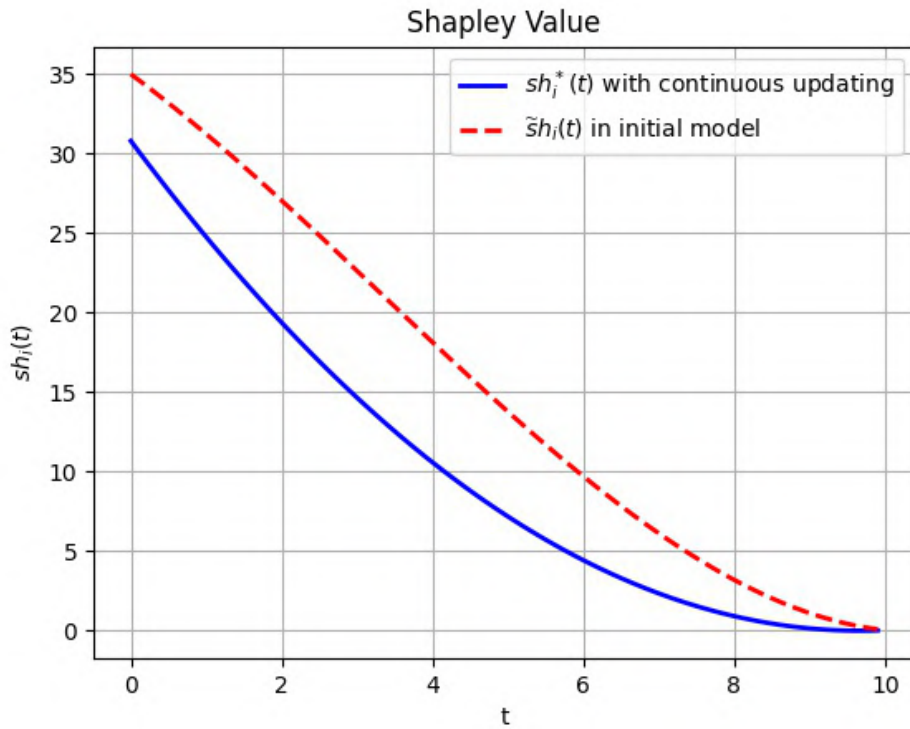


Figure 4.24: Shapley value for player i in original game (red line) and Shapley value with continuous updating $sh_i^*(t)$ (blue line).

According to Figure 4.24, in the more realistic case (continuous updating), a player with continuous updating can get less allocation from the coalition compared to the original game model. This fact can be explained as follows. At an early

stage, the pollution emitted into the atmosphere is greater than in the original game model, and in the late stage, the pollution with continuous updating is less than in the original game model. Thus, starting from $t = 0$, the players obtain more in the original game model, but they all get 0 at the end of the same period. Hence, as pollution intensifies, the production benefits of different countries gradually decrease.

4.3.4 Linear-Quadratic Case of Cooperative Differential Game with Continuous Updating

The linear-quadratic case of this class of games is particularly important for practical problems arising in human-machine interaction engineering. In this section, it is particularly interesting that the open-loop strategies are used to construct the optimal ones, but subsequently, we obtain strategies in the feedback form. These strategies are employed to define the notions of Shapley value and Nash equilibrium as optimality principles for cooperative and non-cooperative cases, respectively, and present the optimal strategies in the linear-quadratic case. Additionally, for the construction of the characteristic function with continuous updating, the open-loop form of Nash equilibrium with continuous updating is presented.

In this section, we will use only the loop-based strategies to construct the cooperative strategies with continuous updating $u_{ol}^*(t, x)$ and the Nash equilibrium strategies $u_{ol}^{NE}(t, x)$ to construct the characteristic function with continuous updating. Therefore, we will adopt the simplified notations $u^*(t, x)$ and $u^{NE}(t, x)$, respectively.

4.3.4.1 Cooperative Strategies with Continuous Updating

To construct the cooperative game model with continuous updating, we need defining the cooperative strategies and related trajectory with continuous updating. Sufficient conditions for the existence of cooperative strategies in the linear-quadratic differential game model with continuous updating are presented below.

Theorem 4.3.7 *For an N -person linear-quadratic differential game with $Q_i \geq 0$ and $R_{ij} \geq 0$ ($i, j \in N, i \neq j$), let there exist a solution set $\{Z^t, i \in N, t \geq t_0\}$ for the*

matrix Riccati differential equations

$$\begin{aligned}\dot{Z}^t(\tau) &= -\bar{T}A'Z^t(\tau) - \bar{T}Z^t(\tau)A + Z^t(\tau)SZ^t(\tau) - Q, \\ Z^t(1) &= 0,\end{aligned}\tag{4.245}$$

where $S = \bar{T}^2BR^{-1}B'$ and $Q = \sum_{i \in N} Q_i$, and $B = [B_1, \dots, B_n]$ and $R = \{R_{ij}\}_{i,j=1}^n$ are block matrices. Then, the differential game with continuous updating admits a cooperative solution with continuous updating given by

$$u^*(t, x) = -R^{-1}B'Z^t(0)\bar{T}x.$$

Proof For proving this theorem, we introduce the change of variables

$$\begin{aligned}s &= t + \bar{T}\tau, \\ y^t(\tau) &= x^t(t + \bar{T}\tau), \\ v_i^t(\tau, y) &= u_i(t + \bar{T}\tau, x), \quad i \in N.\end{aligned}\tag{4.246}$$

Substituting (5.12) into the motion equations (4.49) and the payoff function (4.50), we obtain

$$\dot{y}^t(\tau) = \bar{T}Ay^t(\tau) + \sum_{i=1}^N \bar{T}B_i v_i^t(\tau, y)\tag{4.247}$$

and

$$K^t(y^t, \tau; v^t) = \sum_{i \in N} \int_0^1 \left((y^t(s))' Q_i y^t(s) + \sum_{j=1}^N (v_j^t(s, y))' R_{ij} v_j^t(s, y) \right) ds, \quad i \in N.\tag{4.248}$$

Theorem 5.1 from [224] and the existence of a solution for the system of differential equations (4.245) lead to the following cooperative solution of the subgame $\Gamma(x, t, \bar{T})$:

$$v^{t,*}(\tau, y_0) = -R^{-1}B'Z^t(\tau)\bar{T}\Phi^t(\tau)y_0,$$

where

$$\begin{aligned}\frac{d\Phi^t}{d\tau} &= (A - SZ^t(\tau))\Phi^t(\tau), \\ \Phi^t(0) &= E.\end{aligned}$$

Returning to the original variables, we obtain the strategies

$$u^t(s, x) = -R^{-1}B'Z^t \left(\frac{s-t}{\bar{T}} \right) \bar{T}\Phi^t \left(\frac{s-t}{\bar{T}} \right) x.$$

Then the generalized cooperative solution of the game with continuous updating has the form

$$\tilde{u}^*(t, s, x) = -R^{-1}B'Z^t \left(\frac{s-t}{\bar{T}} \right) \bar{T}\Phi^t \left(\frac{s-t}{\bar{T}} \right) x. \quad (4.249)$$

We apply the procedure (4.11) to determine the Nash equilibrium with continuous updating using the generalized Nash equilibrium (4.249), $s = t$:

$$u^*(t, x) = -R^{-1}B'Z^t(0)\bar{T}x, \quad t \in [t_0, +\infty), \quad i \in N. \quad (4.250)$$

The proof of the theorem is complete. □

4.3.4.2 Nash Equilibrium Strategies with Continuous Updating for Constructing Characteristic Function

To construct the characteristic function with continuous updating in the next section, we need constructing the open-loop-based strategies with continuous updating. Related optimality conditions are presented below.

Theorem 4.3.8 *For an N -person linear-quadratic differential game with $Q_i \geq 0$ and $R_{ij} \geq 0$ ($i, j \in N, i \neq j$), let there exist a solution set $\{M_i^t, i \in N, t \geq t_0\}$ for the matrix Riccati differential equations*

$$\frac{dM_i^t(\tau)}{d\tau} + \bar{T}M_i^t(\tau)A + \bar{T}A'M_i^t(\tau) + Q_i - \bar{T}^2 M_i^t(\tau) \sum_{j \in N} B_j R_{jj}^{-1} B_j' M_j^t(\tau) = 0,$$

$$M_i^t(1) = 0, \quad i \in N.$$

(4.251)

Then the differential game with continuous updating admits an open-loop-based Nash equilibrium with continuous updating given by

$$u_i^{NE}(t, x) = -R_{ii}^{-1}B_i'M_i^t(0)\bar{T}x(t), \quad i \in N.$$

Proof For proving this theorem, we introduce the change of variables

$$\begin{aligned} s &= t + \bar{T}\tau, \\ y^t(\tau) &= x^t(t + \bar{T}\tau), \\ v_i^t(\tau, y) &= u_i(t + \bar{T}\tau, x), \quad i \in N. \end{aligned} \quad (4.252)$$

Substituting (4.252) into the motion equations (4.49) and the payoff function (4.50), we obtain

$$\dot{y}^t(\tau) = \bar{T}Ay^t(\tau) + \sum_{i=1}^N \bar{T}B_i v_i^t(\tau, y) \quad (4.253)$$

and

$$K_i^t(y^t, \tau; v^t) = \int_0^1 (y^t(s))' Q_i y^t(s) + \sum_{j=1}^N (v_j^t(s, y))' R_{ij} v_j^t(s, y) ds, \quad i \in N. \quad (4.254)$$

Theorem 6.12 from [80] and the existence of a solution for the system of differential equations (4.251) lead to the following open-loop-based Nash equilibrium strategies in the subgame $\Gamma(x, t, \bar{T})$:

$$v_i^{t, NE}(\tau, y_0) = -R_{ii}^{-1} B_i' M_i^t(\tau) \bar{T} \Phi^t(\tau) y_0,$$

where

$$\begin{aligned} \frac{d\Phi^t}{d\tau} &= \left(A - \sum_{i \in N} B_i R_{ii}^{-1} B_i' \right) \Phi^t(\tau), \\ \Phi^t(0) &= E. \end{aligned}$$

Returning to the original variables, we obtain the strategies

$$u_i^t(s, x) = -R_{ii}^{-1} B_i' M_i^t \left(\frac{s-t}{\bar{T}} \right) \bar{T} \Phi^t \left(\frac{s-t}{\bar{T}} \right) x.$$

Then the generalized open-loop Nash equilibrium in the game with continuous updating has the form

$$\tilde{u}_i^{NE}(t, s, x) = -R_{ii}^{-1} B_i' M_i^t \left(\frac{s-t}{\bar{T}} \right) \bar{T} \Phi^t \left(\frac{s-t}{\bar{T}} \right) x. \quad (4.255)$$

We apply the procedure (4.11) to determine the Nash equilibrium with continuous updating using the generalized Nash equilibrium (4.255), $s = t$:

$$u_i^{NE}(t, x) = -R_{ii}^{-1} B_i' M_i^t(0) \bar{T} x, \quad t \in [t_0, +\infty), \quad i \in N. \quad (4.256)$$

The proof of the theorem is complete. □

Remark 3 *Note that the open-loop-based solution with continuous updating has a feedback form, i. e., the open-loop-based Nash equilibrium with continuous updating explicitly depends on the current state. This fact is due to the way the solution is constructed—as a value of the generalized open-loop Nash equilibrium.*

4.3.4.3 Characteristic Function for Subgame on Interval $[t, t + \bar{T}]$

Consider coalition S in the n -player differential game $\Gamma(x, t, \bar{T})$ (4.49) (4.50). The characteristic function is defined as the total payoff of coalition S in the Nash equilibrium $u^{NE} = (u_1^{NE}, \dots, u_{n_S}^{NE})$ in the game $\Gamma^S(x, t, \bar{T})$ with the following set of players: coalition S (acting as one player) and the players from the set $N \setminus S$, i. e., in the game of $n_S = |N \setminus S| + 1$ players.

We construct the auxiliary game $\Gamma^S(x, t, \bar{T})$. Let the first player in this game be the player associated with coalition S for convenience, and let the other players be the ones from the set $N \setminus S$ renumbered in some way. We relabel the matrices for $N \setminus S$ as $A^S = A$, $B_i^S = B_{k_i}$, $Q_i^S = Q_{k_i}$, and $R_{i,j} = R_{k_i, k_j}$, where $i, j = \overline{2, n_S}$ and i is the new number of player k_i from $\Gamma(x, t, \bar{T})$ in the game $\Gamma^S(x, t, \bar{T})$. Some matrices for the coalition player have a block structure: $B_1^S = [B_{m_1} \dots B_{m_c}]$, $R_{1,1}^S = \text{diag}(R_{m_1, m_1}, \dots, R_{m_c, m_c})$, $R_{i,1}^S = \text{diag}(R_{k_i, m_1}, \dots, R_{k_i, m_c})$; the others are the sum of the corresponding matrices from $\Gamma(x, t, \bar{T})$: $Q_1^S = \sum_{m \in S} Q_m$, $R_{1,i}^S = \sum_{m \in S} R_{m, k_i}$, where $m_1, \dots, m_s \in S$, $i = \overline{2, n_S}$.

Thus, the motion equations of $\Gamma^S(x, t, \bar{T})$ have the form

$$\begin{aligned} \dot{x}^t(s) &= A^S x^t(s) + B_1^S u_1^t(s, x^t) + \dots + B_n^S u_n^t(s, x^t), \\ x^t(t) &= x. \end{aligned}$$

The payoff function of player $i \in N^S$ in the game $\Gamma^S(x, t, \bar{T})$ is defined as

$$K_i^{S,t}(x^t, t, \bar{T}; u^t) = \int_t^{t+\bar{T}} \left((x^t(s))' Q_i^S x^t(s) + \sum_{j=1}^n (u_j^t(s, x^t))' R_{ij}^S u_j^t(s, x^t) \right) ds,$$

where $x^t(s)$, $u^t(s, x)$ are the trajectory and strategies in the game $\Gamma^S(x, t, \bar{T})$.

Lemma 4.3.1 *For an N -person linear-quadratic differential game $\Gamma(x, t, \bar{T})$ with $Q_i^S \geq 0$ and $R_{ij}^S \geq 0$ ($i, j \in N, i \neq j$), let there exist a solution set $\{M_i^S, i \in N_S, t \geq t_0\}$ for the matrix Riccati differential equations*

$$\begin{aligned} \frac{dM_i^S(\tau)}{d\tau} + \bar{T} M_i^S(\tau) A^S + \bar{T} (A^S)' M_i^S(\tau) + Q_i^S - \\ - \bar{T}^2 M_i^S(\tau) \sum_{j \in N_S} B_j^S (R_{jj}^S)^{-1} (B_j^S)' M_j^S(\tau) = 0, \end{aligned} \quad (4.257)$$

$$M_i^S(1) = 0, \quad i \in N_S.$$

Then the characteristic function of the game $\Gamma(x, t, \bar{T})$ has form

$$\begin{aligned} V^t(S, x, \xi, t + \bar{T}) = \int_{\xi}^{t+\bar{T}} \left((x^*(s, t, x))' Q_1^S x^*(s, t, x) + \right. \\ \left. + \sum_{j=1}^{n_S} (u_j^t(s, x))' R_{1j}^S u_j^t(s, x) \right) ds, \end{aligned} \quad (4.258)$$

where

$$u_i^t(s, x) = - (R_{ii}^S)^{-1} (B_i^S)' M_i^S \left(\frac{s-t}{\bar{T}} \right) \bar{T} \Phi^S \left(\frac{s-t}{\bar{T}} \right) x, \quad (4.259)$$

$\Phi^S(\tau)$ is the solution of

$$\begin{aligned} \frac{d\Phi^S}{d\tau} = \left(A^S - \sum_{i \in N_S} B_i^S (R_{ii}^S)^{-1} (B_i^S)' \right) \Phi^S(\tau), \\ \Phi^S(0) = E, \end{aligned}$$

$x^*(s, t, x)$ is the solution of

$$\begin{aligned}\dot{x}^t(s) &= A^S x^t(s) + B_1^S u_1^t(s, x) + \dots + B_n^S u_n^t(s, x), \\ x^t(t) &= x.\end{aligned}\tag{4.260}$$

Proof We have defined the characteristic function as the total payoff of coalition S in the Nash equilibrium in the game $\Gamma^S(x, t, \bar{T})$.

Similar to the proof of Theorem 4.3.8, we use the change of variables (4.252), Theorem 6.12 from [80], and the existence of a solution for the system of differential equations (4.257). As a result, the open-loop-based Nash equilibrium strategies in the subgame $\Gamma^S(x, t, \bar{T})$ have the form (4.259).

According to building the auxiliary game $\Gamma^S(x, t, \bar{T})$, the total payoff of coalition S is calculated as $K_1^{S,t}(x^t, t, \bar{T}, u^t)$. Then the function (4.258) is characteristic, and the system dynamics evolve by (4.260). \square

4.3.4.4 Characteristic Function for Games with Continuous Updating

Suppose that the function $\tilde{V}^t(S; \tilde{x}_i^*(s), s, t + \bar{T})$, $S \subseteq N$, is continuously differentiable with respect to $s \in [t, \bar{T}]$ and integrable with respect to $t \in [t_0, +\infty)$. We define the characteristic function in the game model with continuous updating $V(S; x^*(t), t)$ as in (4.156).

Theorem 4.3.9 *For coalition S in an N -person linear-quadratic differential game with $Q_i^S \geq 0$ and $R_{ij}^S \geq 0$ ($i, j \in N, i \neq j$), let there exist a solution set $\{M_i^S, i \in N_S, t \geq t_0\}$ for the matrix Riccati differential equations (4.257). Then the characteristic function for the game with continuous updating has form*

$$V(S, x, t, T) = \int_t^T (x^*(s, t, x))' \left(Q_1^S - \bar{T}^2 \sum_{j=1}^{n_S} P_j' R_{1j}^S P_j \right) x^*(s, t, x) ds, \tag{4.261}$$

where

$$P_j = (R_{jj}^S)^{-1} (B_j^S)' M_j^S(0), \tag{4.262}$$

$x^*(s, t, x)$ is the solution of

$$\begin{aligned} \dot{x}^t(s) &= \left(A^S - \sum_{i=1}^{n_S} B_i^S P_i \right) x^t(s), \\ x^t(t) &= x. \end{aligned} \quad (4.263)$$

Proof By definition, we have the general form (4.156) for the characteristic function of the game with continuous updating. Lemma 4.3.1 gives the characteristic function (4.258) of the subgame $\Gamma(x, t, \bar{T})$. Substituting (4.258) into (4.156), we get

$$\begin{aligned} V(S, x, t, T) &= \int_t^T -\frac{d}{ds} V^\tau(S; \tilde{x}_\tau^*(s), s, \tau + \bar{T})|_{s=\tau} d\tau = \\ &= \int_t^T \left((x^*(s, t, x))' Q_1^S x^*(s, t, x) + \right. \\ &\quad \left. + \sum_{j=1}^{n_S} (u_j^s(s, x^*(s, t, x)))' R_{1j}^S u_j^s(s, x^*(s, t, x)) \right) ds, \end{aligned} \quad (4.264)$$

Taking into account $\Phi^S(0) = E$, we have

$$u_i^s(s, x) = - (R_{ii}^S)^{-1} (B_i^S)' M_i^S(0) \bar{T} x. \quad (4.265)$$

Substituting (4.265) into (4.264) with (4.262), we get

$$\begin{aligned} V(S, x, t, T) &= \int_t^T \left((x^*(s, t, x))' Q_1^S x^*(s, t, x) - \right. \\ &\quad \left. - \bar{T}^2 \sum_{j=1}^{n_S} (x^*(s, t, x))' P_j' R_{1j}^S P_j x^*(s, t, x) \right) ds. \end{aligned} \quad (4.266)$$

We write (4.266) in terms similar to (4.261). Taking into account (4.265) and (4.260), we finally describe the system dynamics as (4.263). \square

Below, any cooperative solution with continuous updating will be defined using the characteristic function with continuous updating.

4.3.4.5 Differential Game Model of Public Stock of Knowledge with Continuous Updating

4.3.4.6 Common Description

Consider a model with two individuals investing in a public stock of knowledge (see also Dockner et al. [50]). Let $x(t)$ be the stock of knowledge at a time instant t , and $u_i(t)$ be the investment of player i in public knowledge at the instant t . Assume that the stock of knowledge evolves according to the accumulation equation

$$\dot{x}(t) = -\beta x(t) + u_1(t, x_0) + u_2(t, x_0), \quad x(0) = x_0, \quad (4.267)$$

where β is the depreciation rate. Assume that each player obtains a quadratic utility from consuming the stock of knowledge, and the cost of investment increases quadratically with the investment effort. That is, the cost function of both players is given by

$$K_i(x_0, t_0, T; u) = \int_0^T (-q_i x^2(t) + r_i u_i^2(t, x_0)) dt, \quad i = 1, 2.$$

4.3.4.7 Game Model with Continuous Updating

Now consider the case of continuous updating. Suppose that at each time instant $t \in [t_0, +\infty)$, two individuals use information about the motion equations and payoff functions on the interval $[t, t + \bar{T}]$. As the current time t evolves, the interval defining the information shifts as well. The motion equations in the game model with continuous updating have the form

$$\dot{x}^t(s) = -\beta x^t(s) + u_1^t(s, x) + u_2^t(s, x), \quad x^t(t) = x, \quad t \in [t_0, +\infty).$$

4.3.4.7.1 Non-cooperative Case

The payoff function of player $i \in N$ in the game model with continuous updating is defined as

$$K_i^t(x^t, t, \bar{T}; u^t) = \int_t^{t+\bar{T}} \left(- (x^t(s))^2 q_i + (u_i^t(s, x))^2 r_i \right) ds, \quad i = 1, 2.$$

As an example, consider the symmetric case in which $r_1 = r_2 = r$ and $q_1 = q_2 = q$. According to Theorem 4.3.8 defining the form of open-loop Nash equilibrium with continuous updating, at the first step we need to solve the following differential equation:

$$\begin{cases} \dot{k}(\tau) = 2\beta\bar{T}k(\tau) + \frac{2\bar{T}k^2(\tau)}{r} + q, \\ k(1) = 0. \end{cases} \quad (4.268)$$

The solution of (4.268) is

$$k(\tau) = \frac{r(\beta - v)}{2\bar{T}} \left(\frac{2v}{v - \beta + (v + \beta)e^{2v\bar{T}(1-\tau)}} - 1 \right), \quad (4.269)$$

where $v = \sqrt{\beta^2 - \frac{2q}{r}}$. According to Theorem 4.3.8, the open-loop-based Nash equilibrium with continuous updating has the form

$$\tilde{u}_i^{NE}(t, x) = -\frac{k(0)x\bar{T}}{r}. \quad (4.270)$$

Substituting (4.269) into (4.270), we obtain

$$\tilde{u}_i^{NE}(t, x) = \frac{\beta - v}{2} \left(\frac{2v}{v - \beta + (v + \beta)e^{2v\bar{T}(1-\tau)}} - 1 \right) x. \quad (4.271)$$

Substituting (4.271) into (4.267), we obtain $\tilde{x}^{NE}(t)$ as the solution of the equation

$$\dot{\tilde{x}}^{NE}(t) = -\beta\tilde{x}^{NE}(t) + \tilde{u}_1^{NE}(t, x) + \tilde{u}_2^{NE}(t, x), \quad \tilde{x}^{NE}(0) = x_0. \quad (4.272)$$

4.3.4.7.2 Cooperative Case

The payoff function of the player in the cooperative game model with continuous updating is defined as

$$K^t(x^t, t, \bar{T}; u^t) = \sum_{i=1}^2 \int_t^{t+\bar{T}} \left(- (x^t(s))^2 q_i + (u_i^t(s, x))^2 r_i \right) ds.$$

Consider the symmetric case $r_1 = r_2 = r$, $q_1 = q_2 = q$ again. According to Theorem 2 defining the cooperative strategies with continuous updating, at the first step we need to solve the following differential equation:

$$\begin{cases} \dot{k}(\tau) = 2\beta\bar{T}k(\tau) + \frac{2\bar{T}k^2(\tau)}{r} + 2q, \\ k(1) = 0. \end{cases} \quad (4.273)$$

The solution of (4.273) is

$$k(\tau) = \frac{r(\beta - v_1)}{2\bar{T}} \left(\frac{2v_1}{v_1 - \beta + (v_1 + \beta)e^{2v_1\bar{T}(1-\tau)}} - 1 \right), \quad (4.274)$$

where $v_1 = \sqrt{\beta^2 - \frac{4q}{r}}$. According to Theorem 4.3.7, the cooperative strategies with continuous updating have the form

$$\tilde{u}_i^*(t, x) = -\frac{k(0)x\bar{T}}{r}. \quad (4.275)$$

Substituting (4.274) into (4.275), we obtain:

$$\tilde{u}_i^*(t, x) = \frac{\beta - v_1}{2} \left(\frac{2v_1}{v_1 - \beta + (v_1 + \beta)e^{2v_1\bar{T}(1-\tau)}} - 1 \right) x. \quad (4.276)$$

Substituting (4.276) into (4.267), we obtain $\tilde{x}^*(t)$ as the solution of the equation

$$\dot{\tilde{x}}^*(t) = -\beta\tilde{x}^*(t) + \tilde{u}_1^*(t, x) + \tilde{u}_2^*(t, x), \quad \tilde{x}^*(0) = x_0. \quad (4.277)$$

Taking into account (4.276) and (4.277), we get the characteristic function

$$\tilde{V}(S, x, t, T) = \sum_{i \in S} \int_t^T \left(-(\tilde{x}^*(s))^2 q_i + (\tilde{u}_i^*(s, x))^2 r_i \right) ds. \quad (4.278)$$

Substituting (4.278) into (4.164), we obtain the Shapley value with continuous updating $\widetilde{Sh}_i(\tilde{x}^*(t), t, T)$ for this example.

4.3.4.8 Game Model on Infinite Interval

Consider the classical approach to Nash equilibrium in the game on an infinite interval $[0, +\infty)$. The motion equations have the form

$$\dot{x}(t) = -\beta x(t) + u_1(t, x) + u_2(t, x), \quad x(0) = x_0. \quad (4.279)$$

The payoff function of player $i \in N$ is defined as

$$K_i(x_0; u) = \lim_{T \rightarrow \infty} \int_0^T \left(-q_i x^2(t) + r_i u_i^2(t, x) \right) dt, \quad i = 1, 2.$$

4.3.4.8.1 Non-cooperative Case

According to [224], in the symmetric case ($r_1 = r_2 = r$, $q_1 = q_2 = q$), the open-loop Nash equilibrium strategies have the form

$$u_i^*(t, x_0) = -\frac{kx_0}{r} e^{-(\beta + \frac{2k}{r})t}, \quad (4.280)$$

where k is the solution of

$$\frac{2k^2}{r} + 2\beta k + q = 0.$$

Substituting (4.280) into (4.279), we obtain $x^{NE}(t)$ as the solution of the equation

$$\dot{x}^{NE}(t) = -\beta x^{NE}(t) - \frac{2kx_0}{r} e^{-(\beta + \frac{2k}{r})t}, \quad x^{NE}(0) = x_0. \quad (4.281)$$

4.3.4.8.2 Cooperative Case

The payoff function in the cooperative case is defined as

$$K(x_0; u) = \sum_{i=1}^2 \lim_{T \rightarrow \infty} \int_0^T (-q_i x^2(t) + r_i u_i^2(t, x)) dt.$$

According to [224], in the symmetric case ($r_1 = r_2 = r$, $q_1 = q_2 = q$), the open-loop Nash equilibrium strategies have the form

$$u_i^*(t, x_0) = -\frac{kx_0}{r} e^{-(\beta + \frac{2k}{r})t}, \quad (4.282)$$

where k is the solution of

$$\frac{2k^2}{r} + 2\beta k + 2q = 0.$$

Substituting (4.282) into (4.279), we obtain $x^*(t)$ as the solution of the equation

$$\dot{x}^*(t) = -\beta x^*(t) - \frac{2kx_0}{r} e^{-(\beta + \frac{2k}{r})t}, \quad x^*(0) = x_0. \quad (4.283)$$

Taking into account (4.282) and (4.283), we get the characteristic function

$$V(S, x, t, T) = \sum_{i \in S} \int_t^T \left(-(x^*(s))^2 q_i + (u_i^*(s, x))^2 r_i \right) ds. \quad (4.284)$$

Substituting (4.284) into (4.164), we obtain the Shapley value $Sh_i(x^*(t), t, T)$ on the infinite interval.

4.3.4.9 Numerical Simulation

Consider the numerical simulation results for the game model presented above on the interval $[0, 8]$, i.e., $t_0 = 0$ and $T = 8$. At the initial instant $t_0 = 0$, the stock of knowledge is 100, i.e. $x_0 = 100$. The other parameters of the models are $\beta = 0.9$, $r = 6$, $q = -1$, and $\bar{T} = 3$. In Figure 4.25, the comparison of the Nash equilibrium

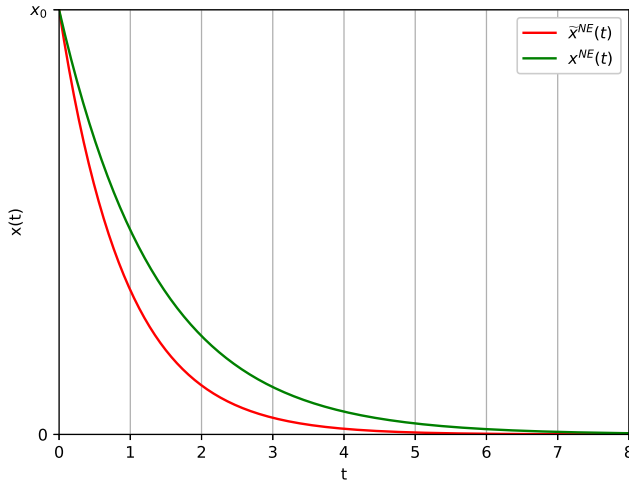


Figure 4.25: $\tilde{x}^{NE}(t)$ (4.356) — red lower line, $x^{NE}(t)$ (4.328) — green upper line.

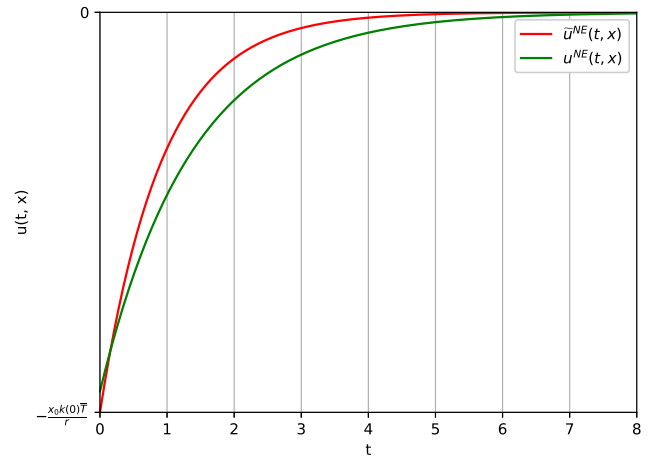


Figure 4.26: $\tilde{u}^{NE}(t)$ (4.271) — red upper line, $u^{NE}(t)$ (4.280) — green lower line.

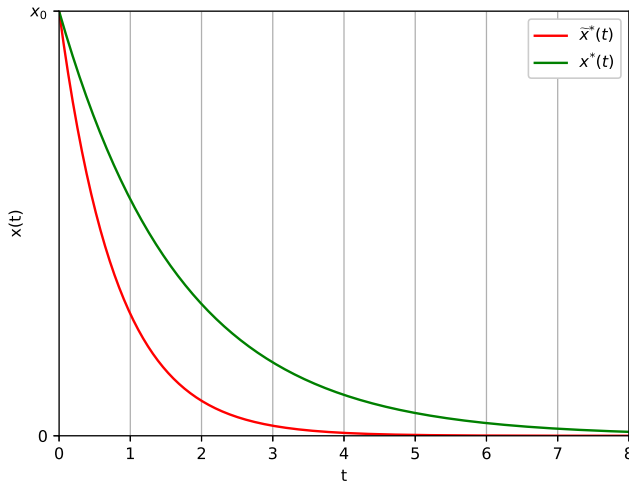


Figure 4.27: $\tilde{x}^*(t)$ (4.277) — red lower line, $x^*(t)$ (4.283) — green upper line.

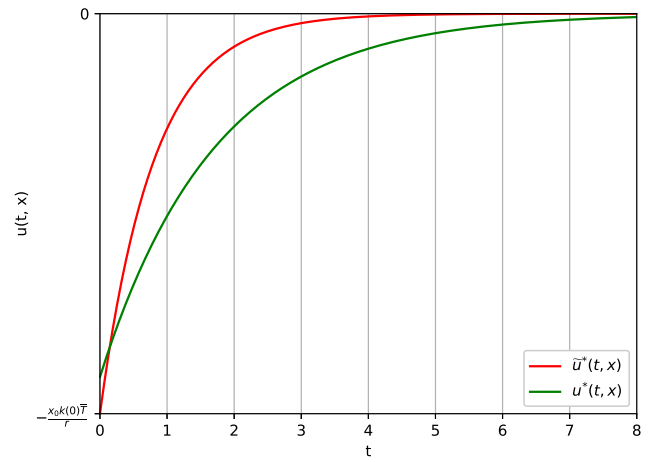


Figure 4.28: $\tilde{u}^*(t)$ (4.276) — red upper line, $u^*(t)$ (4.282) — green lower line.

with continuous updating (red lines) and Nash equilibrium in the game on an infinite interval is presented. In Figure 4.26 similar results are presented for the strategies. In Figures 4.27-4.28, similar comparisons of the cooperative solutions are presented. In Figure 4.29, the comparison of the payoff functions in the non-cooperative case is presented. In Figure 4.30, the comparison of the Shapley value in these two cases is presented.

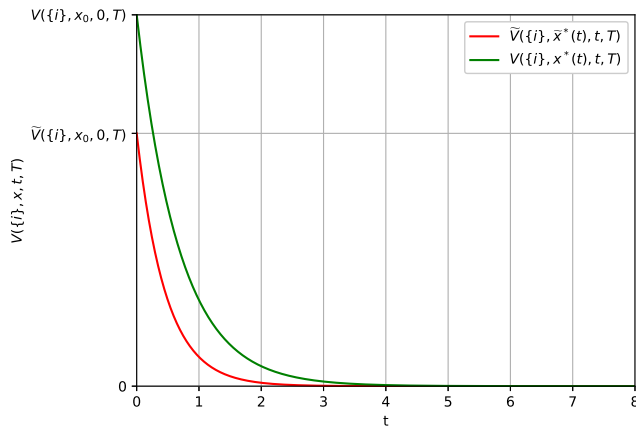


Figure 4.29: Payoff function with continuous updating $\tilde{V}(\{i\}, \tilde{x}^*(t), t, T)$ — red lower line, payoff function on infinite interval $V(\{i\}, x^*(t), t, T)$ — green upper line.

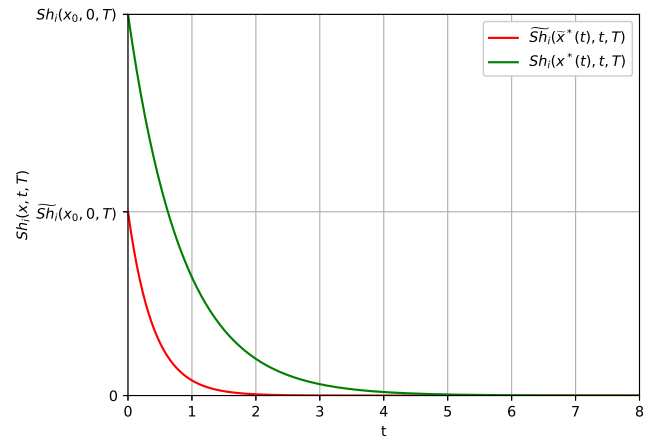


Figure 4.30: Shapley value with continuous updating $\tilde{Sh}_i(\tilde{x}^*(t), t, T)$ — red lower line, Shapley value on infinite interval $Sh_i(x^*(t), t, T)$ — green upper line.

4.4 On Class of Non-Transferable Utility Differential Games with Continuous Updating

This section considers the class of cooperative differential games with the non-transferable utility and continuous updating. The process to construct the Pareto optimal strategy with continuous updating and the Pareto trajectory is described. Another important contribution is that the property of subgame consistency is adopted for the class of games with continuous updating. A resource extraction game model is used as an example. The Pareto optimal strategies and corresponding trajectory are constructed, and the set of Pareto optimal strategies satisfying the subgame consistency property is presented. The feedback-based Pareto optimal strategies are used in this section, and therefore a simplified notation is adopted.

4.4.1 Pareto Optimal Strategies with Continuous Updating

Under continuously updated information, it is important to model the behavior of players. To do this, we use the concept of Pareto optimality. However, for the class of differential games with continuous updating, we would like to have it in the following form:

- for any fixed $t \in [t_0, +\infty)$, $u^P(t, x) = (u_1^P(t, x), \dots, u_n^P(t, x))$ coincides with the Pareto optimal strategy profile in the game (4.5), (4.6) defined on the interval $[t, t + \bar{T}]$ in the instant t .

However, direct application of the classical approaches for finding of the Pareto optimal strategies is not possible. To construct such strategies, we consider the concept of generalized Pareto optimal strategies as a principle of optimality:

$$\tilde{u}^P(t, x; s, x_t) = (\tilde{u}_i^P(t, x; s, x_t), i = 1 \dots n), \quad t \in [t_0, +\infty], \quad s \in [t, t + \bar{T}]. \quad (4.285)$$

We will use this concept for constructing the strategies $u^P(t, x)$.

Definition 4.4.1 *A strategy profile $\tilde{u}^P(t, x; s, x_t) = (\tilde{u}_1^P(t, x; s, x_t), \dots, \tilde{u}_n^P(t, x; s, x_t))$ is a generalized Pareto optimal strategy profile in the game with continuous updating if for any fixed $t \in [t_0, +\infty)$, the strategy profile $\tilde{u}^P(t, x; s, x_t)$ is Pareto optimal in the game $\Gamma(x, t, \bar{T})$.*

Definition 4.4.2 *A strategy profile $u^P(t, x)$ is called Pareto optimal with continuous updating if*

$$u^P(t, x) = \tilde{u}^P(t, x; s, x_t)|_{s=t}, \quad t \in [t_0, +\infty), \quad (4.286)$$

where $\tilde{u}^P(t, s, x)$ is the generalized Pareto optimal strategy profile in the sense of Definition (4.4.1).

The trajectory $x^*(t)$ corresponding to the Pareto optimal strategy profile with continuous updating $u^P(t, x)$ can be obtained from the system (4.8).

4.4.2 Hamilton-Jacobi-Bellman Equation with Continuous Updating

To construct the Pareto optimal strategies and corresponding trajectories with continuous updating, we need to consider the following optimization problem for every vector of weights $\alpha : \alpha_i \in (0, 1), \sum_{i=1}^n \alpha_i = 1$ [234]. Later, we will denote by $u^\alpha(t, x)$ the Pareto optimal strategy profile and by $\tilde{u}^\alpha(t, s, x)$ the generalized Pareto

optimal strategy profile:

$$\sum_{i=1}^n \alpha_i K_i^t(x, t; u^t) \longrightarrow \max_{u_1^t, \dots, u_n^t} \text{ subject to (4.5),} \quad (4.287)$$

where $t \in [t_0, +\infty]$ is the current time instant. To solve (4.287) for a fixed vector of weights α (to define strategy profile $u^\alpha(t, x)$), we need to determine the generalized Pareto optimal strategy profiles $\tilde{u}^\alpha(t, x; s, x_t)$. We will employ a modernized version of dynamic programming. Combining all possible $u^\alpha(t, x)$ for all possible weights α , we obtain the set of Pareto optimal strategy profiles with continuous updating. Classically, the weights α define the agreement between the players, but they can be reconsidered. In this section, the weights α are fixed at the beginning of the game, $\alpha(t) = \alpha$.

We denote by $W^\alpha(t; s, x)$ the Bellman function in the subgame starting at the time instant s of the game starting at the current time t :

$$W^\alpha(t; s, x) = \max_{u_1^t, \dots, u_n^t} \sum_{i=1}^n \alpha_i K_i^t(x, s, u^t) \text{ subject to (5.4).} \quad (4.288)$$

The Hamilton-Jacobi-Bellman (HJB) equation has the following form.

Theorem 4.4.1 $\tilde{u}^\alpha(t, x; s, x_t)$ is the generalized Pareto optimal strategy profile in the differential game with continuous updating if there exist functions $W^\alpha(t; s, x) : [t_0, +\infty) \times [t, t + \bar{T}] \times R \rightarrow R$, continuously differentiable with respect to s and x , that satisfy the following system of partial differential equations (4.289):

$$\begin{aligned} -W_s^\alpha(t; s, x) &= \max_{\phi_i} \left\{ \sum_{i=1}^n \alpha_i g^i(s, x, \tilde{u}_{-i}^\alpha) + W_x^{\alpha_i}(t; s, x) f(s, x, \tilde{u}_{-i}^\alpha) \right\} \\ &= \alpha_i g^i(s, x, \tilde{u}^\alpha) + W_x^\alpha(t; s, x) f(s, x, \tilde{u}^\alpha), \\ W^\alpha(t; t + \bar{T}, x) &= 0, \quad i \in N. \end{aligned} \quad (4.289)$$

where $\tilde{u}_{-i}^\alpha(\phi_i) = (\tilde{u}_1^{\alpha_1}, \dots, \phi_i, \dots, \tilde{u}_n^{\alpha_n})$.

Proof According to the definition of generalized Pareto optimal strategy profile, $\tilde{u}^\alpha(t, x; s, x_t)$ should be Pareto optimal for any fixed t .

Fixing t in the statement of Theorem 4.4.1 and, in particular, in (4.289), we obtain the classical sufficient conditions for a Pareto optimal strategy profile in the differential game with prescribed duration $[t, t + \bar{T}]$ presented in [80]. Therefore, for any fixed t , the conditions for the definition of generalized Pareto optimal strategy profile are satisfied. The theorem is proved. \square

We consider only the class of generalized Pareto optimal strategy profiles such that for the Pareto optimal strategy profile with continuous updating, the solution of the system (4.8) satisfies the conditions of existence, uniqueness, and continuability of A. F. Filippov [135]. If the generalized Pareto optimal strategy profile $\tilde{u}^\alpha(t, x; s, x_t)$ can be obtained from equations (4.289), we obtain the desired strategy profile $u^\alpha(t, x)$ using the procedure (4.286).

4.4.3 Subgame Consistency with Continuous Updating

Under cooperation with non-transferable payoffs, the players negotiate to establish an agreement (optimality principle) to play the cooperative game. In particular, the optimality principle has to satisfy group optimality (i) and individual rationality (ii) along the chosen trajectory $x^*(t)$ (in our case, the Pareto optimal trajectory). Subgame consistency requires that an extension of the solution to a later starting time and any possible state brought about by the prior optimal behaviour of the players would remain optimal as well. Both group optimality and individual rationality are required. Group optimality requires the players to seek cooperative strategies (controls) yielding a Pareto optimal solution. The solution has to satisfy individual rationality: all players would obtain a higher payoff in the cooperative case compared to the individual behavior.

According the procedure (4.286), the Pareto optimal strategies and trajectories with continuous updating satisfy the group optimality property. But the individual rationality property is not always satisfied [234]. For the class of games with continuous updating, the individual rational property has the following form:

$K_i^{t, NE}(x^*(t), t; \tilde{u}^{NE}(t, x; s, x_t)) \leq K_i^{t, \alpha}(x^*(t), t; \tilde{u}^\alpha(t, x; s, x_t))$, for $\forall i$ and $\forall t$, where $K_i^{t, NE}(x^*(t), t; \tilde{u}^\alpha(t, x; s, x_t))$ is the individual payoff of player i in the Nash equilibrium in the game defined on the interval $[t, t + \bar{T}]$ starting along the Pareto

optimal trajectory $x^*(t)$, and $K_i^{t,\alpha}(x^*(t), t; \tilde{u}^\alpha(t, x; s, x_t))$ is the individual payoff under cooperation (4.286) in the game starting along the Pareto optimal trajectory $x^*(t)$. Subgame consistency can be formally stated as follows.

Definition 4.4.3 *A Pareto optimal strategy profile $u^\alpha(t, x)$ is called subgame consistent if the corresponding generalized Pareto optimal strategy profile $\tilde{u}^\alpha(t, x; s, x_t)$ is such that the following conditions are satisfied:*

- (i) *Group optimality: $(K_i^{t,\alpha}(x^*(t), t; \tilde{u}^\alpha(t, x; s, x_t)), i \in N)$ is Pareto optimal;*
- (ii) *Individual rationality:*

$$K_i^{t,NE}(x^*(t), t; \tilde{u}^{NE}(t, x; s, x_t)) \leq K_i^{t,\alpha}(x^*(t), t; \tilde{u}^\alpha(t, x; s, x_t)), \quad \forall i \text{ and } \forall t.$$

Suppose that there exists a set A of weights α such that conditions (i) and (ii) are satisfied. The set of Pareto optimal strategies $u^\alpha(t, x)$, where $\alpha \in A$, will be called the subgame consistent cooperative solution with continuous updating.

4.4.4 Differential Game of Non-renewable Resource Extraction

As an illustrative example, consider the differential game model with continuous updating for extracting a nonrenewable resource (see [50]).

4.4.4.0.1 Original Game Model

We denote by $x(t)$ the state vector indicating the resource stock at a time instant t . Let $u_i(t, x)$ be the extraction rate of player i at a time instant t if the resource stock is equal to x . Assume that $u_i(t, x) \geq 0$, and if $x(t) = 0$, then the only feasible rate of extraction is $u_i(t, x) = 0$.

The dynamics of the stock are given by the equation

$$\dot{x}(t) = - \sum_{i=1}^n b_i u_i(t, x), \quad x(t_0) = x_0, \quad (4.290)$$

where $b_i > 0$ for all $i = 1, \dots, n$, and $x_0 > 0$. The payoff function of player i has

the form

$$K_i(x_0, T - t_0) = \int_{t_0}^T \ln u_i(t, x) dt, \quad i \in N. \quad (4.291)$$

4.4.4.1 Pareto Optimal Strategies with Continuous Updating

According to Section 4.4.1, to determine the Pareto optimal strategies in the game with continuous updating, we consider the family of auxiliary subgames $\Gamma(x, t, t + \bar{T})$ with duration \bar{T} starting at a time instant t from a state x . To define the Pareto optimal strategies $\tilde{u}^\alpha(t, x; s, x_t)$ in the auxiliary subgame $\Gamma(x, t, t + \bar{T})$, we use the dynamic programming technique. We denote by $W^\alpha(t; s, x)$ the Bellman function in the subgame for the current time instant t starting at s :

$$W^\alpha(t; s, x) = \max_{u_1^t, \dots, u_n^t} \sum_{i=1}^n \alpha_i K_i(x, s; u_t) \quad (4.292)$$

subject to : $\dot{x}_t(s) = - \sum_{i=1}^n b_i u_i^t(s, x_t), \quad x_t(t) = x.$

The Hamilton-Jacobi-Bellman (HJB) equation has the following form:

$$-\frac{\partial W^\alpha(t; s, x)}{\partial s} = \max_{u_1^t, \dots, u_n^t} \left\{ \sum_{i=1}^n \alpha_i \ln u_i^t(s, x) - \frac{\partial W^\alpha(t; s, x)}{\partial x} \sum_{i=1}^n b_i u_i^t(s, x) \right\},$$

$$\lim_{s \rightarrow t + \bar{T}} W^\alpha(t; s, x) = 0. \quad (4.293)$$

The solution of (4.293) will be found in the form $W^\alpha(t; s, x) = A(t, s) \ln x + B(t, s)$. The partial derivatives are given by

$$W_s^\alpha(t; s, x) = \dot{A}(t, s) \ln x + \dot{B}(t, s),$$

$$W_x^\alpha(t; s, x) = \frac{A(t, s)}{x}, \quad (4.294)$$

where $\dot{A}(t, s)$ and $\dot{B}(t, s)$ are the derivatives with respect to s . Maximizing the expression in the right-hand side of (4.293) and substituting the result into (4.294),

we obtain

$$\tilde{u}_i^\alpha(t, s, x) = \frac{x\alpha_i}{b_i A(t, s)}, \quad \sum_{i=1}^n \alpha_i = 1. \quad (4.295)$$

Substituting (4.294) and (4.295) into (4.293), we obtain the following system of differential equations:

$$\dot{A}(t, s) \ln x + \dot{B}(t, s) = -\ln x + \ln A(t, s) + \ln b^N - \sum_{i=1}^n \alpha_i \ln \alpha_i + n, \quad (4.296)$$

$$\lim_{s \rightarrow t+\bar{T}} A(t, s) \ln x + B(t, s) = 0,$$

where $b^N = \prod_{i=1}^n b_i$. Then we get:

$$\dot{A}(t, s) = -1, \quad \lim_{s \rightarrow t+\bar{T}} A(t, s) = 0, \quad (4.297)$$

$$\dot{B}(t, s) = \ln A(t, s) + \ln b^N - \sum_{i=1}^n \alpha_i \ln \alpha_i + n, \quad \lim_{s \rightarrow t+\bar{T}} B(t, s) = 0.$$

The solution of (4.296) has the form

$$B(s, t) = -(t + \bar{T} - s) \left(-\sum_{i=1}^n \alpha_i \ln \alpha_i + \ln b^N + n + \ln n(t + \bar{T} - s) \right), \quad (4.298)$$

$$A(s, t) = n(t + \bar{T} - s), \quad s \in [t, t + \bar{T}), \quad t \in [t_0, +\infty). \quad (4.299)$$

Finally, we obtain the Pareto optimal strategies in the auxiliary subgame $\Gamma(x, t, t + \bar{T})$:

$$\tilde{u}_i^\alpha(t, x; s, x_t) = \frac{x\alpha_i}{b_i n(t + \bar{T} - s)}, \quad s \in [t, t + \bar{T}), \quad t \in [t_0, +\infty), \quad \sum_{i=1}^n \alpha_i = 1. \quad (4.300)$$

In addition,

$$\begin{aligned} W^\alpha(t; s, x) &= \\ &= (t + \bar{T} - s) \left(\ln b^N + \sum_{i=1}^n \alpha_i \ln \alpha_i - n - n \ln \frac{x}{n(t + \bar{T} - s)} \right), \quad s \in [t, t + \bar{T}). \end{aligned} \quad (4.301)$$

Following the procedure (4.286), we construct the Pareto optimal strategies with continuous updating

$$u^\alpha(t, x) = \tilde{u}^\alpha(t, x; s, x_t)|_{s=t} = \left(\frac{x\alpha_i}{b_i\bar{T}}, i = 1 \dots n \right), s \in [t, t + \bar{T}], \sum_{i=1}^n \alpha_i = 1. \quad (4.302)$$

Substituting (4.302) into (4.290), we derive the Pareto optimal trajectory $x^*(t)$ with continuous updating:

$$x^*(t) = x_0 \exp \frac{-t}{\bar{T}}. \quad (4.303)$$

As we can see, the Pareto optimal strategy profile (4.302) depends on the weight α , but the Pareto optimal trajectory (4.303) does not. The reason is that the sum of weights α equals 1.

4.4.4.2 Subgame Consistency of Pareto Optimal Solution

In this section, we want to construct the subgame consistent Pareto optimal strategy profile $u^\alpha(t, x)$ using the conditions presented in Definition 4.4.3. For this purpose, we need to calculate $K_i^{t,\alpha}(x^*(t), t; \tilde{u}^\alpha(t, x; s, x_t))$ and $K_i^{t,NE}(x^*(t), t; \tilde{u}^{NE}(t, x; s, x_t))$. First, we need to calculate $\tilde{x}_t^\alpha(s)$ and $\tilde{x}_t^{NE}(s)$, which are the imaginary or expected Pareto optimal and Nash equilibrium trajectories in the game defined on the interval $[t, t + \bar{T}]$ starting along the Pareto optimal trajectory $x^*(t)$. Taking into account $x_0 = x^*(t)$, $x(t) = x_t(s)$, we substitute the Pareto optimal strategy profile (4.300) into (4.290) to obtain

$$\tilde{x}_t^\alpha(s) = x^*(t) \cdot \frac{t + \bar{T} - s}{\bar{T}} \cdot \exp^{\frac{1}{n}}, \quad s \in [t, t + \bar{T}]. \quad (4.304)$$

Then it is possible to find the Pareto optimal strategy profile $\tilde{u}^\alpha(t, x; s, \tilde{x}_t^\alpha(s))$ along the trajectory (4.304) on the interval $[t, t + \bar{T}]$:

$$\tilde{u}^\alpha(t, x; s, \tilde{x}_t^\alpha(s)) = \frac{x^*(t) \cdot \alpha_i \exp^{\frac{1}{n}}}{b_i n \bar{T}}, \quad s \in [t, t + \bar{T}], \quad \sum_{i=1}^n \alpha_i = 1. \quad (4.305)$$

Now we calculate the Pareto optimal payoff of player i at the time instant t with continuous updating by substituting (4.305) into the payoff function (4.291) with

continuous updating (with the limits t and $t + \bar{T}$):

$$K_i^{t,\alpha}(x^*(t), t; \tilde{u}^\alpha(t, x; s, x_t)) = \int_t^{t+\bar{T}} \ln \frac{x^*(t) \cdot \alpha_i \exp^{\frac{1}{n}}}{b_i n \bar{T}} ds. \quad (4.306)$$

Using the definition from the paper [78], we can find the Nash equilibrium strategies with continuous updating for this specific game model. The payoff function of player i at a time instant s in the Nash equilibrium in the game defined on the interval $[t, t + \bar{T}]$:

$$\begin{aligned} K_i^{t,NE}(x^*(t), t; \tilde{u}^{NE}(t, x; s, x_t)) &= \int_t^{t+\bar{T}} \ln \tilde{u}^\alpha(t, x; s, \tilde{x}_t^{NE}(s)) ds = \\ &= \int_t^{t+\bar{T}} \ln \frac{x^*(t) \cdot (t+\bar{T}-s)}{b_i (t+\bar{T}-s)} ds = \int_t^{t+\bar{T}} \ln \frac{x^*(t)}{b_i} ds. \end{aligned} \quad (4.307)$$

Let us define the set of α satisfying (4.308), which therefore guarantees subgame consistency of the Pareto optimal strategies in this game model with continuous updating:

$$K_i^{t,\alpha}(x^*(t), t; \tilde{u}^\alpha(t, x; s, x_t)) \geq K_i^{t,NE}(x^*(t), t; \tilde{u}^{NE}(t, x; s, x_t)), \quad \sum_{i=1}^n \alpha_i = 1. \quad (4.308)$$

The solution is

$$\begin{aligned} \ln \frac{x^*(t) \cdot \alpha_i \exp^{\frac{1}{n}}}{b_i n \bar{T}} &\geq \ln \frac{x^*(t)}{b_i}, \\ \ln \alpha_i &\geq \ln \frac{x^*(t)}{b_i} - \ln \frac{x^*(t) \cdot \exp^{\frac{1}{n}}}{b_i n \bar{T}}, \\ \alpha_i &\geq \frac{n \bar{T}}{\exp^{\frac{1}{n}}}, \quad \sum_{i=1}^n \alpha_i = 1. \end{aligned} \quad (4.309)$$

This solution shows under which conditions the Pareto optimal strategy profile with continuous updating satisfies the subgame consistency property.

Consider a numerical example with $n = 2$, $x_0 = 5$, $b_i = 2$, $\bar{T} = 0.5$, $\alpha_i \in \{0.1, 0.2, \dots, 1\}$. In Figure 4.31, we compare the trajectories in the original game and the game with continuous updating. In Figure 4.32, we see the difference between the Pareto optimal strategies in the original game and the game with continuous

updating for $\alpha_1 = 0.664$, $\alpha_2 = 0.336$. Figure 4.33 shows the Pareto optimal strategy of player i for different α . In Figure 4.34, the payoff functions (4.306) for different weights α and the Nash equilibrium payoff (4.307) are presented. Clearly, the red line (4.307) divides the blue lines (4.306) in the figure into two parts. The set of blue lines above the red line is what we want to find (4.309). The blue lines (4.306) below the red line (4.307) do not satisfy the subgame consistency property.

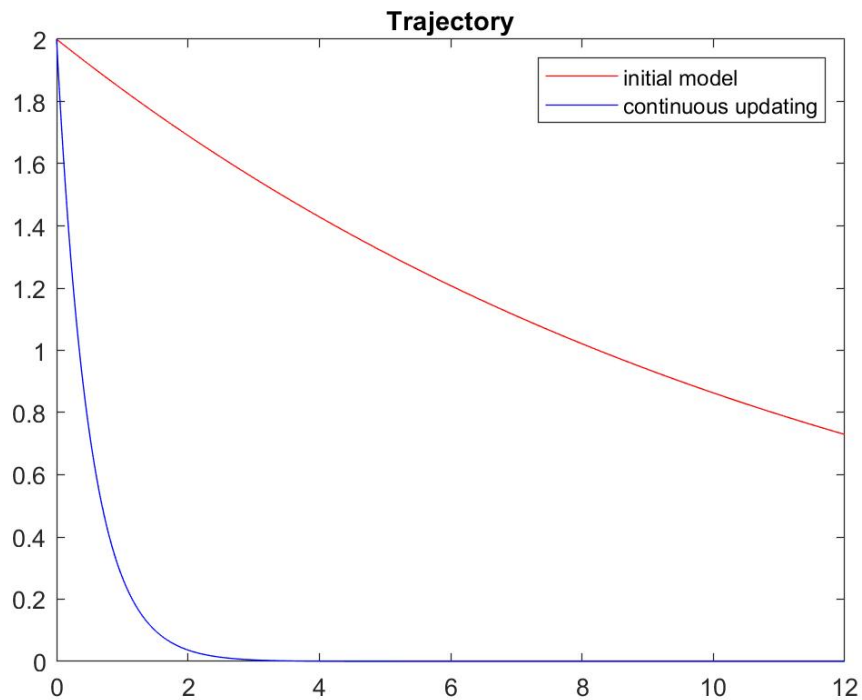


Figure 4.31: Pareto optimal trajectory with continuous updating (blue line) and Pareto optimal trajectory in original game (red line).

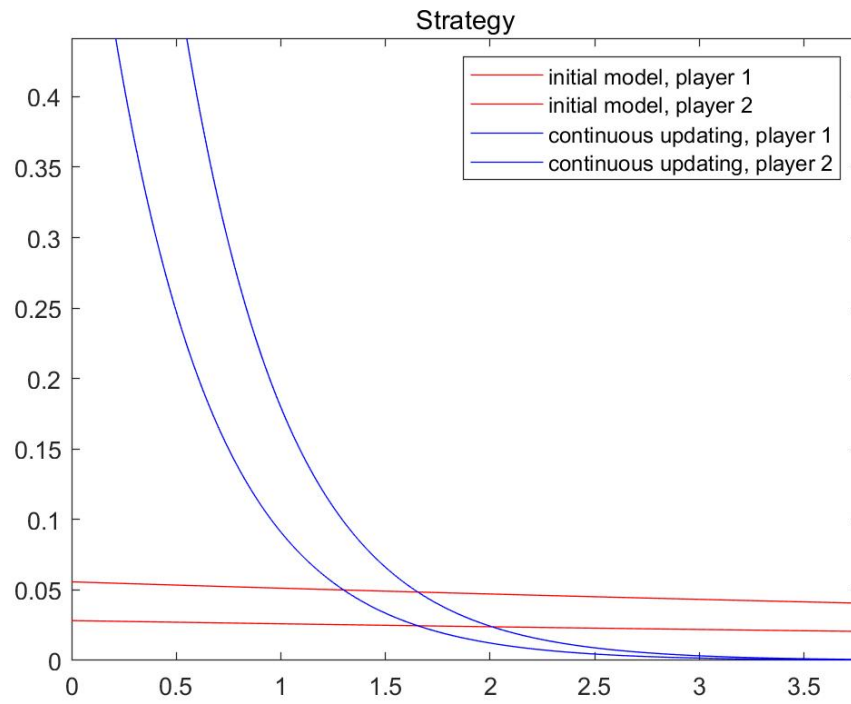


Figure 4.32: Pareto optimal strategies of players 1 and 2 in original game and game model with continuous updating for $\alpha_1 = 0.664$ and $\alpha_2 = 0.336$.

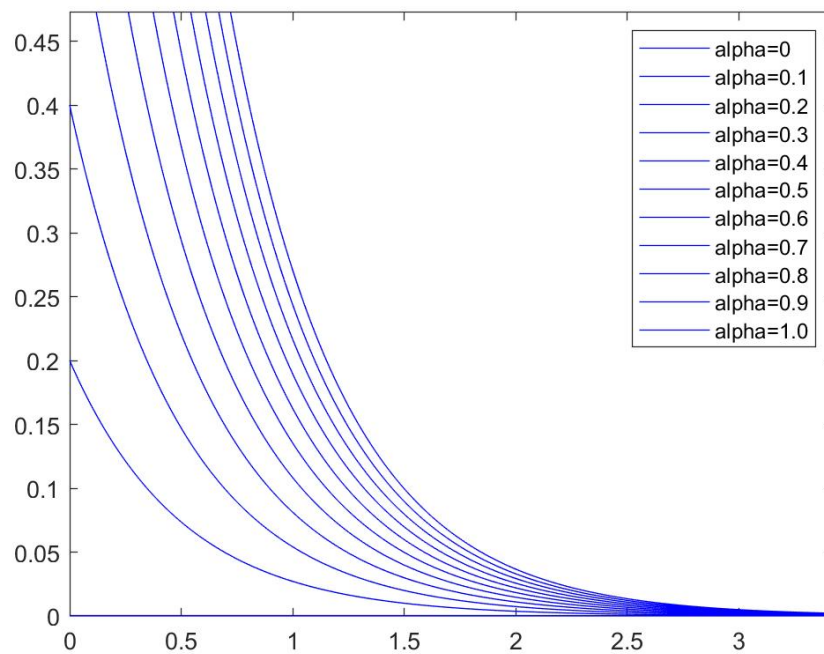


Figure 4.33: Pareto optimal strategy of player i with continuous updating for different weights $\alpha_i = (0.1, 0.2, \dots, 1)$.

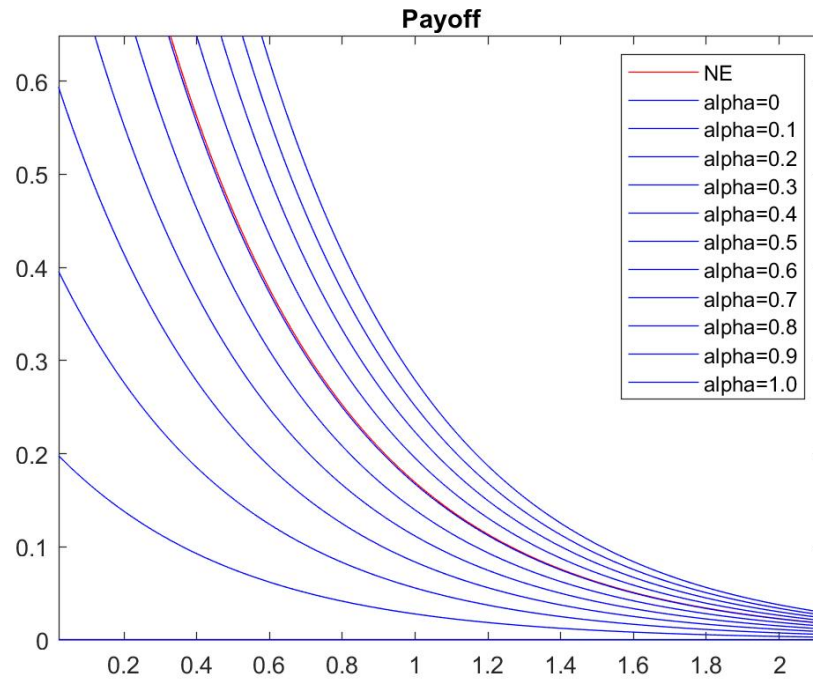


Figure 4.34: Payoff function (4.306) of player i corresponding to Pareto optimal strategy profile (blue lines) and payoff function (4.307) of player i corresponding to Nash equilibrium (red lines) with continuous updating.

4.5 Differential Game Model of Resource Extraction with Continuous and Dynamic Updating

This section describes the model of resource extraction for the case of dynamic and continuous updating. It is proved that the optimal control (cooperative strategies) and feedback Nash equilibrium strategies uniformly converge to the corresponding strategies in the game model with continuous updating as the number of updating instants tends to infinity. Similar results are presented for the optimal trajectory (cooperative trajectory), the equilibrium trajectory, and the corresponding payoffs.

Below we will use only the feedback-based strategies to construct the cooperative strategies with continuous updating $u_{fb}^*(t, x)$ and the Nash equilibrium strategies $u_{fb}^{NE}(t, x)$ to construct the characteristic function with continuous updating. Therefore, we will simplify the notation and use $u^*(t, x)$ and $u^{NE}(t, x)$, respectively.

4.5.1 Original Game Model of Resource Extraction

Consider the classical game-theoretical model of non-renewable resource extraction with symmetric players [168]. Here the amount of resource directly depends on the rates of extraction, which are selected by the companies (players). The game involves n symmetric players (which form the set N) with utility functions depending on their extraction rates $h_i(t, x, u_i) = \log u_i$. The game starts at the instant t_0 and terminates at T , i.e., the game is defined on the interval $[t_0, T]$. The amount of resource at the beginning of the game is x_0 .

We denote by $x(t) \in R^1$ the amount of resource available to the players at the instant t and by $u_i(t, x)$ the strategy of player $i \in N$, which is the resource extraction rate defined for any t and x . We will seek for the feedback strategies, assuming that $\forall t, u_i(t, x) \geq 0$ and $x(t) = 0$ imply $u_i(t, x) = 0$. The amount of resource $x(t)$ is a function of time t that depends on the rates of extraction (strategies) of players $u_i(t, x)$:

$$\begin{aligned} \dot{x} &= - \sum_{i=1}^n u_i(t, x), \\ x(t_0) &= x_0. \end{aligned} \tag{4.310}$$

The payoff function of player $i \in N$ has the form

$$K_i(x_0, T - t_0) = \int_{t_0}^T \log(u_i(\tau, x)) d\tau, i \in N. \tag{4.311}$$

Assume that for any n -tuple of strategies $u_1(\cdot), \dots, u_n(\cdot)$, the conditions of existence, uniqueness, and continuability of solutions (4.310) are satisfied [160]. Taking into account the symmetry of players, we put $u(t, x) = u_i(t, x)$ for each $i \in N$.

In the next section, the optimal strategies (controls) are calculated for two basic classes of differential games, i.e., cooperative and noncooperative differential games.

4.5.1.1 Cooperative Differential Game Model

Consider the cooperative version of the non-renewable resource extraction game [168]. The players unite in the grand coalition $S = N$ and, acting as one player, maximize the joint payoff. The corresponding optimal control problem is formulated in the following way:

$$\sum_{i=1}^n K_i(x_0, T - t_0) = n \int_{t_0}^T \log(u(\tau, x)) d\tau \rightarrow \max_u \quad (4.312)$$

subject to

$$\begin{aligned} \dot{x} &= -nu(t, x), \\ x(t_0) &= x_0 > 0, \\ u(t, x) &\geq 0. \end{aligned} \quad (4.313)$$

To solve the optimization problem (4.312), (4.313), we use the method of dynamic programming proposed in this Thesis. We define the Bellman function as the maximum value of functional (4.311) in the subgame $\Gamma(x, T - t)$ starting at the instant t from the position x :

$$V(t, x) = \max_u \left\{ \sum_{i=1}^n K_i(x, T - t) \right\} = \max_u \left\{ n \int_t^T \log u(\tau, x) d\tau \right\}. \quad (4.314)$$

The following result was established in the paper [2]: if there exists a continuously differentiable function $V(t, x)$ satisfying the Hamilton-Jacobi-Bellman equation

$$\begin{aligned} -V_t(t, x) &= \max_u \{ n \log u - nuV_x(t, x) \}, \\ V(T, x) &= 0, \end{aligned} \quad (4.315)$$

then the control $u^*(t, x)$ determined by maximizing the right-hand side of (4.315) is optimal in the control problem (4.312), (4.313).

From the first-order optimality condition for (4.315) we obtain

$$u^* = \frac{1}{V_x(t, x)}.$$

Substituting the result into (4.315) yields

$$\begin{aligned} V_t(t, x) &= n \log V_x(t, x) + n, \\ V(T, x) &= 0. \end{aligned}$$

We define the Bellman function in the form

$$V(t, x) = A(t) \log x + B(t).$$

Substituting it into (4.345), we obtain

$$\begin{aligned} \dot{A}(t) \log x + \dot{B}(t) &= n \log A(t) - n \log x + n, \\ A(T) = B(T) &= 0. \end{aligned} \tag{4.316}$$

The solution of (4.316) are the functions

$$\begin{aligned} A(t) &= n(T - t), \\ B(t) &= -n(T - t) \log n(T - t). \end{aligned} \tag{4.317}$$

Finally, we obtain the following expressions: for the Bellman function,

$$V(t, x) = n(T - t) \log \frac{x}{n(T - t)}, t \in [t_0, T]; \tag{4.318}$$

for the optimal control,

$$u^*(t, x) = \frac{1}{V_x(t, x)} = \frac{x}{n(T - t)}, t \in [t_0, T]. \tag{4.319}$$

Substituting the optimal control into the motion equation (4.313), we obtain a differential equation for the trajectory corresponding to the optimal control:

$$\begin{aligned} \dot{x}(t) &= -\frac{x(t)}{T - t}, \\ x(t_0) &= x_0. \end{aligned} \tag{4.320}$$

The solution of the Cauchy problem (4.320) has the form

$$x^*(t) = x_0 \frac{T-t}{T-t_0}, \quad t \in [t_0, T]. \quad (4.321)$$

The trajectory $x^*(t)$ and strategy (control) $u^*(t, x)$ are called cooperative.

The cooperative strategies along the cooperative trajectory have the form

$$u^*(t, x^*) = \frac{x_0}{n(T-t)}, \quad t \in [t_0, T]. \quad (4.322)$$

In Figure 4.35, green line represents the cooperative trajectory $x^*(t)$ (4.321); in Figure 4.36, green line represents the corresponding optimal strategies (4.322) in the game of non-renewable resource extraction $\Gamma(x_0, T - t_0)$.

4.5.1.2 Noncooperative Differential Game Model

Consider the case when each player decides on the amount of extracted resources $u_i(t, x)$ individually. As a principle of optimality, we use the feedback Nash equilibrium. We denote by

$$V_i(t, x) = \int_t^T \log(u_i^{NE}(t, x)) d\tau \quad (4.323)$$

the payoff of player i in the feedback Nash equilibrium $u^{NE}(t, x) = (u_1^{NE}(t, x), \dots, u_n^{NE}(t, x))$, in the subgame $\Gamma(x, T - t)$ starting at a time instant t from a position x , where

$$\begin{aligned} \dot{x}(t) &= - \sum_{i=1}^n u_i^{NE}(t, x), \\ x(t) &= x. \end{aligned} \quad (4.324)$$

To find the equilibrium strategies $u_1^{NE}(t, x), \dots, u_n^{NE}(t, x)$, we also use the dynamic programming technique described in [80]. The corresponding system of

Hamilton-Jacobi-Bellman equations has the form

$$-V_t^i(t, x) = \max_{u_i} \left\{ \log u_i - V_x^i(t, x) \left(u_i + \sum_{k=1, k \neq i}^n u_k^{NE} \right) \right\}, \quad (4.325)$$

$$V_i(T, x) = 0, \quad i \in N.$$

According to [80], if there exist continuously differentiable functions $V_i(t, x)$ satisfying (4.325), then $u_i^{NE}(t, x)$ is a feedback Nash equilibrium. We will seek for a Bellman function in the form

$$V_i(t, x) = A_i(t) \log x + B_i(t), \quad i \in N.$$

Solving (4.325), we obtain the Bellman function

$$V_i(t, x) = (T - t) \left(\log \frac{x}{T - t} - n + 1 \right), \quad t \in [t_0, T], \quad i \in N. \quad (4.326)$$

Thus, the feedback Nash equilibrium is

$$u_i^{NE}(t, x) = \frac{x}{T - t}, \quad t \in [t_0, T], \quad i \in N, \quad (4.327)$$

and the corresponding trajectory $x^{NE}(t)$ is

$$x^{NE}(t) = x_0 \left(\frac{T - t}{T - t_0} \right)^n, \quad t \in [t_0, T]. \quad (4.328)$$

The equilibrium strategies $(u_1^{NE}(t, x), \dots, u_n^{NE}(t, x))$ along the trajectory $x^{NE}(t)$ are

$$u_i^{NE}(t, x^{NE}) = x_0 \frac{(T - t)^{n-1}}{(T - t_0)^n}, \quad t \in [t_0, T], \quad i \in N. \quad (4.329)$$

In Figure 4.37, green line represents the feedback Nash equilibrium strategies $u^{NE}(t, x)$ (4.329); in Figure 4.38, green line represents the corresponding equilibrium trajectory $x^{NE}(t)$ (4.328).

4.5.2 Game Model with Dynamic Updating

To model the behavior of players under dynamically updated information, the interval $[t_0, T]$ is split into l subintervals with length $\Delta t = \frac{T-t_0}{l}$, and the behavior of players on each subinterval $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, l$, is described by the notion of truncated subgame.

4.5.2.1 Cooperative Game Model with Dynamic Updating

First, consider the cooperative version of the limited resource extraction game with dynamic updating. Since the structure of the truncated subgame on each interval $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ corresponds to the original game defined on the interval $[t_0, T]$, the solution of each subgame j is defined in the similar way. The main difference consists in the model parameters, namely, $t_0 + j\Delta t$ is the initial instant of the subgame, $t_0 + j\Delta t + \bar{T}$ is the terminal instant of the subgame, and x_0^j is the amount of resource at the beginning of the truncated subgame. Thus, the cooperative strategies $u_j^*(x, t)$ in each subgame $\bar{\Gamma}_j(x_0^j, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ have the form

$$u_j^*(t, x) = \frac{x}{n(t_0 + j\Delta t + \bar{T} - t)}, \quad t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}], \quad (4.330)$$

and the corresponding cooperative trajectory $x_j^*(t)$ is

$$x_j^*(t) = x_0^j \frac{t_0 + j\Delta t + \bar{T} - t}{\bar{T}}, \quad t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]. \quad (4.331)$$

Note that x_0^j depends on the value of trajectory in the previous truncated subgame $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$:

$$x_0^j = x_{j-1}(t_0 + j\Delta t), \quad \text{where } x_0^0 = x_0.$$

Theorem 4.5.1 *The resulting cooperative trajectory $\hat{x}^*(t)$ in the game model with*

dynamic updating has the following form:

$$\hat{x}^*(t) = x_0 \left(1 - \frac{\Delta t}{\bar{T}}\right)^j \frac{t_0 + j\Delta t + \bar{T} - t}{\bar{T}}, \quad t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}], \quad j = 0, \dots, l. \quad (4.332)$$

Proof To derive the explicit formula for the resulting trajectory using (4.331), we need defining the parameter x_0^j for any truncated subgame $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$.

Consider the trajectory sections with numbers $j - 1$ and j :

$$\begin{aligned} x_{j-1}^*(t) &= x_0^{j-1} \frac{t_0 + (j-1)\Delta t + \bar{T} - t}{\bar{T}}, \\ x_j^*(t) &= x_0^j \frac{t_0 + j\Delta t + \bar{T} - t}{\bar{T}}, \end{aligned}$$

where $x_0^j = x_{j-1}^*(t_0 + j\Delta t)$. Then the following relation holds:

$$\begin{aligned} x_0^j &= x_0^{j-1} \left(1 - \frac{\Delta t}{\bar{T}}\right), \\ x_0^j &= x_0^{j-2} \left(1 - \frac{\Delta t}{\bar{T}}\right)^2, \\ &\dots, \\ x_0^j &= x_0^0 \left(1 - \frac{\Delta t}{\bar{T}}\right)^j. \end{aligned}$$

Taking into account $x_0^0 = x_0$, we finally obtain (4.333):

$$x_0^j = x_0 \left(1 - \frac{\Delta t}{\bar{T}}\right)^j. \quad (4.333)$$

Substituting (4.333) into (4.331), we arrive in formula (4.332):

$$\hat{x}^*(t) = x_0 \left(1 - \frac{\Delta t}{\bar{T}}\right)^j \frac{t_0 + j\Delta t + \bar{T} - t}{\bar{T}}, \quad t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}], \quad j = 0, \dots, l.$$

□

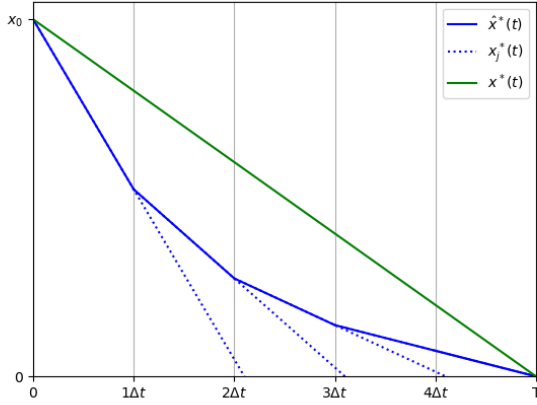


Figure 4.35: Resulting cooperative trajectory $\hat{x}^*(t)$ (4.332) (blue line) and cooperative trajectory in original game $x^*(t)$ (4.321) (green line).

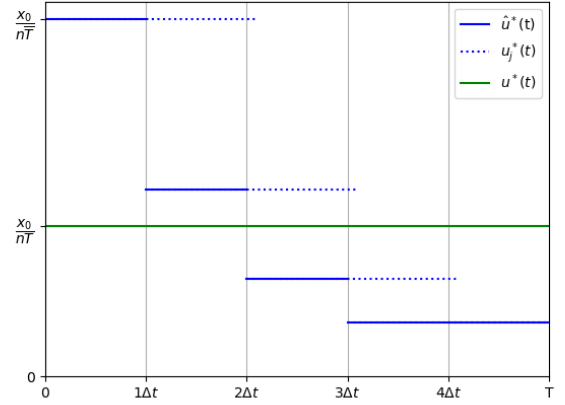


Figure 4.36: Resulting cooperative strategies $\hat{u}^*(t, \hat{x}^*)$ (4.334) (blue line) and cooperative strategies in original game $u^*(t, x^*)$ (4.322) (green line).

As a result, the cooperative strategies along the trajectory (4.332) have the form

$$u_j^*(t, x_j^*) = \frac{x_0}{n\bar{T}} \left(1 - \frac{\Delta t}{\bar{T}}\right)^j, t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]. \quad (4.334)$$

We denote the resulting strategies $\hat{u}(t, x)$ (4.110) by $\hat{u}^*(t, x)$, and the corresponding resulting cooperative trajectory by $\hat{x}^*(t)$.

Figure 4.35 (Figure 4.36) presents a comparison of the resulting cooperative trajectory $\hat{x}^*(t)$ (4.332) (the resulting cooperative strategies $\hat{u}^*(t)$ (4.334)) in the game with dynamic updating and the cooperative trajectory $\hat{x}^*(t)$ (4.321) ($u^*(t)$ (4.322)) in the original game with prescribed duration.

□

4.5.2.2 Noncooperative Game Model with Dynamic Updating

Consider now the noncooperative game model with dynamic updating. Assume that the players act individually. As in the original game, we use feedback Nash equilibrium as the optimality principle. Performing calculations similar to those in

Section 4.5.1.2, we obtain:

$$u_j^{NE}(t, x) = \frac{x}{\bar{T} + j\Delta t + t_0 - t}, \quad t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}], j = 0, \dots, l, \quad (4.335)$$

$$x_j^{NE}(t) = x_0^j \left(\frac{\bar{T} + j\Delta t + t_0 - t}{\bar{T}} \right)^n, \quad t \in [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}], j = 0, \dots, l. \quad (4.336)$$

Theorem 4.5.2 *The resulting equilibrium trajectory $\hat{x}^{NE}(t)$ in the game model with dynamic updating has the following form:*

$$\hat{x}^{NE}(t) = x_0 \left(1 - \frac{\Delta t}{\bar{T}} \right)^{jn} \left(\frac{t_0 + j\Delta t + \bar{T} - t}{\bar{T}} \right)^n, \quad t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]. \quad (4.337)$$

Proof The proof is similar to the proof of Theorem 4.5.1. □

The feedback Nash equilibrium strategies along the equilibrium trajectory $\hat{x}^{NE}(t)$ on the interval $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, l$ have the form

$$u_j^{NE}(t, x_j^{NE}(t)) = x_0 \left(1 - \frac{\Delta t}{\bar{T}} \right)^{jn} \frac{(\bar{T} + j\Delta t + t_0 - t)^{n-1}}{\bar{T}^n}. \quad (4.338)$$

4.5.3 Game Model with Continuous Updating

Consider the case when the length of Δt interval between the updating instants is negligibly small. In other words, information updates continuously in time ($\Delta t \rightarrow 0$ or $l \rightarrow \infty$).

At the first step (see the procedure in the paper [78]), we consider the following game model $\Gamma(x, t, t + \bar{T})$ with an initial instant t and a terminal instant $t + \bar{T}$:

$$K_i^t(x, t, t + \bar{T}; u_1^t, \dots, u_n^t) = \int_t^{t+\bar{T}} \log(u_i^t(\tau, x_t)) d\tau, \quad (4.339)$$

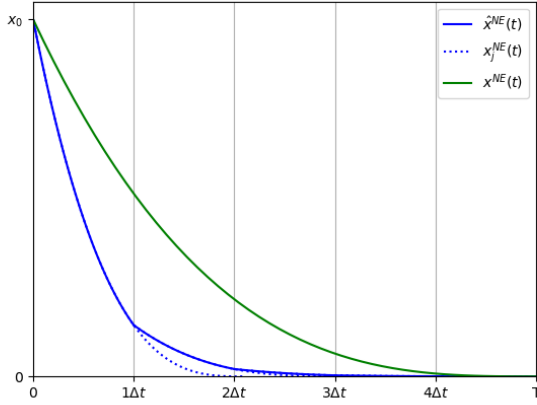


Figure 4.37: Resulting equilibrium trajectory $\hat{x}^{NE}(t)$ (4.337) (blue line) and equilibrium trajectory in original game $x^{NE}(t)$ (4.321) (green Line).

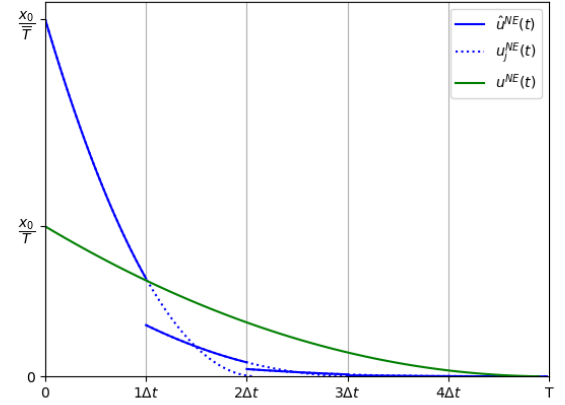


Figure 4.38: Resulting equilibrium strategies $\hat{u}^{NE}(t, \hat{x}^{NE})$ (4.338) (blue line) and equilibrium strategies in original game $u^{NE}(t, x^{NE})$ (4.322) (green line).

subject to

$$\begin{aligned} \dot{x}_t &= - \sum_{i=1}^n u_i^t(\tau, x_t), \\ x_t(t) &= x, \\ x_t(t) &\geq 0, \end{aligned} \tag{4.340}$$

where x is some starting position. Suppose that the parameter t is a fixed constant. We will use the game model (4.339), (4.340) to construct the strategies for a class of games with continuous updating.

Following the procedure presented in [78], we construct Hamilton-Jacobi-Bellman equations to find the cooperative (noncooperative) strategies, namely, the optimal strategies (feedback Nash equilibrium). For this purpose, we need defining the Bellman function for this game model:

$$\begin{aligned} V(t; \tau, x) &= \max_{u^t} \left\{ \sum_{i=1}^n K_i^t(x, t + \bar{T} - \tau; u_1^t, \dots, u_n^t) \right\} = \\ &= \max_{u^t} \left\{ n \int_{\tau}^{t+\bar{T}} \log(u^t(\tau, x_t)) d\tau \right\}, \quad \tau \in [t, t + \bar{T}] \end{aligned} \tag{4.341}$$

$$\left(V_i(t; \tau, x) = K_i^t(x, t + \bar{T} - \tau; u_1^{t,NE}, \dots, u_n^{t,NE}) = \int_{\tau}^{t+\bar{T}} \log(u_i^{t,NE}(\tau, x)) d\tau, \tau \in [t, t + \bar{T}], i \in N \right), \quad (4.342)$$

where the Bellman function $V(t; \tau, x)$ is the maximum value of the total payoff of players (4.339) in the subgame $\Gamma(x, \tau, t + \bar{T})$ starting at the instant $\tau \in [t, t + \bar{T}]$ from the position x

(the Bellman function $V_i(t; \tau, x)$ is the payoff of player $i \in N$ in the feedback Nash equilibrium $u^{t,NE}(\tau, x) = (u_1^{t,NE}(\tau, x), \dots, u_n^{t,NE}(\tau, x))$ in the subgame $\Gamma(x, \tau, t + \bar{T})$ starting at the instant t from the position x).

Using the approach described in Sections 4.5.1.1 and 4.5.1.2, we derive the corresponding Hamilton-Jacobi-Bellman equation for determining the cooperative strategies (feedback Nash equilibrium strategies):

$$\begin{aligned} -V_{\tau}(t; \tau, x) &= \max_{u^t} \{n \log u^t - nu^t V_x(t; \tau, x)\}, \\ V(t; t + \bar{T}, x) &= 0 \end{aligned} \quad (4.343)$$

$$\left(\begin{aligned} -V_{\tau}^i(t; \tau, x) &= \max_{u^t} \{n \log u^t - nu^t V_x^i(t; \tau, x)\}, \\ V(t; t + \bar{T}, x) &= 0 \end{aligned} \right). \quad (4.344)$$

The technique for solving (4.343) and (4.344) is similar to the one used for the classical control problems in Sections 4.5.1.1 and 4.5.1.2. Therefore, to demonstrate the solution technique, we will only present the solution for the cooperative setting (4.343), i.e., determine the cooperative strategies $u^{t,*}(\tau, x)$.

From the first-order optimality condition for (4.343) we obtain

$$u^{t,*} = \frac{1}{V_x(t; \tau, x)}.$$

Substituting this expression into (4.343) yields

$$\begin{aligned} V_t(t; \tau, x) &= n \log V_x(t; \tau, x) + n, \\ V(t; t + \bar{T}, x) &= 0. \end{aligned}$$

We define the Bellman function in the form

$$V(t, \tau, x) = A(t, \tau) \log x + B(t, \tau).$$

Substituting this form into (4.343), we obtain:

$$\begin{aligned} \dot{A}_\tau(t, \tau) \log x + \dot{B}_\tau(t, \tau) &= n \log A(t, \tau) - n \log x + n, \\ A(t, t + \bar{T}) &= B(t, t + \bar{T}) = 0. \end{aligned} \quad (4.345)$$

The solution of (4.345) are the functions

$$\begin{aligned} A(t, \tau) &= n(t + \bar{T} - \tau), \\ B(t, \tau) &= -n(t + \bar{T} - \tau) \log n(t + \bar{T} - \tau), \quad \tau \in [t, t + \bar{T}]. \end{aligned} \quad (4.346)$$

Finally, we obtain an expression for the Bellman function:

$$V(t; \tau, x) = n(t + \bar{T} - \tau) \log \frac{x}{n(t + \bar{T} - \tau)}, \quad \tau \in [t, t + \bar{T}] \quad (4.347)$$

and an expression for the optimal control:

$$u^{t,*}(\tau, x) = \frac{1}{V_x} = \frac{x}{n(t + \bar{T} - \tau)}, \quad \tau \in [t, t + \bar{T}]. \quad (4.348)$$

The control (strategy) $u^{t,*}(\tau, x)$ (4.348) is optimal in the game model $\Gamma(x, t, t + \bar{T})$, which is defined on the interval $[t, t + \bar{T}]$. Therefore, the strategy $u^{t,*}(\tau, x)$ (4.348) for a fixed t is defined on the same interval, i.e., $[t, t + \bar{T}]$. Changing the parameter t , we change the initial instant of the game $\Gamma(x, t, t + \bar{T})$ and, therefore, automatically change the interval on which the optimal control (feedback Nash equilibrium) is calculated. Note that the optimal control (4.348) explicitly depends on the parameter t as well as on the parameters \bar{T} and τ .

Using $u^{t,*}(\tau, x)$ (4.348), we construct the strategy that will be used for proving convergence. Suppose that at the instant t , as a resulting strategy in the game with continuous updating, we use $\tilde{u}^*(t, x) = u^{t,*}(\tau, x)$, where $\tau = t$. This means that at any instant t , the players orient themselves (define the optimal strategies) using the information on the interval $[t, t + \bar{T}]$.

The same procedure can be performed for the equilibrium strategies, which we will denote by $\tilde{u}^{NE}(t, x)$:

$$\tilde{u}^*(t, x) = \frac{x}{n\bar{T}} \quad \left(\tilde{u}^{NE}(t, x) = \frac{x}{\bar{T}} \right). \quad (4.349)$$

Next, we show that the strategies constructed in this way are the strategies in the game with continuous updating, namely, in the case when the interval between the updating instants $\Delta t \rightarrow 0$ (i.e., $l \rightarrow \infty$).

Theorem 4.5.3 *As $\Delta t \rightarrow 0$ (i.e., $l \rightarrow \infty$), the resulting strategies $\hat{u}^*(t, x) = \hat{u}_l^*(t, x)$ (4.334) ($\hat{u}^{NE}(t, x) = \hat{u}_l^{NE}(t, x)$ (4.335)) in the game with dynamic updating uniformly converge to $\tilde{u}^*(t, x)$ ($\tilde{u}^{NE}(t, x)$), which are the resulting strategies with continuous updating (4.349):*

$$\hat{u}_l^*(t, x) \xrightarrow{[t_0, T]} \tilde{u}^*(t, x) \quad \left(\hat{u}_l^{NE}(t, x) \xrightarrow{[t_0, T]} \tilde{u}^{NE}(t, x) \right). \quad (4.350)$$

Proof From the solution of the original game presented in Section 4.5.1 it follows that the strategy of player i in the noncooperative game differs from the corresponding strategy only by the coefficient n . Therefore, in the proof we present only the cooperative case.

We use the following criterion for a uniform convergence of the sequence of functions $\hat{u}_l^*(t, x)$ to the function $\tilde{u}^*(t, x)$ as $l \rightarrow \infty$, where l is the number of updating instants:

$$\lim_{\Delta t \rightarrow 0} \sup_{t \in [t_0, T]} |\hat{u}_l^*(t, x) - \tilde{u}^*(t, x)| = \lim_{l \rightarrow \infty} \sup_{t \in [t_0, T]} |\hat{u}_l^*(t, x) - \tilde{u}^*(t, x)| = 0. \quad (4.351)$$

Suppose that the greatest value of the difference $\hat{u}_l^*(t, x) - \tilde{u}^*(t, x)$ (4.351) belongs to the interval $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$. Then it is necessary to prove the following

equality:

$$\lim_{\Delta t \rightarrow 0} \sup_{t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]} |u_j^*(t, x) - \tilde{u}^*(t, x)| = 0. \quad (4.352)$$

The greatest value of $\hat{u}_l^*(t, x) - \tilde{u}^*(t, x)$ can be achieved in three cases as follows:

- 1 at an inner point of the interval $(t_0 + j\Delta t, t_0 + (j + 1)\Delta t)$; in this case, the optimum of the difference (4.352) can be found using the first-order optimality condition:

$$\begin{aligned} \frac{d}{dt}(u_j^*(t, x) - \tilde{u}^*(t, x)) &= \frac{d}{dt} \left(\frac{x}{n(\bar{T} + j\Delta t + t_0 - t)} - \frac{x}{n\bar{T}} \right) = \\ &= \frac{x}{n(\bar{T} + j\Delta t + t_0 - t)^2}. \end{aligned}$$

The derivative does not reach zero on the interval $(t_0 + j\Delta t, t_0 + (j + 1)\Delta t)$, showing the strict monotonicity of the difference function on the interval $(t_0 + j\Delta t, t_0 + (j + 1)\Delta t)$.

- 2 at the left boundary of the interval $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, i.e., at the instant $t = t_0 + j\Delta t$:

$$\begin{aligned} &\lim_{\Delta t \rightarrow 0} |u_j^*(t_0 + j\Delta t, x) - \tilde{u}^*(t_0 + j\Delta t, x)| = \\ &= \lim_{\Delta t \rightarrow 0} \left| \frac{x}{n(\bar{T} + j\Delta t + t_0 - (t_0 + j\Delta t))} - \frac{x}{n\bar{T}} \right| = \lim_{\Delta t \rightarrow 0} \left| \frac{x}{n\bar{T}} - \frac{x}{n\bar{T}} \right| = 0. \end{aligned}$$

- 3 at the right boundary of the interval $[t_0 + j\Delta t, t_0 + (j + 1)\Delta t]$, i.e., at the instant $t = t_0 + (j + 1)\Delta t$:

$$\begin{aligned} &\lim_{\Delta t \rightarrow 0} |u_j^*(t_0 + (j + 1)\Delta t, x) - \tilde{u}^*(t_0 + (j + 1)\Delta t, x)| = \\ &= \lim_{\Delta t \rightarrow 0} \left| \frac{x}{n(\bar{T} + j\Delta t + t_0 - (t_0 + (j + 1)\Delta t))} - \frac{x}{n\bar{T}} \right| = \\ &\quad \lim_{\Delta t \rightarrow 0} \left| \frac{x}{n(\bar{T} - \Delta t)} - \frac{x}{n\bar{T}} \right| = 0. \end{aligned}$$

Within the interval, the difference (4.352) does not exceed its values at the boundaries of the interval. Moreover, at the left and right boundaries of the interval,

the difference function tends to zero. As a result, the difference function tends to zero within the interval.

In accordance with the criterion (4.351), the sequence of the resultant controls $\hat{u}^*(t, x) = \hat{u}_l^*(t, x)$ with dynamic updating uniformly converges to the control $\tilde{u}^*(t, x)$ with continuous updating as $l \rightarrow \infty$. \square

Theorem 4.5.4 *The cooperative trajectory in the game with dynamic updating $\hat{x}^*(t) = \hat{x}_l^*(t)$ uniformly converges to the trajectory in the game with continuous updating $\tilde{x}^*(t)$ as $\Delta t \rightarrow 0$ (i.e., $l \rightarrow \infty$).*

Proof Note that

$$\begin{aligned}\hat{u}^*(t, x) &= \hat{u}_l^*(t, x) = x\hat{u}^*(t, 1), \\ \tilde{u}^*(t, x) &= x\tilde{u}^*(t, 1).\end{aligned}$$

In the case of dynamic updating, from the motion equation we obtain:

$$\begin{aligned}\frac{d\hat{x}_l^*}{dt}(t) &= -n\hat{u}^*(t, x) = -n\hat{x}_l^*\hat{u}^*(t, 1), \\ \log \hat{x}_l^*(t) &= -n \int_{t_0}^t \hat{u}^*(\tau, 1) d\tau.\end{aligned}$$

In the case of continuous updating,

$$\begin{aligned}\frac{d\tilde{x}^*}{dt}(t) &= -n\tilde{u}^*(t, x) = -n\tilde{x}^*(t)\tilde{u}^*(t, 1), \\ \log \tilde{x}^*(t) &= -n \int_{t_0}^t \tilde{u}^*(\tau, 1) d\tau.\end{aligned}$$

The uniform convergence of the strategies implies

$$\begin{aligned} \hat{u}_i^*(t, x) \underset{[t_0, T]}{\rightrightarrows} \tilde{u}^*(t, x) &\implies n \int_{t_0}^t \hat{u}_i^*(\tau, x) d\tau \underset{[t_0, T]}{\rightrightarrows} n \int_{t_0}^t \tilde{u}^*(\tau, x) d\tau \implies \\ &\implies \exp \left\{ -n \int_{t_0}^t \hat{u}_i^*(\tau, x) d\tau \right\} \underset{[t_0, T]}{\rightrightarrows} \exp \left\{ -n \int_{t_0}^t \tilde{u}^*(\tau, x) d\tau \right\}. \end{aligned} \quad (4.353)$$

Since the function e^{-s} is bounded on the interval $s \in [0, \infty)$, the following holds:

$$\hat{x}_i^*(t, x) \underset{[t_0, T]}{\rightrightarrows} \tilde{x}^*(t, x). \quad (4.354)$$

□

We construct the resulting trajectory with continuous updating by substituting $\tilde{u}^*(t, x)$ ($\tilde{u}^{NE}(t, x)$) into the motion equation (4.310):

$$\begin{aligned} \dot{x} &= -\frac{x}{T}, & \left(\dot{x} &= -n\frac{x}{T} \right). \\ x(t_0) &= x_0 & \left(x(t_0) &= x_0 \right). \end{aligned} \quad (4.355)$$

Then the corresponding solutions are the trajectories in the game with continuous updating:

$$\tilde{x}^*(t) = x_0 e^{-\frac{t-t_0}{T}} \quad \left(\tilde{x}^{NE}(t) = x_0 e^{-n\frac{t-t_0}{T}} \right). \quad (4.356)$$

The optimal strategies (feedback Nash equilibrium strategies) along the resulting trajectory in the game with continuous updating have the form

$$\tilde{u}^*(t, \tilde{x}^*(t)) = \frac{x_0}{nT} e^{-\frac{t-t_0}{T}} \quad \left(\tilde{u}^{NE}(t, \tilde{x}^{NE}(t)) = \frac{x_0}{T} e^{-n\frac{t-t_0}{T}} \right). \quad (4.357)$$

4.5.4 Numerical Simulation

Consider the numerical simulation results for the game model with three symmetric players ($n = 3$) on the interval $[0, 100]$, i.e., $t_0 = 0$ and $T = 100$. At the initial instant $t_0 = 0$, the amount of resource is 2000, i.e., $x_0 = 2000$. Suppose that in the

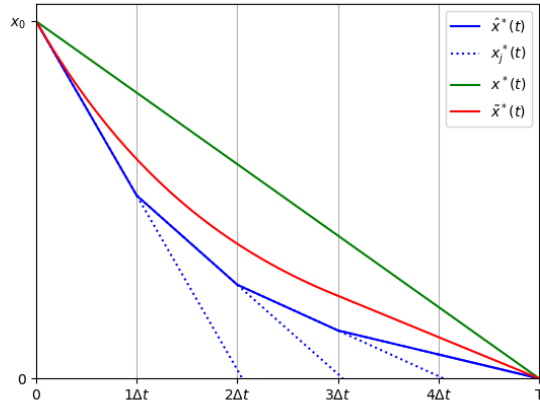


Figure 4.39: Resulting cooperative trajectory $\tilde{x}^*(t)$ (4.356) with continuous updating (red line), resulting cooperative trajectory $\hat{x}^*(t)$ (4.332) with dynamic updating (blue line), and cooperative trajectory in original game $x^*(t)$ (4.321) (green line).

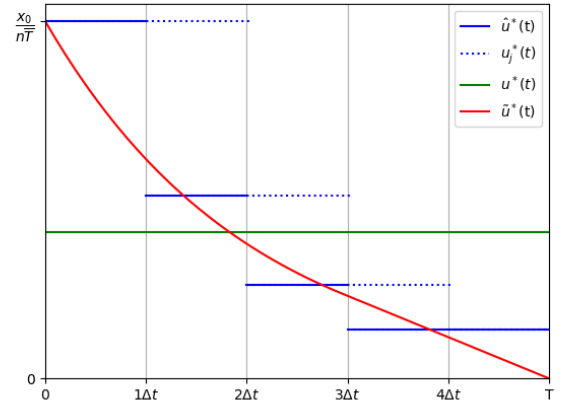


Figure 4.40: Resulting cooperative strategies $\tilde{u}^*(t)$ (4.349) with continuous updating (red line), resulting cooperative strategies $\hat{u}^*(t)$ (4.334) with dynamic updating (blue line), and cooperative strategies in original game $u^*(t)$ (4.322) (green line).

case of dynamic updating (blue solid and dotted lines in Figures 4.39 - 4.42), the intervals between the updating instants are $\Delta t = 20$, which means that $l = 5$. In Figures 4.39 and 4.41, the comparison of the resulting strategies in the original game with prescribed duration (green line) and the game with dynamic updating (blue line) and continuous updating (red lines) in the cooperative and noncooperative cases is presented. In Figures 4.40 and 4.42, similar results are presented for the strategies.

To demonstrate the results of Theorems 4.1 and 4.2 on the convergence of the resulting strategies and resulting trajectory, consider the simulation results for the case of frequent updating, namely, $l = 50$. Figures 4.43-4.46 represent the same solutions as in Figures 4.39-4.42 but for the case $\Delta t = 2$. Therefore, the convergence results are confirmed by the numerical experiments presented below.

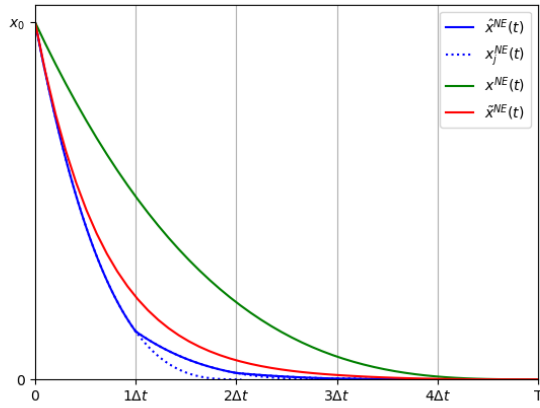


Figure 4.41: Resulting equilibrium trajectory $\tilde{x}^{NE}(t)$ (4.356) with continuous updating (red line), resulting equilibrium trajectory $\hat{x}^{NE}(t)$ (4.337) with dynamic updating (blue line), and equilibrium trajectory in the initial game $x^{NE}(t)$ (4.328) (green line).

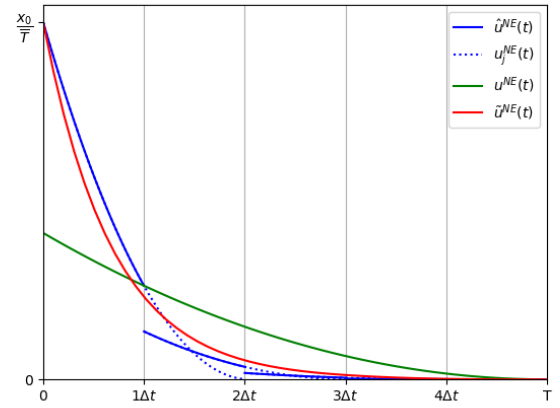


Figure 4.42: Resulting equilibrium strategies $\tilde{u}^{NE}(t)$ (4.349) with continuous updating (red line), resulting equilibrium strategies $\hat{u}^{NE}(t)$ (4.338) with dynamic updating (blue line), and resulting equilibrium strategies in original game $u^{NE}(t)$ (4.329) (green line).

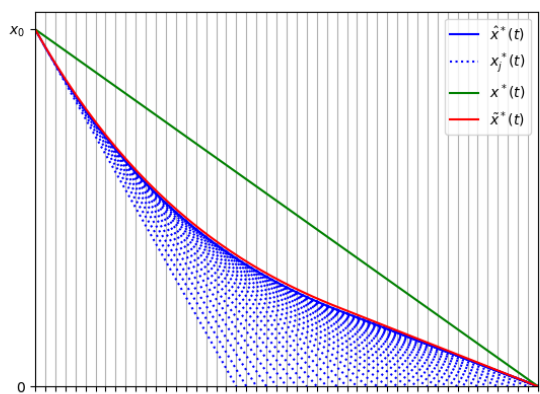


Figure 4.43: $\tilde{x}^*(t)$ (4.356) (red line), $\hat{x}^*(t)$ (4.332) (blue line), and $x^*(t)$ (4.321) (green line).

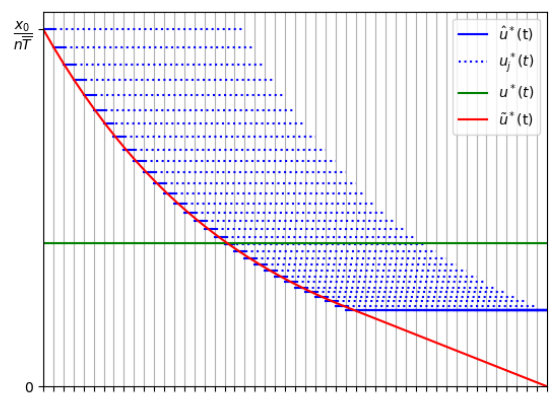


Figure 4.44: $\tilde{u}^*(t)$ (4.349) (red line), $\hat{u}^*(t)$ (4.334) (blue line), and $u^*(t)$ (4.322) (green line).

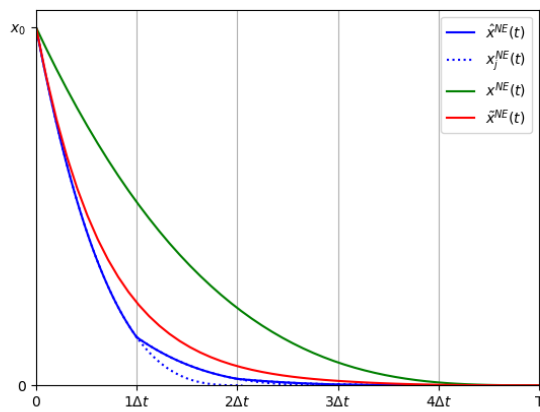


Figure 4.45: $\tilde{x}^{NE}(t)$ (4.356) (red line), $\hat{x}^{NE}(t)$ (4.337) (blue line), and $x^{NE}(t)$ (4.328) (green line).

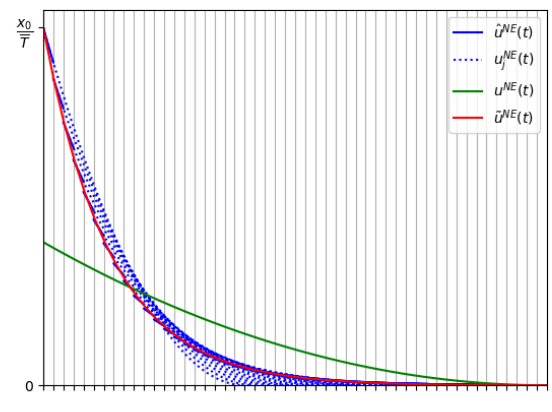


Figure 4.46: $\tilde{u}^{NE}(t)$ (4.349) (red line), $\hat{u}^{NE}(t)$ (4.338) (blue line), and $u^{NE}(t)$ (4.329) (green line).

Chapter 5

Engineering Applications of Continuous Approach to Inverse Optimal Control Problems

The theory of optimal control has received considerable attention to model motion behavior or decision-making of humans. Most approaches are based on a fixed (or infinite) time horizon which implies that all information is available at the beginning of the time interval. Nevertheless, it is reasonable to believe that the human uses information defined by a continuously moving information horizon at each time instant and adapts accordingly. Therefore, in this chapter, an optimal feedback control approach based on the continuous updating paradigm is proposed. The model parameters defining individual human behavior consist of the cost function parameters and the length of the information horizon, which can be identified via a corresponding inverse optimal control approach. This approach is employed to simulations of a potential application example of human behavior identification in a driver assistance system.

5.1 Optimal Control with Continuous Updating

In this section, we present the results concerning optimal control with continuous updating. We first present the classical optimal control problem and then show the continuous updating approach.

5.1.1 Original Optimal Control Problem

Consider the following optimal control problem defined on an interval $[t_0, T]$:

$$J(x_0, t_0; u) = \int_{t_0}^T g[t, x(t), u(t, x)] dt \rightarrow \min_u \quad (5.1)$$

subject to

$$\begin{aligned} \dot{x}(t) &= f(t, x, u), \\ x(t_0) &= x_0, \\ x &\in \text{comp}\mathbb{R}^n, u = u(t, x) \in U \subset \text{comp}\mathbb{R}^m, t \in [t_0, T], \end{aligned} \quad (5.2)$$

where $\text{comp}\mathbb{R}^m$ denotes a compact set in the m -dimensional real space; $g[t, x(t), u(t, x(t))]$ and $f(t, x, u)$ are integrable functions; $x(t) \in \mathbb{R}^n$ is the solution of the Cauchy problem (5.2) with a fixed $u(t, x) \in \mathbb{R}^m$. A control $u(t, x) \in \mathbb{R}^m$ is called admissible if problem (5.2) has a unique and continuous solution.

5.1.2 Optimal Control Problem with Continuous Updating

Using the original optimal control problem defined on the closed time interval $[t_0, T]$, we construct the corresponding optimal control problem with continuous updating.

Consider the following optimal control problem defined on the interval $[t, t + \bar{T}]$, where $0 < \bar{T} < +\infty$:

$$J(x, t; u^t) = \int_t^{t+\bar{T}} g[s, x_t(s), u^t(s, x_t)] ds \rightarrow \min_{u^t} \quad (5.3)$$

subject to

$$\begin{aligned} \dot{x}_t(s) &= f(s, x_t, u^t), \\ x_t(t) &= x, \\ x_t &\in \text{comp}\mathbb{R}^n, u^t = u^t(s, x_t) \in U \subset \text{comp}\mathbb{R}^m, s \in [t, t + \bar{T}], \end{aligned} \quad (5.4)$$

where $u^t(s, x_t) \in \mathbb{R}^m$ and $x_t(s) \in \mathbb{R}^n$ are the optimal control and state trajectories on the interval $[t, t + \bar{T}]$, respectively.

The main characteristic of the optimal control problem with continuous updating is the following: *the current time $t \in [t_0, +\infty)$ evolves continuously, and consequently, the human continuously obtains new information about the motion equation and objective function on the interval $[t, t + \bar{T}]$.*

The control $u(t, x)$ in the optimal control problem with continuous updating has the form

$$u(t, x) = u^t(s, x)|_{s=t}, \quad t \in [t_0, +\infty), \quad (5.5)$$

where $u^t(s, x)$, $s \in [t, t + \bar{T}]$, is the control in the problem defined on the interval $[t, t + \bar{T}]$, and $u^t(s, x)|_{s=t}$ is its part at the first instant $s = t$. The main idea of (5.5) is that, as the current time t evolves, information is updated. Therefore, to model the behavior of the human, we need to consider the control $u^t(s, x)$ only at the points $s = t$.

The trajectory $x(t)$ in the optimal control problem with continuous updating is determined by the system dynamics (5.2), where $u = u(t, x)$ is the control (5.5). Assume that the control with continuous updating (5.5) is admissible.

The essential difference between the control problem with continuous updating and classical optimal control problem defined on the closed interval consists in the following: in the latter problem, the decision maker is guided by the objective received eventually on the interval $[t_0, T]$; in the former problem (with continuous updating), at the time instant t , the system is oriented on the expected objective (5.3), which is calculated using the information on the interval $[t, t + \bar{T}]$ (the information the system has at the instant t).

5.1.3 Optimal Control with Continuous Updating

Under continuously updated information, we use the concept of optimal control in feedback form $u^*(t, x)$. Furthermore, we require that for any fixed current time $t \in [t_0, +\infty)$, $u^*(t, x)$ coincides with the optimal control in problem (5.3) and (5.4), defined on the interval $[t, t + \bar{T}]$ at the instant t .

However, direct application of the classical approaches for optimal control in feedback form is not possible. Consider another current time instant $t + \epsilon$, where $\epsilon \ll \bar{T}$. Due to the requirement above, $u^*(t + \epsilon, x)$ at the instant $t + \epsilon$ must coincide with the optimal control in the problem defined on the interval $[t + \epsilon, t + \epsilon + \bar{T}]$.

Therefore, $u^*(t, x)$, $t \in [t_0, +\infty)$, should be defined as an infinite combination of the optimal controls on the intervals $[t, t + \bar{T}]$ for every current time instant $t \in [t_0, +\infty)$.

To construct such controls, we consider the concept of generalized optimal control in feedback form

$$\tilde{u}^*(t, x; s, x), \quad t \in [t_0, +\infty), \quad s \in [t, t + \bar{T}], \quad (5.6)$$

which will be used below for constructing the control $u^*(t, x)$.

Definition 5.1.1 *A control $\tilde{u}^*(t, x; s, x)$ is a generalized optimal control in the problem with continuous updating if it is optimal, for any fixed $t \in [t_0, +\infty)$, in the problem with the objective function (5.3) and the system dynamics (5.4).*

Using generalized optimal control, we define a solution concept for the optimal control problem with continuous updating.

Definition 5.1.2 *A control $u^*(t, x)$ is called an optimal control with continuous updating if*

$$u^*(t, x) = \tilde{u}^*(t, x; s, x)|_{s=t}, \quad t \in [t_0, +\infty), \quad (5.7)$$

where $\tilde{u}^*(t, x; s, x)$ is the generalized optimal control in the sense of Definition 5.1.1.

5.2 Linear-Quadratic Optimal Control with Continuous Updating

In this section, we give solutions for the optimal control problem with continuous updating according to Definition 5.1.2 for the class of linear-quadratic optimal control.

5.2.1 Problem Formulation in Linear-Quadratic Case with Continuous Updating

According to the general problem statement of the optimal control problem with continuous updating, the linear-quadratic case will have the following form.

Consider the optimal control problem defined on the interval $[t, t + \bar{T}]$, where $0 < \bar{T} < +\infty$:

$$J(x, t; u^t) = \int_t^{t+\bar{T}} x_t^T(s) Q x_t(s) + u^t(s, x_t)^T R u^t(s, x_t) ds \rightarrow \min_{u^t} \quad (5.8)$$

subject to

$$\begin{aligned} \dot{x}_t(s) &= Ax_t(s) + Bu^t(s, x_t), \\ x_t(t) &= x, \\ x_t &\in \mathbb{R}^n, u^t = u^t(s, x_t) \in U \subset \text{comp} \mathbb{R}^m, s \in [t, t + \bar{T}]. \end{aligned} \quad (5.9)$$

The definitions of the generalized optimal control and the optimal control with continuous updating are analogous to Definitions 5.1.1 and 5.1.2, respectively.

5.2.2 Optimal Control with Continuous Updating for LQP

We present an explicit form of the optimal control with continuous updating using the system of Riccati differential equations.

Theorem 5.2.1 *For any initial state x_0 , the linear-quadratic control problem with continuous updating (5.8) and (5.9) has a solution if and only if the Riccati differential equation (5.10) has a symmetric solution $K(\cdot)$ on the interval $[0, 1]$:*

$$\begin{aligned} \dot{K}(\tau) &= -\bar{T}A^T K(\tau) - \bar{T}K(\tau)A + \bar{T}K(\tau)SK(\tau) - \bar{T}Q, \\ K(1) &= 0, \end{aligned} \quad (5.10)$$

where $S = BR^{-1}B^T$, the matrices Q and R are symmetric, and R is positive

definite.

If the linear-quadratic control problem with continuous updating has a solution, then it is unique, and the optimal control in feedback form is

$$u^*(t, x) = -R^{-1}B^TK(0)x. \quad (5.11)$$

Proof To prove the theorem, we introduce the change of variables

$$\begin{aligned} s &= t + \bar{T}\tau, \\ y_t(\tau) &= x(t + \bar{T}\tau), \\ v_t(\tau, y_t) &= u(t + \bar{T}\tau, x). \end{aligned} \quad (5.12)$$

Substituting (5.12) into the motion equations (5.9) and the objective function (5.8), we obtain

$$\dot{y}_t(\tau) = \bar{T}Ay_t(\tau) + \bar{T}Bv_t(\tau, y_t) \quad (5.13)$$

and

$$J(y_t, t; v_t) = \int_0^1 \bar{T}y_t^T(\tau)Q_i y_t(\tau) + \bar{T}(v_t(\tau, y_t))^T R v_t(\tau, y_t) d\tau. \quad (5.14)$$

As is known (e.g., see [224, Theorem 5.1]), a criterion for the existence of optimal control in feedback form is the existence of a symmetric solution for the system of differential equations (5.10). The optimal control has the form

$$v_t^*(\tau, y_t) = -R^{-1}B^TK(\tau)y_t. \quad (5.15)$$

From (5.12) we have

$$\tau = \frac{s - t}{\bar{T}}.$$

Returning to the original variables, we obtain the following control:

$$u^t(s, x) = -R^{-1}B^TK\left(\frac{s - t}{\bar{T}}\right)x.$$

This control is optimal in the problem defined on the interval $[t, t + \bar{T}]$ by construction. Equations (5.13), the function (5.14), and the solution (5.15) have the same form for all values t . Then the generalized optimal control in the control

problem with continuous updating has the form

$$\tilde{u}^*(t, s, x) = -R^{-1}B^TK \left(\frac{s-t}{\bar{T}} \right) x. \quad (5.16)$$

Using (5.7), we set the optimal control with continuous updating equal to the generalized optimal control (5.16), for $s = t$:

$$u^*(t, x) = -R^{-1}B^TK(0)x, \quad t \in [t_0, +\infty). \quad (5.17)$$

Thus, the control (5.15) in the subproblem exists for any initial value y_t ($t \geq t_0$). Hence, the control (5.17) exists in the optimal control problem with continuous updating for any x_0 . \square

Using the optimal control with continuous updating $u^*(t, x)$ (5.11), we propose a corresponding inverse optimal control problem with continuous updating.

5.3 Inverse Optimal Control Problem with Continuous Updating

Suppose that the function $g[t, x(t), u(t, x(t))]$ in (5.1) is parameterized. Therefore, in an inverse optimal control problem, this parametrization is unknown and has to be estimated from the observed control and state trajectories. Furthermore, the value of the information horizon \bar{T} is also unknown. We denote by θ the set of unknown parameters, including the cost function parametrization and the information horizon. The inverse control problem with continuous updating involves the estimation of the unknown parameter set θ based on the observed control $\hat{u}(t)$ and the corresponding trajectory $\hat{x}(t)$ on the interval $[t_0, T]$. The objective is to obtain an estimated optimal control $u_\theta^*(t, x_\theta^*(t))$ and a corresponding trajectory $x_\theta^*(t)$ by minimizing the difference between these model trajectories and the observed ones.

Below, the inverse optimal control problem with continuous updating is solved with an approach analogous to the bi-level approach to standard optimal control problems introduced in [235]. In the cited paper, it was successfully applied with experimental data. The model parameters θ are determined by minimizing the

squared error between the observed and model trajectories. This objective is represented by the optimization problem

$$Div = \int_{t_0}^T \|x_\theta(t) - \hat{x}(t)\|^2 + \|u_\theta(t, x_\theta^*(t)) - \hat{u}(t)\|^2 dt \rightarrow \min_{\theta}, \quad (5.18)$$

where x_θ and u_θ are the optimal control trajectories arising from the minimization of the cost function (5.1) with the parameters θ . Note that in the continuous updating inverse optimal control approach, $u_\theta(t, x_\theta(t))$ is defined as in (5.7), and $x_\theta(t)$ arises from the corresponding solution of the system dynamics. Therefore, we will consider the linear-quadratic optimal control problem with continuous updating, exploiting the explicit form (5.11) for solving (5.18).

5.4 Simulation Results

In this section, we illustrate the proposed optimal control approach with continuous updating and apply the corresponding solution of the inverse problem to identify the behavioral parameters of the driver controlling the lateral dynamics of a vehicle.

5.4.1 Single-track and Steering Model

The simulated system is a single-track and steering model with six states from [236]. The model is a validated linear approximation of steering for a midsize passenger car.

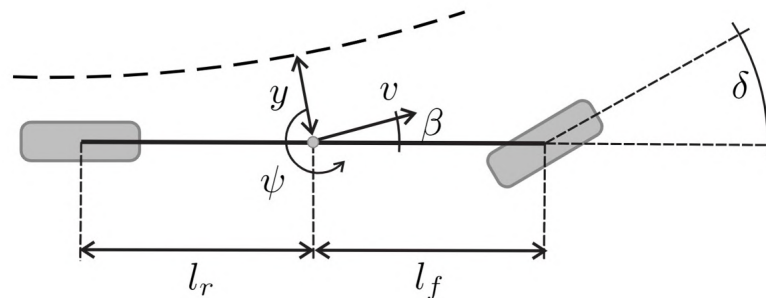


Figure 5.1: Geometric relations of single-track vehicle model.

The system's state vector is given by

$$x(t) = \begin{bmatrix} \beta(t) & \psi(t) & \dot{\psi}(t) & y(t) & \delta(t) & \dot{\delta}(t) \end{bmatrix}^T,$$

where $\beta(t)$ is the side-slip angle, $\psi(t)$ is the yaw angle, $\dot{\psi}(t)$ is the yaw angle velocity, $y(t)$ is the lateral distance to the middle lane, $\delta(t)$ is the steering wheel angle, and $\dot{\delta}(t)$ is the steering wheel angle velocity. All states are in SI units. The control $u(t, x) \in \mathbb{R}$ in the system is the torque applied by the human.

The system dynamics have the form

$$\dot{x}(t) = A x(t) + B u(t) \quad (5.19)$$

with a given initial state $x(t_0) = x_0$ and

$$A = \begin{bmatrix} \frac{-C_f - C_r}{Mv} & \frac{C_r l_r - Mv^2 - C_f l_f}{Mv^2} & 0 & 0 & \frac{C_f}{Mv i_s} & 0 \\ \frac{C_r l_r - C_f l_f}{J_z} & \frac{-C_f l_f^2 - C_r l_r^2}{J_z v} & 0 & 0 & \frac{C_f l_f}{J_z i_s} & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ v & 0 & v & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \frac{-C_s}{J_s} & \frac{-D_s}{J_s} \end{bmatrix}, \quad (5.20)$$

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \frac{1}{J_s} \end{bmatrix}^T.$$

The geometric relations of the model are demonstrated schematically in Figure 5.1. The system's parameters are combined in Table 5.1.

Similar to, e.g., [237], suppose that the driver's objective function has a quadratic form. Therefore, the integrand in (5.1) and (5.8) is

$$g[t, x(t), u(t, x), \theta] = x^T(t) Q x(t) + u(t, x)^T R u(t, x). \quad (5.21)$$

The parameter θ has to be identified based on the generated data using the approach (5.18). The parameter θ consists of the elements of the matrices Q and R and the time horizon \bar{T} :

$$\theta = \{Q, R, \bar{T}\}. \quad (5.22)$$

Table 5.1: System constants of linear single-track steering model [236]

Name	Value	Unit	Description
v	20	$\frac{\text{m}}{\text{s}}$	vehicle velocity
l_f	1.5	m	distance from center of mass to front wheelbase
l_r	1	m	distance from center of mass to rear wheelbase
C_f	137.5	$\frac{\text{kN}}{\text{rad}}$	front tire slip
C_r	137.5	$\frac{\text{kN}}{\text{rad}}$	rear tire slip
M	1500	kg	vehicle mass
J_z	1800	kg m^2	vehicle yaw inertia
D_s	1	$\frac{\text{Nm s}}{\text{rad}}$	damping constant of steering train
C_s	1	$\frac{\text{Nm}}{\text{rad}}$	retraction constant of steering wheel
J_s	0.1	$\frac{\text{Nm s}^2}{\text{rad}}$	inertia of steering train
i_s	16	—	steering transfer constant

5.4.2 Optimal Control with Continuous Updating

To illustrate different solutions arising from optimal control with continuous updating, we first define the ground truth parameters of the objective function of the human driver:

$$Q = \text{diag}(1, 10, 1, 40, 10, 1), \quad R = [1]. \quad (5.23)$$

Let the initial state at $t_0 = 0$ be

$$x_0 = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0.1 \end{bmatrix}.$$

We use the optimal control with continuous updating (Theorem 5.2.1) to generate data sets simulating the measurements of human steering behavior. This has been done by solving (5.10) numerically using `ode45` solver in MATLAB. The data sets, denoted by D1–D4, consist of the optimal state and control trajectories with $t \in [0, 10]$ and correspond to the information horizons

$$\bar{T} = 1, \quad \bar{T} = 1.5, \quad \bar{T} = 2, \quad \bar{T} = 2.5, \quad (5.24)$$

respectively. All data sets have been generated based on the same cost function parameters (5.23) and the system dynamics (5.20). We exemplarily show the corresponding optimal control trajectories in Figure 5.2 as well as the lateral distance to the middle lane and the steering wheel angle in Figure 5.3 and Figure 5.4, respectively.

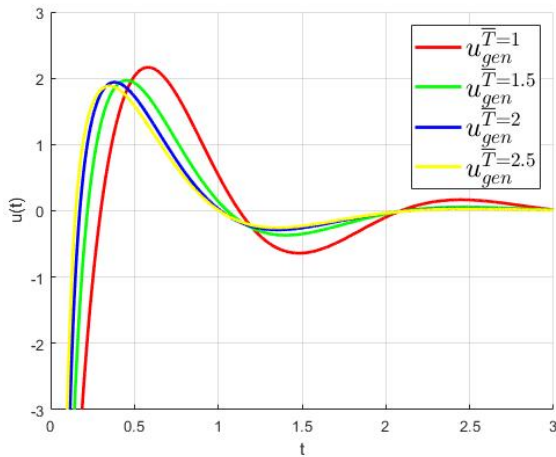


Figure 5.2: Optimal controls with continuous updating with different lengths of the information horizon.

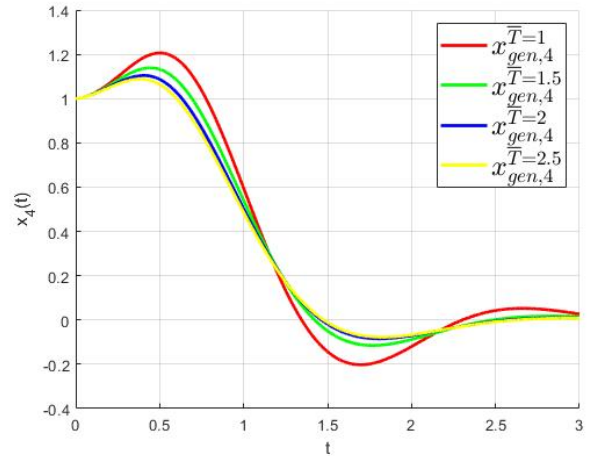


Figure 5.3: $y(t)$ in control problem with continuous updating and different lengths of information horizon.

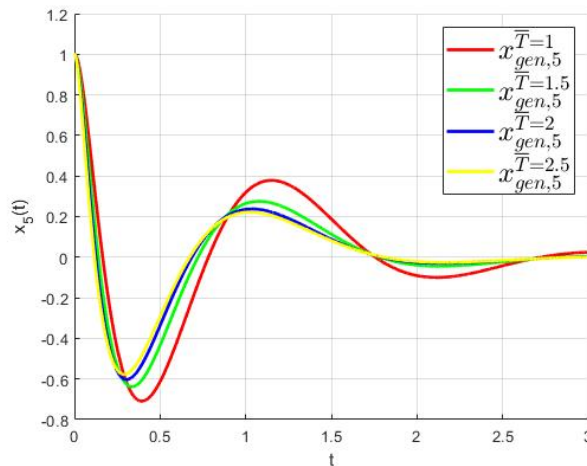


Figure 5.4: $\delta(t)$ in control problem with continuous updating and different lengths of information horizon.

5.4.3 Inverse Optimal Control Problem

Now we solve the inverse optimal control problem (5.18) using the continuous updating approach to obtain an estimate of θ as defined in (5.22). We use the data set D2, i.e., the optimal control trajectories with continuous updating with $\bar{T} = 1.5$, as the observed trajectories $\hat{x}(t)$ and $\hat{u}(t)$ on the interval $[0, 10]$. To show the significance of identifying \bar{T} apart from the cost function matrices, we present the optimal matrices $[Q, R]$ for different fixed values of \bar{T} defined in (5.24). Furthermore, we show the corresponding values of the objective function Div (5.18). To avoid ambiguity of the parameters, we normalize all cost function parameters with respect to $R = [1.00]$. The results are given below.

$$\begin{aligned}
Div^{\bar{T}=1} &= 7.7944 \cdot 10^{-9}, R = 1.00 \\
Q^{\bar{T}=1} &= \text{diag}(17.11, 38.56, 727.49, 91.64, 65.98, 2.50), \\
Div^{\bar{T}=1.5} &= 5.4986 \cdot 10^{-13}, R = 1.00 \\
Q^{\bar{T}=1.5} &= \text{diag}(11.95, 9.91, 1.00, 40.00, 10.01, 1.00), \\
Div^{\bar{T}=2} &= 0.017448, R = 1.00 \\
Q^{\bar{T}=2} &= \text{diag}(0.00, 0.00, 0.00, 18.54, 0.00, 0.34), \\
Div^{\bar{T}=2.5} &= 0.11415, R = 1.00 \\
Q^{\bar{T}=2.5} &= \text{diag}(0.00, 0.00, 0.00, 9.79, 0.00, 0.02).
\end{aligned} \tag{5.25}$$

The inverse optimal control approach has been introduced in Section 5.3, where we also have identified the information horizon \bar{T} . The solution of (5.18) is calculated using the Sequential Quadratic Programming (SQP) method by `fmincon` solver in MATLAB. (The parameters are required to be positive.) The identification results are:

$$\begin{aligned}
\bar{T}_{opt} &= 1.4815 \\
Div^{\bar{T}_{opt}} &= 1.3065 \cdot 10^{-10}, R = 1.00, \\
Q^{\bar{T}_{opt}} &= \text{diag}(177.12, 8.82, 9.84, 41.26, 10.99, 1.03).
\end{aligned} \tag{5.26}$$

Based on the identified parameters $Q^{\bar{T}_{opt}}$ and \bar{T}_{opt} , we generate the trajectories $u_{id}^{\bar{T}_{opt}}$

and $x_{id}^{\bar{T}_{opt}}$ corresponding to the optimal control with continuous updating. Figure 5.5 shows the original optimal control (from the data set D2, black line) and the optimal control with continuous updating from the identified parameters (yellow line). Similarly, in Figure 5.6 and Figure 5.7, the ground truth and identified model trajectories are shown for the lateral distance to the middle lane $y(t)$ and the steering wheel angle $\delta(t)$.

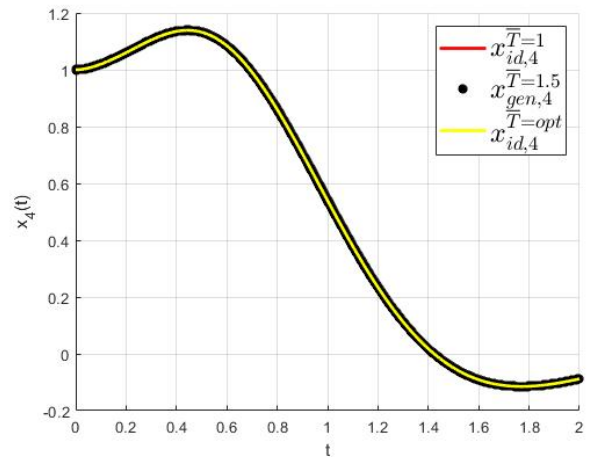
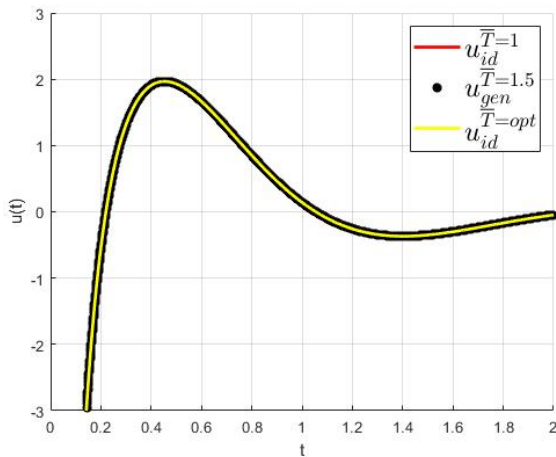


Figure 5.5: Ground truth control (black line) and identified optimal control with continuous updating (yellow line) generated under $\bar{T}_{opt} = 1.4815$ and $Q^{\bar{T}_{opt}}$.

Figure 5.6: Ground truth $y(t)$ (black line) and identified optimal control with continuous updating (yellow line) generated under $\bar{T}_{opt} = 1.4815$ and $Q^{\bar{T}_{opt}}$.

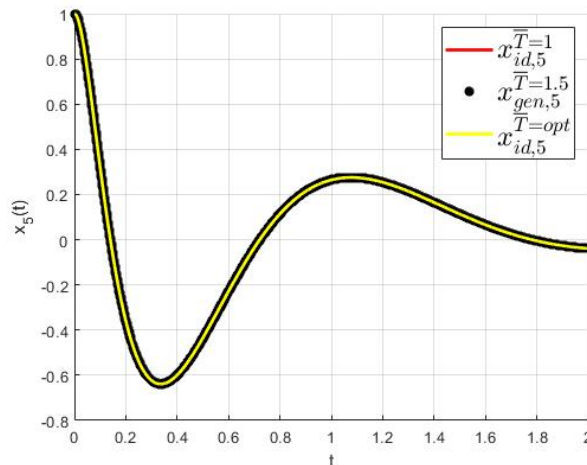


Figure 5.7: Ground truth $\delta(t)$ (black line) and identified optimal control with continuous updating (yellow line) generated under $\bar{T}_{opt} = 1.4815$ and $Q^{\bar{T}_{opt}}$.

5.4.4 Discussion

According to Figure 5.2, the optimal control trajectories are different depending on the assumed length of the information horizon \bar{T} . Additionally, Figure 5.3 and Figure 5.4 show that more information makes the stabilization of the states more effective.

The solutions (5.25) indicate the following: if the human acts with a moving information horizon, then the assumption of a fixed—possibly wrong—information horizon may lead to suboptimal results. This fact shows the importance of identifying the optimal value of information horizon \bar{T} based on the observed data, which can be performed using the inverse optimal control approach. The identified cost function parameters $Q^{\bar{T}_{opt}}$ with optimal \bar{T} are closest to the ground truth ones (5.23) but differ especially in the first diagonal element of the matrix Q (the parameter of the slip angle deviation). In this maneuver, the value of the slip angle does not affect the other states considerably. Therefore, it is a challenge to identify the corresponding parameters correctly. Nevertheless, the identified parameters still adequately describe the ground truth trajectories.

5.5 Conclusion

In this section, an optimal control approach and the corresponding inverse problem based on continuous updating have been presented. The decision-maker updates his behavior based on the new information available, which arises from shifting the time horizon. The simulation results have illustrated the applicability of the approach and highlighted the importance of identifying the real value of information horizon \bar{T}_{opt} . The proposed approach enables more profound modeling of engineering systems with humans. In future works, inverse optimal control with continuous updating will be tested on real measured data from different human drivers.

Conclusion

Let us give a brief overview of the results obtained in this work. The introduction has provided an overview of the literature on the topic of study and discusses its relevance, theoretical and practical value, and scientific novelty.

In Chapter 1, a novel cooperative solution for differential games IDP-core has been introduced. The solution has been constructed using a special form of conditions for imputation distribution procedures and possesses the strong time-consistency property by construction. Also, a numerical algorithm to analyze the non-emptiness property of the IDP-core has been presented.

In Chapter 2, the class of dynamic games with information updating has been introduced. Cooperative and noncooperative cases have been considered, and conditions for related optimality principles such as Nash equilibrium and cooperative strategies have been given. The conditions are based on Bellman equations for the dynamic setting. Additionally, the random horizon updating procedure has been introduced for the dynamic updating setting. The characteristic function for the dynamic updating setting has been defined; the relationship between the cooperative solutions in each truncated subgame and the game defined on the entire interval has been established. The dynamic oligopoly marketing model of advertising has been considered for simulation.

Chapter 3 has presented the class of cooperative differential games with dynamic updating. Several dynamic updating models have been considered, such as the classical model where the players have certain information about the payoff functions and motion equations and the updating model with random information horizon and uncertain stochastic future. For the construction of cooperative strategies and the characteristic function with dynamic updating, the conditions for the optimality are the Hamilton-Jacobi-Bellman equations. For the cooperative solution with dynamic

updating, the property of Δt -time consistency has been proved. The notion of characteristic function with dynamic updating has been introduced; the connection between the cooperative solutions in each truncated subgame and the game defined on the entire time interval has been proved.

Chapter 4 has been devoted to the class of cooperative differential games with continuous updating. In this class, the updating procedure is performed continuously in time. Updating has been implemented using the classical approach where the players have certain information about the payoff function and motion equation. Optimality principles for the cooperative and Nash equilibrium strategies, both for the open-loop based and closed-loop based ones, have been presented in the form of Hamilton-Jacobi-Bellman equations, Pontryagin's Maximum principle with continuous updating, and conditions for a special case of the linear-quadratic (autonomous and nonautonomous) continuous updating models have been presented. Additionally, the convergence results have been obtained on the connection between the solutions of the game models with dynamic and continuous updating in the linear-quadratic autonomous and nonautonomous cases and one special classical differential game model of resource extraction. The convergence of the cooperative and Nash equilibrium strategies and related trajectories with dynamic updating to the corresponding strategies and trajectories with continuous updating has been proved. The strong time consistency of the cooperative solution with continuous updating has been established.

Finally, Chapter 5 has presented an application of the continuous updating approach to engineering. A new inverse optimal control problem with continuous updating of information has been formulated and solved. This problem and the corresponding solution are important for determining the human parameters in human-machine engineering systems (such as driving assistance systems, adaptive exoskeletons, etc.). Additional results on the continuous updating approach to the applied inverse optimal control problem of assistant driving systems have been considered.

Thus, the goal and objectives of the research study were fully achieved. The author express the deepest gratitude to D.Sc., head of the department of Mathematical Modeling of Energy Systems SPbU, Prof.V. V. Zakharov for support, help, discussion of the results and a rich research experience.

Further studies can be focused on new types of information updating. Bayesian updating seems to be a promising area for research: it can be applied in different types of human-machine engineering systems. Also, it is important to study the conditions for the uniqueness and existence of the cooperative strategies and Nash equilibrium with continuous updating.

References

1. Isaacs R. Differential Games. —New York : John Wiley and Sons, 1965.
2. Bellman R. Dynamic Programming. —Princeton : Princeton University Press, 1957.
3. Bercovitz L. D. A variational approach to differential games // Adv. in game theory, Ann. of Math. Studies. —1964. —P. 127–175.
4. Leitmann G. Cooperative and Non-Cooperative Many Players Differential Games. —N.Y.:Springer-Verlag, 1974.
5. Fleming W. H. Rishel. R. W. Deterministic and Stochastic Optimal Control. —New York: Springer-Verlag, 1975.
6. A. Friedman. Differential Games. —Wiley, N.Y., 1971.
7. L.S. Pontryagin. On the theory of differential games // Uspekhi matematicheskikh nauk. —1966. —Vol. 21, no. 4 (130). —P. 219–274.
8. L.S. Pontryagin V. G. Boltyansky R. V. Gamkrelidze E. F. Mishchenko. Mathematical theory of optimal processes. —M .: Science, 1976.
9. Petrosyan L. A., Tomsky G. V. Dynamic games and their applications. —Leningrad : Ed. Leningrad State University, 1982.
10. A. Petrosyan L. Differential games of pursuit. —L .: Publishing house Leningrad. University, 1977.
11. N.N. Krasovskiy. Game tasks about the meeting of movements. —M .: Nauka, 1970.

12. Krasovsky N. N. A. N. Kotelnikov. On the differential game for interception // Proceedings of the Mathematical Institute. V.A. Steklova. –2010. –no. T. 268. –P. 168–214.
13. Pshenichny B.N. Ostapenko V.V. Differential games. –Kiev: Nauk. Dumka, 1992.
14. N.N. Krasovskiy A.I. Subbotin. Positional differential games. –M.: Nauka, 1974.
15. A.I. Subbotin A.G. Chentsov. Optimization of guarantees in control problems. –M.: Nauka, 1981.
16. Subbotin A.I. Generalized solutions of differential equations of the 1st order. Prospects for dynamic optimization. –M.: Izhevsk, 2003.
17. Nash J. F. Equilibrium points in n-person games. // Proc. Natl. Acad. Sci. United States. –1950. –no. 36. –P. 48–49.
18. E. M. Vaysbord V. I. Zhukovsky. Introduction to differential games of several persons and their applications. –M.: Sov. radio, 1980.
19. Grigorenko N. L. Differential pursuit games by several objects. –M.: Publishing house of Moscow State University, 1983.
20. N.L.Grigorenko. To the theory of differential games of three persons // Proceedings of the IMM UB RAS. –2006. –Vol. 12, no. 1. –P. 78–86.
21. V.I Zhukovsky A. A. Chikriy. Linear-quadratic differential games. –Kiev: Naukova Dumka, 1994.
22. V. I. Zhukovsky M. E. Salakudze. Some game control problems and their applications. –Tbilisi, 1998.
23. V. I. Zhukovsky N. T. Tynyansky. Equilibrium controls of multicriteria dynamical systems. –M.: Publishing house Moscow State University, 1994.
24. Kleimenov A. F. Tasks of conflict management // App. Mathematics and Mechanics. –1975. –Vol. 39, no. 2. –P. 225–234.

25. Kleimenov A. F. On the cooperative theory of noncooperative positional differential games // Dokl. Academy of Sciences of the USSR. –1990. –Vol. T.39, no. 1. –P. 32–35.
26. Kleimenov A. F. Non-antagonistic positional differential games. –Ekaterinburg : Science, 1993.
27. Kleimenov A. F. Cooperative solutions in the position differential game of many individuals with continuous payment functions // Applied Mathematics and Mechanics. –1990. –T. 54, № 3. –C. 389–394.
28. Kleimenov A. F. On solutions in a nonantagonistic positional differential game // App. Mathematics and Mechanics. –1997. –Vol. 61, no. 5. –P. 739–746.
29. Kononenko A. F. The structure of an optimal strategy in dynamic controlled systems // Zh. Vychisl. mat. and mat. fiz. –1980. –Vol. 20, no. 5. –P. 1105–1116.
30. V. N. Lagunov V.V.Sushkin. Multistep positional games of N persons. –Tver, 1993.
31. Malafeev O.A. On the existence of a equilibrium situation in dynamic games with dependent movements // Zh. Vychisl. mat. and mat. fiz. –1974. –Vol. 14, no. 1. –P. 88–98.
32. Malafeev O.A. On the existence of a equilibrium situation in differential coalition-free games of two persons with independent movements // Bulletin of Leningrad State University. –1980. –no. 7.
33. Malafeev O.A. Stationary strategies in differential games // Zh. Vychisl. mat. and mat. fiz. –1977. –Vol. 17, no. 1. –P. 42–51.
34. Malafeev O.A. Controlled conflict systems. –SPb .: Publishing house of SPbSU, 2000.
35. F.F Nikitin S.V Chistyakov. On differential antagonistic games with unlimited duration // Bulletin of St. Petersburg State University. –1980. –Vol. 1, no. 3. –P. 37–43.

36. Petrov N.N. Proof of the existence of the value of the pursuit game with bounded time // *Differential Uravn.* –1970. –Vol. 6, no. 5. –P. 784–797.
37. Petrov N.N. On the existence of the meaning of the game of pursuit // *DAN SSSR.* –1970. –Vol. 190, no. 6. –P. 1289–1291.
38. G. Chentsov A. On the game problem of convergence at a given moment of time // *Mat. Sat.* –1976. –Vol. 99, no. 3. –P. 394–420.
39. G. Chentsov A. On an example of an irregular differential game // *PMM.* –1976. –Vol. 40, no. 6. –P. 1113–1116.
40. Chistyakov S. V. Dynamic aspect of the solution of classical cooperative games // *Dokl. RAS.* –1993. –Vol. 330, no. 6. –P. 707–709.
41. Chistyakov S. V. On non-cooperative differential games // *Reports of the USSR Academy of Sciences.* –1981. –T. 259, № 5. –C. 1052–1055.
42. Chistyakov S. V. On the construction of strongly dynamically stable solutions cooperative differential games // *Vestn. SPbSU.* –1992. –no. 1. –P. 50–54.
43. Chistyakov S. V. Program iterations and universal delta - optimal strategy in positional differential game // *Dokl. AN SSSR.* –1991. –Vol. 319, no. 6. –P. 1333–1336.
44. S. V. Chistyakov FF Nikitin. Existence and uniqueness theorem for the solution of the generalized Isaacs-Bellman equation // *Differential Equations.* –2007. –Vol. 43, no. 6. –P. 757–766.
45. Subbotina N. N. The method of characteristics for Hamilton-Jacobi equations and applications to dynamical optimization // *Journal of math. sciences.* –2007. –Vol. 135, no. 3. –P. 2955–3091.
46. Zhukovskii V.I. *Lyapunov Functions in Differential Games.* –London and New York: Taylor Francis, 2003.
47. E. J. Dockner Van Long N. International pollution control: Cooperative versus noncooperative strategies // *Journal of Environmental Economics and Management.* –1993. –Vol. 25. –P. 13–29.

48. H. S. Bierman L. Fernandez. Game Theory with Economic Applications. –Massachusetts, USA : Addison: Wesley Publishing Company, Inc., 1993.
49. Gibbons R. Game Theory for Applied Economists. –Princeton, New Jersey : Princ. Univ. Press, 1992.
50. Dockner E. J., Jorgensen S., et al. Differential Games in Economics and Management Science. –Cambridge : Cambridge University Press, 2000.
51. L. A. Petrosyan N. A. Zenkevich E. V. Shevkoplyas. Game Theory. –SPb: BHV-Petersburg, 2012.
52. V.V Zakharov L.A. Petrosyan. Introduction to mathematical ecology. –L .: Publishing house of Leningrad State University, 1986.
53. V.V Zakharov L.A. Petrosyan. Mathematical models in ecology. –SPb .: Publishing house of St. Petersburg. University, 1996.
54. V. A. Gorelik A. F. Kononenko. Game-theoretic models of decision making in ecological and economic systems. –M., 1982.
55. V. V. Mazalov A. N. Rettieva. Equilibrium according to Nash in the problems of environmental protection // Mat. modeling. –2006. –Vol. 18, no. 5. –P. 73–90.
56. Rettieva A. N. Cooperative regulating condition in the problem of separation of biological resources // UBS. –2009. –Vol. 26.1. –P. 366–384.
57. Breton M., Zaccour G., Zahaf M. A differential game of joint implementation of environmental projects // Automatica. –2005. –Vol. 41, no. 10. –P. 1737–1749.
58. P. Chander H. Tulkens. The Core of an Economy with Multilateral Environmental Externalities // International Journal of Game Theory. –1997. –Vol. 23. –P. 379–401.
59. Chander P. The Gamma-Core and Coalition Formation // international journal of game theory. –2007. –no. 35. –P. 539–556.

60. P. Dasgupta R. Gilbert J. Stiglitz. Strategic considerations in invention and innovation: the case of natural resources // *Econometrica*. –1983. –Vol. 51. –P. 1439–1448.
61. E. J. Dockner A. Gaunersdorfer. On the Profitability of Horizontal Mergers in Industries with Dynamic Competition // *Japan and the World Economy*. –2001. –no. 4. –P. 195–216.
62. J. A. Filar P. S. Gaertner. A regional allocation of world CO2 emission reductions // *Mathematics and Computers in Simulation*. –1997. –Vol. 43. –P. 269–275.
63. B. C. Giri S. K. Goyal. Recent trends in modeling of deteriorating inventory // *European Journal of Operational Research*. –2001. –Vol. 134, no. 1. –P. 1–16.
64. A. Haurie G. Zaccour. Differential game models of global environmental management // *Annals of Dynamic Games, Boston*. –1994. –P. 124–132.
65. al A. Haurie.et. *Games and dynamic games*. –World Scientific Books, 2012.
66. S. Jørgensen G. Martin-Herran G. Zaccour. Agreeability and Time Consistency in Linear-State Differential Games // *Journal of Optimization Theory and Applications*. –2003. –Vol. 119, no. 1. –P. 49–63.
67. S. Jørgensen G. Zaccour. *Developments in Differential Game Theory and Numerical Methods: Economic and Management Applications // Computational Management Science*. –2007. –Vol. 4, no. 2. –P. 159–182.
68. S. Jørgensen G. Sorger. Feedback Nash equilibria in a problem of optimal fishery management // *Journal of Optimization Theory and Applications*. –1990. –Vol. 64. –P. 293–310.
69. Jorgensen S., Yeung D. W. K. Inter- and intragenerational renewable resource extraction // *Annals of Operations Research*. –1999. –Vol. 88. –P. 275–289.
70. S. Jørgensen G. Zaccour. Time-Consistent Side Payment in a Dynamic Game of Downstream Pollution // *Journal of Economic Dynamics and Control*. –2001. –Vol. 25. –P. 1973–1987.

71. Kaitala. V. Equilibria in a stochastic resource management game under imperfect information // *European Journal of Operational Research*. –1993. –Vol. 71. –P. 439–453.
72. V. Kaitala M. Pohjola. Sustainable international agreements on greenhouse warming // *A game theory study*, in Carraro and Filar (eds.), *Control and Game Theoretic Models of the Environment*, *Annals of the International Society of Dynamic Games*. –1995. –Vol. II.
73. J. Krawczyk G. Zaccour. Management of pollution from decentralized agents by the local government // *International Journal of Environment and Pollution*. –1999. –Vol. 12, no. 2/3. –P. 343–357.
74. Sorger G. Competitive dynamic advertising: a modification of the case game // *Journal of Economic Dynamics and Control*. –1989. –Vol. 13. –P. 55–80.
75. Zakharov V. One approach to allocating damage to environment. –*System Modelling and Optimization*, New York.: Springer-Verl., 1994.
76. Petrosian O. L. Looking Forward Approach in Cooperative Differential Games. // *International Game Theory Review*. –2016. –no. 2.
77. Petrosian O., Barabanov A. Looking Forward Approach in Cooperative Differential Games with Uncertain-Stochastic Dynamics // *Journal of Optimization Theory and Applications*. –2017. –Vol. 172, no. 1. –P. 328–347.
78. Petrosian O. L., Tur A. V. Hamilton-Jacobi-Bellman Equations for Non-cooperative Differential Games with Continuous Updating. // *Communications in Computer and Information Science*. –2019. –Vol. 1090. –P. 178–191.
79. Kuchkarov I. I., Petrosian O. L. On class of linear quadratic non-cooperative differential games with continuous updating. // *Lecture Notes in Computer Science*. –2019. –Vol. 11548. –P. 635–650.
80. Basar T., Olsder G. J. *Dynamic noncooperative game theory*. –London : Academic Press, 1995.

81. Shevkoplyas E. V. Optimal Solutions in Differential Games with Random Duration // Journal of Mathematical Sciences. –2014. –Vol. 199, no. 6. –P. 715–722.
82. Petrosyan L. A., Murzov N. V. Game-theoretic Problems in Mechanics // Lithuanian Mathematical Collection. –1966. –№ 3. –C. 423–433.
83. Pontryagin L. S. On Theory of Differential Games // Successes of Mathematical Sciences. –1966. –№ 26, 4 (130). –C. 219–274.
84. Goodwin G. C., Seron M. M., Dona J. A. Constrained Control and Estimation: An Optimisation Approach. –New York : Springer, 2005.
85. Kwon W. H., Han S. H. Receding Horizon Control: Model Predictive Control for State Models. –New York : Springer, 2005.
86. Rawlings J. B., Mayne D. Q. Model Predictive Control: Theory and Design. –Madison : Nob Hill Publishing, 2009.
87. Wang L. Model Predictive Control System Design and Implementation Using MATLAB. –New York : Springer, 2005.
88. Bemporad A., Morari M., et al. The explicit linear quadratic regulator for constrained systems // Automatica. –2002. –Vol. 38, no. 1. –P. 3–20.
89. Hempel A., Goulart P., Lygeros J. Inverse Parametric Optimization With an Application to Hybrid System Control // IEEE Transactions on Automatic Control. –2015. –Vol. 60, no. 4. –P. 1064–1069.
90. Kwon W., Bruckstein A., Kailath T. Stabilizing state-feedback design via the moving horizon method // 21st IEEE Conference on Decision and Control. –1982.
91. Kwon W., Pearson A. A modified quadratic cost problem and feedback stabilization of a linear system // IEEE Transactions on Automatic Control. –1977. –Vol. 22(5). –P. 838–842.
92. Mayne D., Michalska H. Receding horizon control of nonlinear systems // IEEE Transactions on Automatic Control. –1990. –Vol. 35(7). –P. 814–824.

93. Shaw L. Nonlinear control of linear multivariable systems via state-dependent feedback gains // IEEE Transactions on Automatic Control. –1979. –Feb. –Vol. 24, no. 1. –P. 108–112. –Conference Name: IEEE Transactions on Automatic Control.
94. Haurie A. Feedback equilibria in differential games with structural and modal uncertainties // Advances in Large Scale Systems. –1984.
95. Joao B. R. do Val, Basar T. Receding horizon control of jump linear systems and a macroeconomic policy problem // Journal of Economic Dynamics & Control. –1999. –Vol. 23. –P. 1099–1131.
96. Van Den Broek W. A. Moving horizon control in dynamic games // Journal of Economic Dynamics & Control. –2002. –Vol. 26. –P. 937–961.
97. Sethi S., Sorger G. A Theory of Rolling Horizon Decision Making // Annals of Operations Research. –1991. –Vol. 29. –P. 387–416.
98. Vasin A., Divtsova A. The repeated game modelling an agreement on protection of the environment // In VIII Moscow International Conference on Operations Research (ORM2018), MAKS Press Moscow. –2018. –Vol. 1. –P. 261–263.
99. Vasin A., Divtsova A. Game-theoretic model of agreement on limitation of transboundary atmospheric pollution // Mat. Teor. Igr Pril. –2017. –Vol. 9, no. 1. –P. 27–44.
100. Petrosian O. L., Gromova E. V., Pogochev S. V. Strong time-consistent subset of core in cooperative differential games with finite time horizon // Mathematical theory of games and its applications. –2016. –Vol. 8, no. 4. –P. 79–106.
101. Petrosian O. L., Gromova E. V., Pogochev S. V. Strong time-consistent subset of core in cooperative differential games with finite time horizon // Automation and Remote Control. –2018. –Vol. 79, no. 10. –P. 1912–1928.
102. Wolf D. A., Zakharov V. V., Petrosian O. L. On the existence of IDP-core in cooperative differential games // Mathematical theory of games and its applications. –2017. –Vol. 9, no. 4. –P. 18–38.

103. Zakharov V. V., O-Hun K. Linear programming approach in cooperative games // Korean Math. Soc. –1997. –Vol. 34, no. 2. –P. 423–435.
104. Petrosian O. L., Zakharov V. V. IDP-Core: Novel Cooperative Solution for Differential Games // Mathematics. –2020. –Vol. 8, no. 5.
105. Petrosian O. L., Pogozhev S. V. Looking Forward Approach with Random Horizon in Cooperative Differential Games // Contributions to Game Theory and Management. –2020. –Vol. 13. –P. 360–387.
106. Petrosian O. L., Kuchkarov I. I. About the Looking Forward Approach in Cooperative Differential Games with Transferable Utility // Frontiers of Dynamic Games: Game Theory and Management. –2019. –P. 175–208.
107. Petrosian O. L., Nastych M. A., Li Y. The Looking Forward Approach in a Differential Game Model of the Oil Market with Non-transferable Utility // Frontiers of Dynamic Games. –2020. –P. 215–244.
108. Gromova E., Petrosian O. Control of Informational Horizon for Cooperative Differential Game of Pollution Control // 2016 International Conference Stability and Oscillations of Nonlinear Control Systems (Pyatnitskiy's Conference). –2016. –P. 1–4.
109. Petrosian O. L., Gromova E. V. Cooperative Differential Games with Dynamic Updating // IFAC-PapersOnLine. –2018. –Vol. 51, no. 32. –P. 413–417.
110. Petrosian O. L., Nastych M. A., Volf D. A. Non-cooperative Differential Game Model of Oil Market with Looking Forward Approach // Frontiers of Dynamic Games. –2018. –Vol. 25. –P. 189–202.
111. Petrosian O. L., Nastych M. A., Volf D. A. Differential game of oil market with moving informational horizon and non-transferable utility. // Proceedings of the Constructive Nonsmooth Analysis and Related Topics (dedicated to the memory of VF Demyanov) (CNSA), Saint-Petersburg, Russia, 22-27 May. –2017. –P. 1–4.

112. Petrosian O. L. Looking Forward Approach in Cooperative Differential Games with infinite-horizon // Vestnik S.-Petersburg Univ. Ser. 10. Prikl. Mat. Inform. Prots. Upr. –2016. –no. 4. –P. 18–30.
113. Petrosian O. L., Shi L., et al. Moving Information Horizon Approach for Dynamic Game Models // Mathematics. –2019. –Vol. 7, no. 12. –P. 1–31.
114. Shi L., Petrosian O. L. A Dynamic Oligopoly Marketing Model of Advertising // Contributions to Game Theory and Management. –2018. –Vol. 11. –P. 207–223.
115. Shi L., Petrosian O. L., Boiko A. V. Looking Forward Approach for dynamic cooperative advertising game model. // Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes. –2019. –Vol. 15, no. 2. –P. 221–234.
116. Yeung D. W. K., Petrosian O. L. Cooperative Stochastic Differential Games with Information Adaptation // International Conference on Communication and Electronic Information Engineering (CEIE 2016). –2016. –P. 375–381.
117. Kuchkarov I. I., Petrosian O. L. Open-Loop Based Strategies for Autonomous Linear Quadratic Game Models with Continuous Updating // Lecture Notes in Computer Science. –2020. –Vol. 12095. –P. 212–230.
118. Petrosian O. L., Tur A. V., Wang Z. Cooperative differential games with continuous updating using Hamilton–Jacobi–Bellman equation // Optimization Methods and Software. –2020. –Vol. 1275. –P. 256–270.
119. Wang Z., Petrosian O. L. On class of non-transferable utility cooperative differential games with continuous updating // Journal of Dynamics and Games. –2020. –Vol. 7, no. 4. –P. 291–302.
120. Petrosian O. L., Tur A. V., Zhou J. Pontryagin’s Maximum Principle for Non-cooperative Differential Games with Continuous Updating // Communications in Computer and Information Science. –2020. –Vol. 1275. –P. 256–270.

121. Kuchkarov I. I. Feedback based strategies for autonomous linear quadratic cooperative differential games with continuous updating // Contributions to Game Theory and Management. –2020. –Vol. 13. –P. 244–251.
122. Kuchkarov I. I., Petrosian O. L. Non-autonomous Linear Quadratic Non-cooperative Differential Games with Continuous Updating (to be published) // Journal of Global Optimization. –2021.
123. Zhou J., Tur A. V., et al. Transferable Utility Cooperative Differential Games with Continuous Updating using Pontryagin Maximum Principle // Mathematics. –2020. –Vol. 9, no. 2.
124. Petrosian O. L., Tikhomirov D. S., Gao H. Differential Game Model of Resource Extraction with Continuous and Dynamic Updating (to be published) // Journal of Systems Science and Complexity. –2021.
125. Petrosian O. L., Inga J., et al. Optimal Control and Inverse Optimal Control with Continuous Updating for Human Behavior Modeling // IFAC-PapersOnLine. –2020. –Vol. 53, no. 2. –P. 6670–6677.
126. Petrosian O. L., Tikhomirov D. S., Kuchkarov I. I. et al. About One Differential Game Model with Dynamic Updating // Automation and Remote Control. –2020. –no. 81. –P. 1733–1750.
127. Petrosian O. L., Tur A. V., et al. Cooperative differential game model with continuous updating // Mat. Teor. Igr Pril. –2020. –Vol. 12, no. 2. –P. 82–109.
128. Petrosian O. L., Tikhomirov D. S., et al. About one differential game model with dynamic updating // Mat. Teor. Igr Pril. –2019. –Vol. 11, no. 2. –P. 40–67.
129. Yeung D. W. K., Petrosian O. L. Infinite Horizon Dynamic Games: A New Approach via Information Updating // International Game Theory Review. –2017. –Vol. 19, no. 4. –P. 1–23.
130. Rufus I. Differential Games. –New York : John Wiley and Sons, 1965.

131. Fleming W. H. The convergence problem for differential games // Adv. in game theory, Ann. of Math. Studies. –1964. –P. 175–195.
132. Krasovsky N. N. Control of Dynamical System. Problem of minimum guaranteed result. –Moscow : Science, 1985.
133. Kononenko A. F. On equilibrium positional strategies in non-antagonistic differential games // Reports of the USSR Academy of Sciences. –1976. –T. 231, № 2. –C. 285–288.
134. Lukyanenko A., Gurtov A. Towards behavioral control in multi-player network games // 2009 International Conference on Game Theory for Networks. –2009. –P. 1–8.
135. Filippov A. F. Introduction to the theory of differential equations (in Russian). –Moscow : Editorial URSS, 2004.
136. Gillies D. B. Some theorems on n person games // Ph.D. thesis. –1953.
137. Edgeworth F. Y. Mathematical physics. // Kegan Paul, London. –1881.
138. Nikolaevskiy I., Lukyanenko A., Gurtov A. Nash Bargaining Solution Allocation is Not Suitable for Datacenter Jobs // International Game Theory Review. –2017. –Vol. 19, no. 1. –P. 1–22.
139. Kleimenov A. F. To the cooperative theory of non-coalition positional games // Reports of the USSR Academy of Sciences, Moscow. –1990.
140. Kleimenov A. F. Cooperative solutions in the position differential game of many individuals with continuous payment functions // Applied Mathematics and Mechanics. –1990. –Vol. 54, no. 3. –P. 389–394.
141. Yeung D. W. K., Petrosyan L. A. Subgame Consistent Economic Optimization: An Advanced Cooperative Dynamic Game Analysis // Springer, New York. –2012.
142. Petrosyan L. A., Danilov N. N. Cooperative differential games and their applications. –Publishing house of Tomsk University, 1985.

143. Tolwinski B., Haurie A., Leitmann G. Cooperative equilibria in differential games // Journal of Mathematical Analysis and Applications. –1986. –Vol. 119, no. 1. –P. 182–202.
144. Zaccour G. Time Consistency in Cooperative Differential Games: A Tutorial // INFOR: Information Systems and Operational Research. –2008. –Vol. 46, no. 1. –P. 81–92.
145. Gao H., Petrosian L. A., et al. Cooperation in two-stage games on undirected networks // Journal of Systems Science and Complexity. –2017. –Vol. 30(3). –P. 680–693.
146. Gromova E. V, Petrosyan L. A. Strongly dynamically stable cooperative solution in one differential game of harmful emissions management // Managing large systems. –2015. –Vol. 55. –P. 140–159.
147. Vorob'ev N. N. Game Theory. Lectures for Economists and Systems Scientists. –1977.
148. Scarf H. E. The core of an n person game // *Economica*. –1967. –Vol. 1, no. 35. –P. 50–69.
149. Billera L. J. Some theorems on the core of n person game // *SIAM Journal of Applied Mathematics*. –1970. –Vol. 18, no. 3. –P. 567–579.
150. Shapley L. S. On balanced games without side payments // *Mathematical programming*. –1972. –P. 261–290.
151. Bondareva O. N. Some applications of linear programming methods to the theory of cooperative games // *Problems of Cybernetics*. –1963. –№ 10. –C. 119–140.
152. Shapley L. S. On balanced sets and cores // *Naval Research Logistic Quarterly*. –1967. –Vol. 14. –P. 453–460.
153. Zakharov V. V., Dementieva M. Multistage Cooperative Games and Problem of Time Consistency // *International Game Theory Review*. –2004. –Vol. 6, no. 1. –P. 157–170.

154. Zakharov V. V., Akimova A. N. Geometric Properties of the Core, Subcore, Nucleolus. –2002. –Vol. 8. –P. 281–289.
155. Owen G. Game theory. –New York : Academic Press, 1982.
156. Von Neumann J., Morgenstern O. Theory of Games and Economic Behavior. –Princeton : Princeton University Press, 1970.
157. Haurie A. A note on nonzero-sum differential games with bargaining solutions // Journal of Optimization Theory and Applications. –1976. –Vol. 18, no. 1. –P. 31–39.
158. Petrosyan L. A. Time consistency of solutions for differential games with many participants // Vestnik of Leningrad University. Ser. 1: Mathematics, Mechanics, Astronomy. –1977. –№ 4. –C. 46–52.
159. Petrosyan L. A. Strong time consistent differential optimality principles // Vestnik of Leningrad University. Ser. 1: Mathematics, Mechanics, Astronomy. –1993. –№ 4. –C. 35–40.
160. Petrosyan L. A., Danilov N. N. Time consistency of solutions for non-antagonistic differential games with transferable payoffs // Vestnik of Leningrad University. Ser. 1: Mathematics, Mechanics, Astronomy. –1979. –№ 1. –C. 52–59.
161. Petrosyan L. A. Strongly time-consistent differential optimality principles // Vestnik St. Petersburg Univ. Math. Series 1. Mathematics. Mechanics. Astronomy. –1993. –Vol. 26. –P. 40–46.
162. Petrosyan L. A., Danilov N. N. Cooperative differential games and their applications. –1985.
163. Parilina E. M., Zaccour G. Node-Consistent Shapley Value for Games Played over Event Trees with Random Terminal Time // Journal of Optimization Theory and Applications. –2017. –Vol. 175, no. 1. –P. 236–254.
164. Parilina E. M., Zaccour G. Node-consistent core for games played over event trees // Automatica. –2015. –Vol. 53. –P. 304–311.

165. Yeung D., Petrosyan L. A. Subgame Consistent Solutions of a Cooperative Stochastic Differential Game with Nontransferable Payoffs // Journal of optimization theory and applications. –2005. –Vol. 124, no. 3. –P. 701–724.
166. Petrosyan L. A., Yeung D. A Time-Consistent Solution Formula for Bargaining Problem in Differential Games // Journal of optimization theory and applications. –2014. –Vol. 16, no. 4. –P. 1–11.
167. Petrosyan L. A., Danilov N. N. Stability of solutions in non-zero sum differential games with transferable payoffs // Vestnik Leningrad State University. Series 1. Mathematics. Mechanics. Astronomy. –1979. –P. 52–59.
168. Dockner E., Jorgensen S., et al. Differential Games in Economics and Management Science // Cambridge University Press, Cambridge. –2001.
169. Yeung D. W. K., Petrosyan L. A. Subgame Consistent Economic Optimization: An Advanced Cooperative Dynamic Game Analysis // Springer Science & Business Media: Berlin/Heidelberg, Germany. –2011.
170. Owen G. Game Theory. –Emerald Group Publishing Limited: Bingley, UK, 2013.
171. Nash J. F. Non-cooperative games // Ann. Math. –1951. –no. 54. –P. 286–295.
172. Yeung D. W. K. On differential games with a feedback Nash equilibrium. // J. Optim. Theory Appl. –1994. –no. 82. –P. 181–188.
173. Bellman R. Dynamic Programming. –2003.
174. Chander P., Tulkens H. A core-theoretic solution for the design of cooperative agreements on transfrontier pollution. // Public Goods, Environmental Externalities and Fiscal Competition. –2006. –P. 176–193.
175. Petrosjan L. A. Differential Pursuit Games // World Scientific Publ. Co. Pte. Ltd.: Singapore. –1993.
176. Yeung D. W. K., Petrosjan L. A. Cooperative Stochastic Differential Games // Springer Science & Business Media: Berlin/Heidelberg, Germany. –2006.

177. Petrosyan L. A., Danilov N. N. Stability of solutions in non-zero sum differential games with transferable payoffs. // Vestn. Leningr. Univ. –1979. –no. 1. –P. 52–59.
178. Yeung D. W. K., Petrosyan L. A. Subgame consistent cooperative solution of dynamic games with random horizon. // J. Optim. Theory Appl. –2011. –no. 150. –P. 78–97.
179. Prasad A., Sethi S. P., Naik P. A. Optimal Control of an Oligopoly Model of Advertising. // IFAC Proc. Vol. –2009. –no. 42. –P. 91–96.
180. Basar T., Olsder G. J. Dynamic Noncooperative Game Theory // London, Academic Press. –1995.
181. Krasovskii N. N., Kotel'nikova A. N. On a differential interception game // Proceedings of the Steklov Institute of Mathematics. –2010. –no. 268. –P. 161–206.
182. Bellman R. Dynamic Programming // Princeton. –1957.
183. Subbotin A. I. Generalization of the main equation of differential game theory // Journal of Optimization Theory and Applications. –1984. –Vol. 43. –P. 103–133.
184. Subbotin A. I. Generalized Solutions of First Order PDEs // Basel, Birkhauser. –1995.
185. Chander P., Tulkens H. A core-theoretic solution for the design of cooperative agreements on transfrontier pollution // International Tax and Public Finance. –1995. –Vol. 2, no. 2. –P. 279–293.
186. Chander P. The gamma-core and coalition formation // International Journal of Game Theory. –2007. –Vol. 35, no. 4. –P. 539–556.
187. Petrosyan L. A., Danilov N. N. Stability of solutions in non-zero sum differential games with transferable payoffs // Vestnik Leningrad State University. –1979. –Vol. 1. –P. 52–59.
188. Yeung D. W. K., Petrosian L. A. Subgame-consistent Economic Optimization // New York, Springer. –2012.

189. Yeung D. W. K., Petrosyan L. A. Subgame-consistent Economic Optimization // Springer. –2012.
190. Petrosyan L. A., Zaccour G. Time-consistent Shapley value allocation of pollution cost reduction // J. Econ. Dyn. Control. –2003. –Vol. 3. –P. 381–398.
191. Petrosyan L. A., Danilov N. N. Stability of solutions in non-zero sum differential games with transferable payoffs // Vestn. Leningr. Univ. –1979. –Vol. 1. –P. 52–59.
192. Petrosyan L. A. Time-consistency of solutions in multi-player differential games // Vestnik Leningrad State University. –1977. –T. 4. –C. 46–52.
193. Jorgensen S., Yeung D. W. K. Inter- and intragenerational renewable resource extraction // Annals of Operations Research. –1999. –P. 275–289.
194. Shevkoplyas E. V. Time-consistency Problem Under Condition of a Random Game Duration in Resource Extraction // Contributions to Game Theory and Management. –2009. –Vol. 2. –P. 461–473.
195. Shevkoplyas E. V. Stable cooperation in differential games with random duration // Mat. Teor. Igr Pril. –2010. –Vol. 2, no. 3. –P. 79–105.
196. Shevkoplyas E. V. The Shapley value in cooperative differential games with random duration // Ann. Dyn. Games. –2011.
197. Shevkoplyas E. V. The Hamilton-Jacobi-Bellman Equation for a Class of Differential Games with Random Duration // Automation and Remote Control. –2014. –Vol. 75. –P. 959–970.
198. Petrosjan L. A., Shevkoplyas E. V. Cooperative solutions for games with random duration // Game Theory Appl. –2003. –Vol. 9. –P. 125–139.
199. Kostyunin S. Yu., Shevkoplyas E. V. On simplification of integral payoff in differential games with random duration // Vestnik S.-Petersburg Univ. Ser. 10. Prikl. Mat. Inform. Prots. Upr. –2011. –Vol. 4. –P. 47–56.
200. Eiben A. E., Smith J. E. Introduction to Evolutionary Computing // Berlin, Springer. –2003.

201. Dahl C., Yucel M. Testing alternative hypotheses of oil producer behavior // *The Energy Journal*. –1991. –Vol. 12, no. 4. –P. 117–138.
202. Polasky S. Do oil producers act as 'oil'igopolists? // *Journal of Environmental Economics and Management*. –1992. –Vol. 23, no. 3. –P. 216–247.
203. Danielsen A. L., Kim S. OPEC stability: An empirical assessment // *Energy Economics*. –1998. –Vol. 10, no. 3. –P. 174–184.
204. Smith J. L. Inscrutable OPEC? Behavioral tests of the cartel hypothesis // *The Energy Journal*. –2005. –Vol. 26, no. 1. –P. 51–82.
205. Youhanna S. J. A Note on Modelling Opec Behavior 1983-1989: A Test of the Cartel and Competitive Hypotheses // *The American Economist*. –1994. –Vol. 38, no. 2. –P. 78–84.
206. Gulen S. G. Is OPEC a cartel? Evidence from cointegration and causality tests // *The Energy Journal*. –1996. –Vol. 17, no. 2. –P. 43–57.
207. Molchanov P. A statistical analysis of OPEC quota violations // *Economics*. Duke University, Durham, NC. –2003. –P. 1–31.
208. Bockem S. Cartel formation and oligopoly structure: a new assessment of the crude oil market // *Applied Economics*. –2004. –Vol. 36, no. 12. –P. 1355–1369.
209. Plaut S. E. OPEC is Not a Cartel // *Challenge*. –1981. –Vol. 24, no. 5. –P. 18–24.
210. Al-Sultan A. M. Alternative Models for OPEC Behavior // *Journal of Energy and Development*. –1995. –Vol. 18, no. 2. –P. 263–281.
211. Moran T. H. Managing an oligopoly of would-be sovereigns: the dynamics of joint control and self-control in the international oil industry past, present, and future // *International Organization*. –1987. –Vol. 41, no. 4. –P. 575–607.
212. Krasner S. D. The great oil sheikdown // *Foreign Policy*. –1973. –no. 13. –P. 123–138.

213. Ludkovski M., Sircar R. Differential Games // Commodities, Energy and Environmental Finance. Fields Institute Communications. –2015. –Vol. 74. –P. 317–333.
214. Salant S. Exhaustible Resources and Industrial Structure: A Nash-Cournot Approach to the World Oil Market // The Journal of Political Economy. –1976. –Vol. 84. –P. 1079–1094.
215. Berger K., Hoel M., et al. The Oil Market as an Oligopoly. Discussion paper // Central Bureau Statistics. –1988. –Vol. 32. –P. 1–19.
216. Cellini R., Lambertini L. A differential oligopoly game with differentiated goods and sticky prices // European Journal of Operational Research. –2007. –Vol. 176. –P. 1131–1144.
217. Basar T., Olsder G. J. Dynamic Noncooperative Game Theory. –Philadelphia, Pennsylvania : SIAM, 1999.
218. Fershtman C., Nitzan S. Dynamic voluntary provision of public goods // European Economic Review. –1991. –Vol. 35, no. 5. –P. 1057–1067.
219. Wirl F. Dynamic voluntary provision of public goods: extension for nonlinear strategies // European Journal of Political Economy. –1996. –Vol. 12, no. 3. –P. 555–560.
220. Basar T., Olsder G. J. Dynamic Noncooperative Game Theory // SIAM: Philadelphia, United States. –1999. –Vol. 23.
221. Leitmann G., Schmitendorf W. E. Some Sufficiency Conditions for Pareto Optimal Control // ASME Journal of Dynamical Systems, Measurement, and Control. –1973. –Vol. 95, no. G. –P. 356–361.
222. Jørgensen S. A Differential Games Solution to a Logarithmic Advertising Model // Journal of the Operational Research Society. –1982. –no. 33. –P. 425–432.
223. Jørgensen S. Sufficiency and game structure in Nash open-loop differential games // Journal of Optimization Theory and Applications. –1986. –Vol. 1, no. 50. –P. 189–193.

224. Engwerda J. LQ Dynamic Optimization and Differential Games. –New York : Willey, 2005.
225. Eisele T. Nonexistence and nonuniqueness of open-loop equilibria in linear-quadratic differential games // Journal of Optimization Theory and Applications. –1982. –Aug. –Vol. 37, no. 4. –P. 443–468. –Access mode: <https://doi.org/10.1007/BF00934951>.
226. Van Aarle B., Bovenberg A., Raith M. Monetary and Fiscal Policy Interaction and Government Debt Stabilization // Journal of Economics. –1995. –06. –Vol. 62. –P. 111–140.
227. Polyanin A., Zaitsev V. Handbook of Exact Solutions for Ordinary Differential Equations. –Boca Raton : CRC Press, 2002. –10. –ISBN: 9781584882978.
228. Petrosyan L. A., Zaccour G. Time-consistent Shapley value allocation of pollution cost reduction // J. Econ. Dyn. Control. –2003. –Vol. 27, no. 3. –P. 381–398.
229. Shapley L. A value for n-person games // In: H. W. Kuhn and A. W. Tucker, Eds., Contributions to the Theory of Games II, Annals of Mathematics Studies. –1953. –Vol. 28. –P. 307–317.
230. Petrosjan L. A., Zaccour G. Time-consistent Shapley value allocation of pollution cost reduction // Journal of Economic Dynamics & Control. –2003. –Vol. 27. –P. 381–398.
231. Başar T. Dynamic Noncooperative Game Theory, 2nd Edition. // Siam. –1999. –Vol. 23.
232. Leitmann G., Schmitendorf W. Some Sufficiency Conditions for Pareto-Optimal Control // Journal of Dynamic Systems, Measurement, and Control. –1973. –Vol. 95. –P. 356–361.
233. Long N. V. Pollution control: a differential game approach // Annals of Operations Research. –1992. –Vol. 37. –P. 283–296.
234. Yeung D. W. K., Petrosyan L. A. et al. Subgame Consistent Cooperation // Theory and Decision Library C. –2016.

235. Mombaur K., Truong A., Laumond J. P. From human to humanoid locomotion—an inverse optimal control approach // *Autonomous Robots*. —2010. —Apr. —Vol. 28, no. 3. —P. 369–383.
236. Flad M., Otten J., et al. Steering driver assistance system: A systematic cooperative shared control design approach // *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. —2014. —. —P. 3585–3592.
237. Cole D. J., Pick A. J., Odhams A. M. C. Predictive and linear quadratic methods for potential application to modelling driver steering control // *Vehicle System Dynamics*. —2006. —Mar. —Vol. 44, no. 3. —P. 259–284.