



МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ М. В. ЛОМОНОСОВА

Факультет вычислительной математики и кибернетики

Кафедра системного анализа

Егоров Кирилл Юлианович

# Математическое моделирование движений руки, держащей предмет

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

*Научный руководитель*

к.ф.-м.н., доцент И. В. Востриков

Москва, 2023

# Содержание

<b>1</b>	<b>Введение . . . . .</b>	<b>3</b>
<b>2</b>	<b>Математическое моделирование . . . . .</b>	<b>5</b>
2.1	Планарная модель руки, держащей предмет . . . . .	5
2.2	Уравнение динамики . . . . .	6
2.3	Учёт энергетических затрат . . . . .	9
<b>3</b>	<b>Постановка задачи . . . . .</b>	<b>10</b>
3.1	Непрерывная постановка задачи . . . . .	10
3.2	Дискретизация задачи . . . . .	11
<b>4</b>	<b>Синтез оптимального управления . . . . .</b>	<b>12</b>
4.1	Описание метода . . . . .	12
4.2	Синтез оптимальной поправки . . . . .	13
4.3	Регуляризация оптимальной поправки . . . . .	16
4.4	Алгоритм . . . . .	18
<b>5</b>	<b>Синтез начального референсного управления . . . . .</b>	<b>21</b>
5.1	Описание метода . . . . .	21
5.2	Синтез управления . . . . .	22
5.3	Алгоритм . . . . .	25
<b>6</b>	<b>Примеры . . . . .</b>	<b>28</b>
6.1	Целевое положение схвата . . . . .	28
6.2	Целевое положение и скорость схвата . . . . .	29
6.3	Обход препятствия . . . . .	30
<b>7</b>	<b>Заключение. . . . .</b>	<b>33</b>
	Список литературы . . . . .	35
	Приложения . . . . .	37
<b>A</b>	<b>Таблица параметров . . . . .</b>	<b>37</b>

# 1 Введение

Работа посвящена построению оптимальных траекторий движения в модели бионической руки человека, держащего предмет. Целью работы является построение эффективного вычислительного метода для управления сложными биомеханическими системами. Для разработки и тестирования построенного метода была предложена соответствующая математическая модель.

Исследования в области биологического движения имеют огромное практическое значение: они позволяют частично восстановить двигательную функцию у людей с ограниченными возможностями, чем существенно улучшить качество их жизни. Последние технические достижения в области роботизированных протезов и функциональной электронной стимуляции парализованных мышц позволяют начать внедрение данной теории. Более того, сложные бионические устройства становятся бесполезными без соответствующего знания о грамотном управлении ими.

Актуальность данной сферы исследований подтверждается наличием большого числа работ последнего времени, улучшающих существующие методы решения задач, связанных с биологическим движением, таких как [6], [16], а также работ, предлагающих новые математические модели для отдельных аспектов движения, например, [8].

В работе предложена математическая модель руки человека, держащего предмет, как планарного трёхсекционного математического маятника. Выведена динамика данной физической системы. Для возможности анализа поведения системы методами оптимального управления постулирован принцип оптимальности и рассмотрены представленные в литературе ([4], [15]) формализации оптимизационного энергетического критерия.

Использование методов оптимального управления способно привести фундаментальным открытиям в области биологической моторики: от описания свойств функций отдельных мышц, до исследования контроля мышц нервной системой при выполнении целевых задач, — поскольку данные методы напрямую работают с причинами движений, выраженными в форме оптимизационных критериев.

По результатам математического моделирования была поставлена задача оптимального целевого управления нелинейной системой в непрерывной и дискретной формах. В качестве управления выбрано изменение крутящего

момента каждого из сочленений. В работе полагается отсутствие ограничений на управление и известное полное фазовое состояние системы.

Для решения задачи в дискретной постановке были рассмотрены известные базовые методы решения задачи оптимального управления нелинейными системами [11], [5]. В качестве основного метода в данной работе применяется итеративный метод, предполагающий последовательное построение серии линейно-квадратичных регуляторов для системы и функционала качества, аппроксимированных вдоль заданной траектории.

При разработке метода особое внимание было уделено аспектам, не достаточно подробно изложенным в литературе: способу регуляризации оптимальной поправки и способу построения начальной референсной траектории без опоры на мнение эксперта в предметной области.

Полученное в результате программное решение, реализующее предложенный метод, было применено для рассмотрения конкретных постановок задачи — классических задач биомеханического движения: задача целевого положения схвата, задача обхода препятствия. Они служат возможности сравнения результатов работы предложенного метода с имеющимися в литературе, например, с [10], [14].

Результаты, полученные в рамках работы над диссертацией, были представлены в качестве доклада на научной конференции «Ломоносовские чтения – 2023» [17].

## 2 Математическое моделирование

### 2.1 Планарная модель руки, держащей предмет

Рассмотрим руку человека, держащего стержень. В некотором приближении можно считать, что мы имеем трёхсекционный математический маятник. Для каждого из 3-х сочленений нам известны:

1. Масса сочленения  $m_i$ ,  $i = 1, 2, 3$ ;
2. Линейная плотность сочленения  $\rho_i = \rho_i(x)$ ,  $0 \leq x \leq l_i$ ,  $i = 1, 2, 3$ ;
3. Длина сочленения  $l_i$ ,  $i = 1, 2, 3$ ;
4. Угол поворота сочленения  $\theta_i$ ,  $i = 1, 2, 3$  относительно оси абсцисс  $Oe_1$ .

Также считаем, что положение плечевого сустава фиксировано для определённости в точке  $(0, 0)$ . На Рис. 1 приведена схема с примером данного маятника и соответствующая позиция человека.

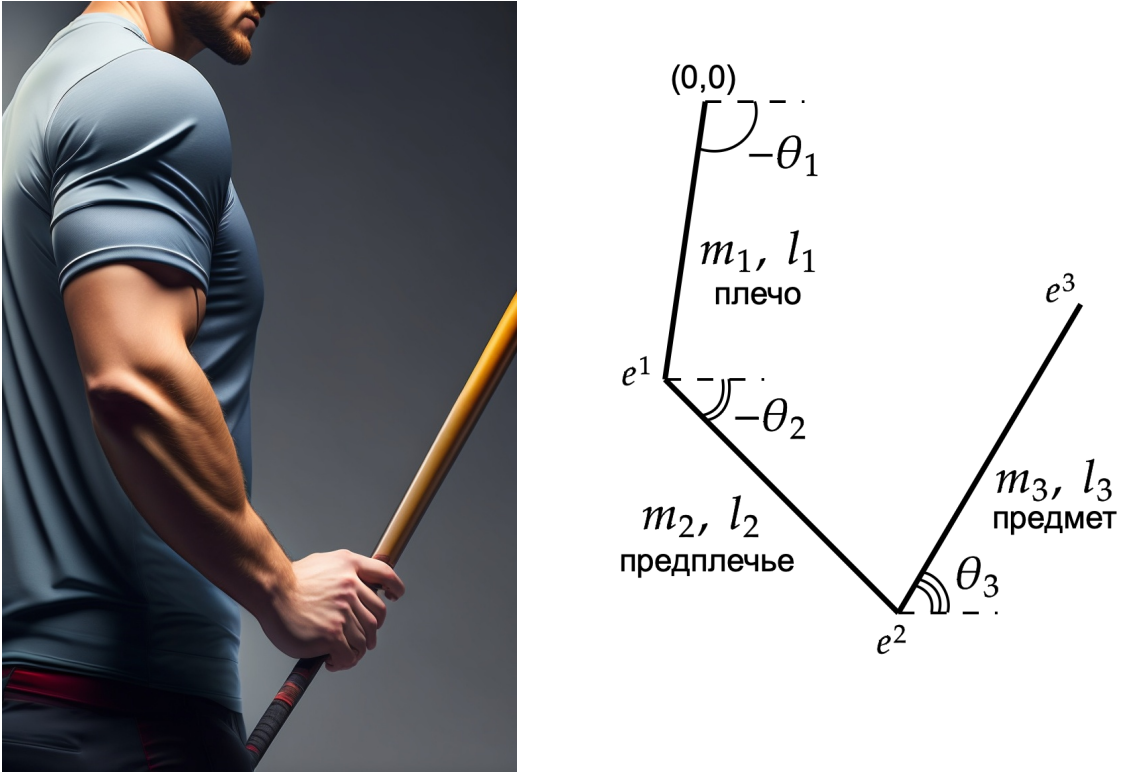


Рис. 1: Иллюстрация предложенной модели. Рисунок слева сгенерирован нейросетью *Lexica Aperture* по текстовому запросу и приведён для визуального соответствия сочленений маятника на схеме с частями тела человека.

В связи с тем, что мы имеем пространство состояний высокой размерности, далее в работе при построении графиков численного решения нам будет удобно рассматривать не траектории фазовых переменных по отдельности, а траекторию *схвата* в картезианской системе координат.

**Определение 1.** *Схватом* будем называть крайнюю точку последнего, в нашем случае третьего, сочленения приведённого маятника. Позицию схвата в картезианской системе координат обозначим за

$$e^3 \in \mathbb{R}^2.$$

Сразу выпишем выражения для позиций крайних точек сочленений:

$$e^i = \begin{bmatrix} \sum_{k=1}^i l_k \cos \theta_k \\ \sum_{k=1}^i l_k \sin \theta_k \end{bmatrix}, \quad i = 1, 2, 3. \quad (2.1)$$

## 2.2 Уравнение динамики

Для получения уравнения динамики рассматриваемой физической системы воспользуемся методом Эйлера–Лагранжа [19]. Идея метода состоит в проведении следующих последовательных шагов:

1. Выбор обобщённых координат;
2. Получение выражения для кинетической  $K$  и потенциальной  $\Pi$  энергий системы, записанных в обобщённых координатах;
3. Получение выражения для лагранжиана системы  $\mathcal{L}$ ;
4. Составление системы уравнений движения, соответствующих каждой обобщённой координате.

Обобщёнными координатами для нашей системы выберем углы поворота сочленений  $\theta_i$ ,  $i = 1, 2, 3$ . Далее перейдем к выражению энергий через обобщенные координаты. Для подсчета кинетической энергии воспользуемся теоремой Кёнинга [18].

**Теорема 1 (Кёнинг).** *Кинетическая энергия тела есть энергия поступательного движения центра масс плюс энергия вращательного движения относительно центра масс*

$$K = \frac{1}{2} m \|v_c\|^2 + \frac{1}{2} \omega^T I \omega, \quad (2.2)$$

где  $m$  — полная масса тела,  $I$  — тензор инерции тела,  $v_c$  — линейная скорость центра масс,  $\omega$  — скорость вращения тела относительно центра масс.

Далее в работе мы будем полагать, что каждое из сочленений представляет собой однородный стержень длины  $l_i$  массы  $m_i$ . В таком случае получаем следующие значения для положения центра масс  $c^i \in \mathbb{R}^2$   $i$ -ого сочленения:

$$c^i = \begin{bmatrix} \sum_{j=1}^{i-1} l_j \cos \theta_j + \frac{l_i}{2} \cos \theta_i \\ \sum_{j=1}^{i-1} l_j \sin \theta_j + \frac{l_i}{2} \sin \theta_i \end{bmatrix}.$$

Выражения для момента инерции и скорости вращательного движения относительно центра масс для стержня получаются соответственно:

$$I_i = \int_{(m_i)} r^2 dm = \rho_i \int_{(l_i)} r^2 dl = \frac{m_i l_i^2}{12},$$

$$\omega_i = 2\dot{\theta}_i.$$

Потенциальная энергия  $i$ -ого сочленения рассчитывается по формуле

$$\Pi_i = m_i g c_2^i,$$

где  $g$  — ускорение свободного падения.

Общая кинетическая и потенциальная энергии системы рассчитываются как сумма энергий каждого из сочленений:

$$K = \sum_{i=1}^3 K_i = \sum_{i=1}^3 \left( \frac{m_i \|\dot{c}^i\|^2}{2} + \frac{m_i l_i^2 |\dot{\theta}_i|^2}{6} \right),$$

$$\Pi = \sum_{i=1}^3 \Pi_i = \sum_{i=1}^3 m_i g c_2^i.$$

Теперь введём лагранжиан системы

$$\mathcal{L} = K - \Pi$$

и построим систему уравнений Эйлера–Лагранжа:

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\theta}_i} - \frac{\partial \mathcal{L}}{\partial \theta_i} = \tau_i, \quad i = 1, 2, 3, \quad (2.3)$$

где  $\tau_i$  — момент силы, действующий на  $i$ -ое сочленение, который доступен для управления.

Продифференцировав члены из левой части уравнения (2.3), получим уравнение динамики для рассматриваемой системы:

$$M(\theta)\ddot{\theta} + L(\theta, \dot{\theta}) = \tau, \quad (2.4)$$

где  $M(\theta) = M^T(\theta) \in \mathbb{R}^{3 \times 3}$  — матрица инерции системы,  $L(\theta, \dot{\theta}) \in \mathbb{R}^3$  — вектор центробежных и кориолисовых сил. На Рис. 2 приведено численное моделирование свободного падения системы в соответствии с полученным уравнением динамики (2.4).

*Замечание 1.* Матрица инерции  $M(\theta)$  является положительно-определённой, поскольку кинетическая энергия системы  $K$  всегда неотрицательна:

$$K(\theta, \dot{\theta}) = \frac{1}{2} \langle \dot{\theta}, M(\theta)\dot{\theta} \rangle > 0 \text{ для любого } \dot{\theta} \neq 0.$$

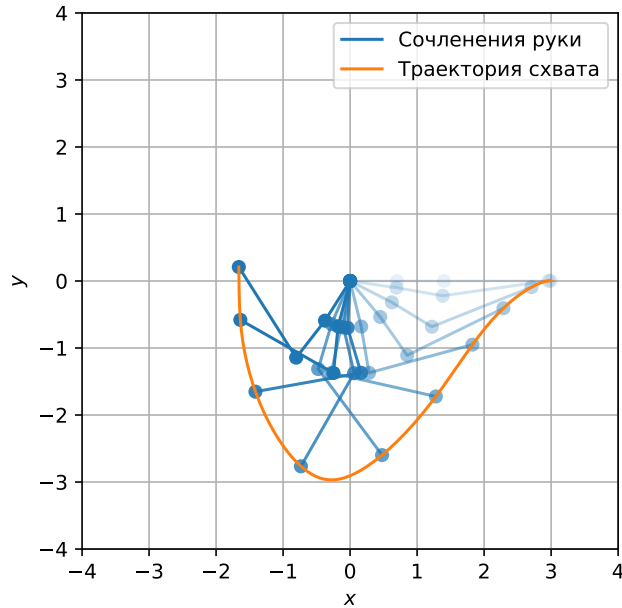


Рис. 2: Траектория системы в свободном падении из начального положения  $\theta^{\text{start}} = [0 \ 0 \ 0]^T$ ,  $\dot{\theta}^{\text{start}} = [0 \ 0 \ 0]^T$  на временном интервале  $0 \leq t \leq 1$ . Положения системы, соответствующие более раннему времени, показаны бледнее. Для изображения положений используется равномерное разбиение временного интервала.



## 2.3 Учёт энергетических затрат

Для моделирования биологического движения необходимо выяснить, какими принципами руководствуется мозг при выборе траектории для некоторого целевого движения. Существует бесконечное число возможных путей и профилей скорости для перемещения руки из одной точки в другую, и каждая траектория может быть достигнута несколькими возможными комбинациями углов между сочленениями. При этом нервная и моторно-двигательные системы человека для выбора одной конкретной траектории анализируют большой объем информации, поступающий от всех органов чувств.

В силу того, что нервная система человека есть результат оптимизационных процессов: эволюции, адаптации к условиям среды, обучения, мы постулируем следующий биологический принцип оптимальности.

**Утверждение 1** (Биологический принцип оптимальности). *Выбираемые нервной системой схемы движения являются оптимальными для поставленной задачи.*

Применение данного принципа позволяет не только моделировать движения методами оптимального управления, но и анализировать их причины.

В работе [13] было показано, что оптимизации проводятся с целью уменьшения затрат энергии. Однако общего подхода к формализации энергетических затрат пока не выработано. Так, например, в работе [4] предлагается минимизировать *рывок* схвата, то есть

$$\int_{t_{\text{start}}}^{t_{\text{final}}} \left\| \frac{d^3 e^3}{dt^3} \right\|^2 dt \longrightarrow \min,$$

а в работе [15] — изменение крутящего момента

$$\int_{t_{\text{start}}}^{t_{\text{final}}} \left\| \frac{d\tau}{dt} \right\|^2 dt \longrightarrow \min. \quad (2.5)$$

Причём существуют и другие менее популярные варианты, например, [3].

Мы будем использовать для формализации энергетических затрат выражение (2.5), поскольку данный критерий напрямую зависит от динамики руки и лучше согласуется с эмпирическими данными, чем модель минимального рывка [2].

## 3 Постановка задачи

### 3.1 Непрерывная постановка задачи

Поставим задачу целевого управления для модели, построенной в Разделе 2. Для этого рассмотрим расширенное фазовое пространство с состоянием

$$x = [\theta \ \dot{\theta} \ \tau]^T \in \mathbb{R}^9.$$

Тогда уравнение динамики системы (2.4) можно переписать в виде стационарной системы однородных дифференциальных уравнений

$$\dot{x} = A(x) + Bu, \quad (3.1)$$

где

$$A(x) = \begin{bmatrix} x_2 \\ M^{-1}(x_1)(x_3 - L(x_1, x_2)) \\ O \end{bmatrix}, \quad B = \begin{bmatrix} O \\ O \\ I \end{bmatrix}.$$

Считаем, что для данной системы поставлена задача Коши, то есть нам известно начальное состояние системы

$$x(t_{\text{start}}) = x^{\text{start}}. \quad (3.2)$$

*Замечание 2.* Отметим, что для выполнения достаточных условий существования и единственности решения Каратеодори для задачи Коши (3.1)-(3.2) управление  $u$  достаточно брать из класса измеримых на рассматриваемом отрезке  $t_{\text{start}} \leq t \leq t_{\text{final}}$  функций [21].

Для задачи Коши (3.1)-(3.2) поставим задачу поиска управления  $u \in U[t_{\text{start}}, t_{\text{final}}]$ , минимизирующего функционал качества вида:

$$J = q^{\text{final}}(x(t_{\text{final}})) + w_1 \int_{t_{\text{start}}}^{t_{\text{final}}} q(x(t)) dt + w_2 \int_{t_{\text{start}}}^{t_{\text{final}}} r(u(t)) dt, \quad (3.3)$$

где  $q^{\text{final}}$ ,  $q$  отвечают за терминальное и фазовые ограничения соответственно и выбираются в зависимости от конкретной постановки задачи, а  $r$  отвечает за энергетические затраты и в соответствии с (2.5) равна:

$$r(u) = \|u\|^2,$$

а  $w_1, w_2$  — веса соответствующих критериев для данной многокритериальной задачи.

Для дальнейших рассуждений потребуем, чтобы функции  $q^{\text{final}}, q$  были дважды дифференцируемыми. Полученные из модели функции  $A$  и  $r$  заведомо удовлетворяют этому требованию.

## 3.2 Дискретизация задачи

Для удобства дальнейших рассуждений дискретизируем задачу (3.1)-(3.2)-(3.3) по времени  $t_{\text{start}} \leq t \leq t_{\text{final}}$ . Для этого введем равномерную сетку с шагом  $\Delta t$ :

$$\{t_i\}_{i=1}^{N+1}, \quad t_1 = t_{\text{start}}, \quad t_{N+1} = t_{\text{final}}, \quad t_{i+1} - t_i = \Delta t.$$

Тогда, сузив класс допустимых управлений до кусочно-постоянных, получаем дискретный вариант рассматриваемой задачи Коши (3.1)-(3.2):

$$\begin{cases} x^{k+1} = f(x^k, u^k), & k = \overline{1, N}, \\ x^1 = x^{\text{start}}, \end{cases} \quad (3.4)$$

где

$$f(x^k, u^k) = \Delta t (A(x^k) + Bu^k) + x^k.$$

При этом функционал (3.3) для дискретной задачи приобретет вид

$$J = q^{N+1}(x^{N+1}) + \sum_{k=1}^N q^k(x^k) + \sum_{k=1}^N r^k(u^k), \quad (3.5)$$

где

$$q^{N+1} = q^{\text{final}}, \quad q^k = w_1 q \Delta t, \quad r^k = w_2 r \Delta t \quad k = \overline{1, N}.$$

## 4 Синтез оптимального управления

### 4.1 Описание метода

Существует два базовых метода для решения нелинейных задач оптимального управления, к классу которых относится задача (3.4)-(3.5):

1. Метод дифференциального динамического программирования (DDP) [9], [11];
2. Метод итеративного линейно-квадратичного регулятора (iLQR<sup>1</sup>) [5].

Методы идейно схожи:

1. Методы итеративны и используют полную информацию о состоянии системы<sup>2</sup>;
2. На каждой итерации методы используют некоторое *референсное* управление  $\bar{u}$  и соответствующую ему референсную траекторию  $\bar{x}$ ;
3. Вдоль референсной траектории задача полиномиально аппроксимируется;
4. На основании аппроксимированной системы строится некоторая поправка на исходное референсное управление.

**Определение 2.** Под *референсным управлением*  $\bar{u}$  мы будем понимать управление, которое подается на вход каждой итерации соответствующего алгоритма. Под *референсной траекторией* — соответствующую референсному управлению траекторию системы  $\bar{x}$ , либо иногда пару  $(\bar{u}, \bar{x})$ .

Отличие методов заключается в способе получения оптимальной поправки: метод DDP строит поправку как антиградиент гамильтониана аппроксимированной задачи

$$\delta u^k = -\alpha \nabla_u H(\bar{u}^k), \quad 0 < \alpha \leq 1,$$

метод iLQR — как её линейно-квадратичный регулятор.

---

<sup>1</sup>В некоторых источниках, например [12], используется аббревиатура SLQ.

<sup>2</sup>Наличие уравнения наблюдения предполагало бы, что каждая итерация алгоритма проводится не на компьютере, а на реальном объекте. Применительно к нашей модели это означало бы, что человек достигнет цели движения только с некоторой попытки.

Считается, что метод iLQR более надежный, так как в меньшей степени подвержен проблемам, присущим градиентным методам, таким как остановка в локальном минимуме, но сходится за большее число итераций, чем метод DDP. Однако при проведении численного эксперимента для сравнения скорости сходимости на конкретных задачах авторы приходят к противоположным результатам. Лучше всего резюмирует это положение вещей работа [7], в которой проведено сравнение двух методов для трёх классических задач механики, и в каждой задаче методы показывают разную асимптотику сходимости.

В данной работе для построения управления был выбран метод iLQR. Выпишем его основные шаги:

1. На каждой итерации имеем референсную траекторию  $(\bar{u}, \bar{x})$ ;
2. Вдоль референсной траектории линеаризуем задачу Коши и аппроксимируем функционал качества до второго порядка;
3. Строим поправку на управление  $\delta u$  как линейно-квадратичный регулятор аппроксимированной задачи;
4. Если не выполнено терминальное условие

$$|J(\bar{u}) - J(\bar{u} + \delta u)| < \varepsilon, \quad (4.1)$$

то используем поправленное управление  $\bar{u} + \delta u$  в качестве референсного на следующей итерации алгоритма.

Для возможности применения построенного метода важно понимать порядок числа итераций, требующихся для построения оптимального управления. Поэтому далее при построении численных решений конкретных постановок задач будет дополнительно указано, за какое число итераций *сошёлся* приведённый метод.

**Определение 3.** Будем говорить, что метод (алгоритм) *сошёлся на  $i$ -ой итерации*, если построенное на  $i$ -ой итерации управление  $\bar{u} + \delta u$  удовлетворяет терминальному условию (4.1).

## 4.2 Синтез оптимальной поправки

Допустим мы имеем некоторое референсное управление  $\bar{u} = \{\bar{u}^k\}_{k=1}^N$  и соответствующую ему референсную траекторию  $\bar{x} = \{\bar{x}^k\}_{k=1}^{N+1}$ . Введем обо-

значения:

$$\begin{aligned} f_x^k &= \frac{\partial f}{\partial x} \Big|_{(\bar{x}^k, \bar{u}^k)}, \quad f_u^k = \frac{\partial f}{\partial u} \Big|_{(\bar{x}^k, \bar{u}^k)}, \\ q^k &= q(\bar{x}^k), \quad q_x^k = \frac{\partial q}{\partial x} \Big|_{\bar{x}^k}, \quad q_{xx}^k = \frac{\partial^2 q}{\partial x^2} \Big|_{\bar{x}^k}, \\ r^k &= r(\bar{u}^k), \quad r_x^k = \frac{\partial r}{\partial u} \Big|_{\bar{u}^k}, \quad r_{xx}^k = \frac{\partial^2 r}{\partial u^2} \Big|_{\bar{u}^k}. \end{aligned}$$

Тогда, линейризуя вдоль референсной траектории задачу Коши (3.4) и строя квадратичную аппроксимацию вдоль той же траектории функционала качества (3.5), получаем следующую задачу:

$$\begin{cases} \delta x^{k+1} = f_x^k \delta x + f_u^k \delta u, \quad k = \overline{1, N}, \\ \delta x^1 = 0. \end{cases} \quad (4.2)$$

$$\begin{aligned} J_\delta &= q^{N+1} + q_x^{N+1} \tilde{x}^{N+1} + \frac{1}{2} \langle \tilde{x}^{N+1}, q_{xx}^{N+1} \tilde{x}^{N+1} \rangle + \\ &+ \sum_{k=1}^N \left[ q^k + q_x^k \tilde{x}^k + \frac{1}{2} \langle \tilde{x}^k, q_{xx}^k \tilde{x}^k \rangle \right] + \\ &+ \sum_{k=1}^N \left[ r^k + r_u^k \tilde{u}^k + \frac{1}{2} \langle \tilde{u}^k, r_{uu}^k \tilde{u}^k \rangle \right], \quad (4.3) \end{aligned}$$

где

$$\tilde{x}^k = \bar{x}^k + \delta x^k, \quad \tilde{u}^k = \bar{u}^k + \delta u^k.$$

Построим гамильтониан для задачи (4.2)-(4.3):

$$\begin{aligned} H_k &= q^k + q_x^k \tilde{x}^k + \frac{1}{2} \langle \tilde{x}^k, q_{xx}^k \tilde{x}^k \rangle + \\ &+ r^k + r_u^k \tilde{u}^k + \frac{1}{2} \langle \tilde{u}^k, r_{uu}^k \tilde{u}^k \rangle + \\ &+ (\lambda^{k+1})^T (f_x^k \delta x^k + f_u^k \delta u^k), \quad (4.4) \end{aligned}$$

где  $\lambda^{k+1}$  — мультипликаторы Лагранжа.

Согласно принципу максимума Понтрягина [20] оптимальное управление  $\delta u^*$  должно удовлетворять необходимому условию

$$\frac{\partial H_k}{\partial (\delta u^k)} \Big|_{\delta u^k = \delta u^{k*}} = r_u^k + r_{uu}^k (\bar{u}^k + \delta u^{k*}) + (f_u^k)^T \lambda^{k+1} = 0,$$

что дает следующее выражение для поправки:

$$\delta u^{k*} = -(r_{uu}^k)^{-1}[(f_u^k)^T \lambda^{k+1} + r_u^k] - \bar{u}^k. \quad (4.5)$$

И при этом имеет силу сопряженная задача:

$$\begin{cases} \lambda^k = (f_x^k)^T \lambda^{k+1} + q_x^k + q_{xx}^k(\bar{x}^k + \delta x^k) \\ \lambda^{N+1} = q_x^{N+1} + q_{xx}^{N+1}(\bar{x}^{N+1} + \delta x^{N+1}). \end{cases} \quad (4.6)$$

Из (4.5) и (4.6) вытекает

$$\begin{pmatrix} \delta x^{k+1} \\ \lambda^k \end{pmatrix} = \underbrace{\begin{pmatrix} f_x^k & -f_u^k(r_{uu}^k)^{-1}(f_u^k)^T \\ q_{xx}^k & (f_x^k)^T \end{pmatrix}}_{\Phi^k} \begin{pmatrix} \delta x^k \\ \lambda^{k+1} \end{pmatrix} + \underbrace{\begin{pmatrix} -f_u^k(r_{uu}^k)^{-1}r_u^k \\ q_x^k \end{pmatrix}}_{\Gamma^k}. \quad (4.7)$$

**Теорема 2.** *Оптимальная поправка  $\delta u$  для задачи (4.2)-(4.3) вычисляется как*

$$\delta u^k = L_k \delta x^k + d^k, \quad (4.8)$$

где

$$\begin{aligned} L_k &= -(r_{uu}^k + (f_u^k)^T S_{k+1} f_u^k)^{-1} (f_u^k)^T S_{k+1} f_u^k, \\ d^k &= -(r_{uu}^k + (f_u^k)^T S_{k+1} f_u^k)^{-1} (v^{k+1} + r_u^k). \end{aligned}$$

Причем переменные  $S_k$ ,  $v^k$  могут быть найдены в обратном времени из соотношений

$$\begin{aligned} S_k &= \Phi_{21}^k + \Phi_{22}^k S_{k+1} (I - \Phi_{12}^k S_{k+1})^{-1} \Phi_{11}^k, \\ v^k &= \Phi_{22}^k S_{k+1} (I - \Phi_{12}^k S_{k+1})^{-1} (\Phi_{12}^k v^{k+1} + \Gamma_1^k) + \Phi_{22}^k v^{k+1} + \Gamma_2^k \end{aligned} \quad (4.9)$$

с граничными условиями

$$\begin{aligned} S_{N+1} &= q_{xx}^{N+1}, \\ v^{N+1} &= q_x^{N+1} + q_{xx}^{N+1} \bar{x}^{N+1}. \end{aligned} \quad (4.10)$$

**Д о к а з а т е л ь с т в о.** Предположим, что мультипликаторы  $\lambda$  имеют следующую аффинную форму относительно фазовой переменной  $\delta x$

$$\lambda^k = S_k \delta x^k + v^k \quad (4.11)$$

Тогда из граничного условия (4.6) вытекает граничное условие на  $S_k$ ,  $v^k$  (4.10):

$$\lambda^{N+1} = q_x^{N+1} + q_{xx}^{N+1} (\bar{x}^{N+1} + \delta x^{N+1})$$

$\Downarrow$

$$S_{N+1} = q_{xx}^{N+1}, v^{N+1} = q_x^{N+1} + q_{xx}^{N+1} \bar{x}^{N+1}.$$

Теперь подставим (4.11) в выражение (4.7) для  $\delta x^{k+1}$ :

$$\delta x^{k+1} = \Phi_{11}^k \delta x^k + \Phi_{12}^k (S_{k+1} \delta x^{k+1} + v^{k+1}) + \Gamma_1^k.$$

Получаем

$$\delta x^{k+1} = \left( \underbrace{I - \Phi_{12}^k S_{k+1}}_{K_k} \right)^{-1} (\Phi_{11}^k \delta x^k + \Phi_{12}^k v^{k+1} + \Gamma_1^k).$$

Подставим получившееся выражение в (4.7) для  $\lambda^k$ :

$$\begin{aligned} \lambda^k &= S_k \delta x^k + v^k = \Phi_{21}^k \delta x^k + \Phi_{22}^k (S_{k+1} \delta x^{k+1} + v^{k+1}) + \Gamma_2^k = \\ &= \Phi_{21}^k \delta x^k + \Phi_{22}^k (S_{k+1} K_k^{-1} (\Phi_{11}^k \delta x^k + \Phi_{12}^k v^{k+1} + \Gamma_1^k) + v^{k+1}) + \Gamma_2^k. \end{aligned}$$

Таким образом получаем искомые соотношения (4.9):

$$\begin{aligned} S_k &= \Phi_{21}^k + \Phi_{22}^k S_{k+1} K_k^{-1} \Phi_{11}^k, \\ v^k &= \Phi_{22}^k (S_{k+1} K_k^{-1} (\Phi_{12}^k v^{k+1} + \Gamma_1^k) + v^{k+1}) + \Gamma_2^k. \end{aligned}$$

Итоговая формула для оптимальной поправки (4.8) получается прямой подстановкой получившихся соотношений в выражение (4.5). ■

*Замечание 3.* Полученная теорема требует существования обратных матриц для  $K_k, k = \overline{1, N}$ . При этом для нелинейных систем данное условие может не выполняться. Чтобы метод продолжал работать, предлагается в случае нулевого определителя  $\det K_k = 0$ , заменять в формулах (4.8), (4.9) матрицу  $K_k$  на регуляризованную

$$\mathcal{K}_k = K_k + \mu I. \quad (4.12)$$

### 4.3 Регуляризация оптимальной поправки

Согласно Теореме 1 оптимальная поправка имеет следующую аффинную форму

$$\delta u^{k*} = L_k \delta x^k + d^k,$$

где  $L_k$  — коэффициент управления с обратной связью,  $d^k$  — коэффициент управления без обратной связи, возникающий по причине того, что мы имеем дело с отклонениями от заданного состояния.



Данная форма не налагает никаких ограничений на поправку  $\delta u$ . На практике это означает, что на начальных итерациях алгоритма, когда референсная траектория далека от оптимальной, поправка зачастую выводит систему за область действия аппроксимации. Визуально это выражается в том, что на каждой итерации алгоритм выдаёт некоторую случайную траекторию и в конечном итоге не сходится к оптимальной траектории. Чтобы избежать такого эффекта, необходимо регуляризовать коэффициент управления без обратной связи  $d^k$ :

$$\delta u^{k*}(\eta) = L_k \delta x^k + \eta d^k.$$

Теперь остается ответить на вопрос, как выбрать подходящий коэффициент регуляризации  $\eta$ . Это можно сделать двумя способами:

1. Дополнительно поточечно ограничить область допустимых управлений

$$u^k \in \mathcal{U}^k.$$

В таком случае можно выбирать коэффициент из соотношения

$$\eta = \begin{cases} 1, & \text{при } \bar{u}^k + \delta u^{k*} \in \mathcal{U}^k, k = \overline{1, N}, \\ \min \{ \eta \mid \bar{u}^k + \delta u^{k*}(\eta) \in \partial \mathcal{U}^k \}, & \text{иначе.} \end{cases}$$

Такой способ предполагает всего один дополнительный проход алгоритма для поиска минимума, однако накладывает ограничения, не предусмотренные исходной задачей, и не даёт гарантии на то, что в среднем отклонение в конечном счёте не накопится.

2. Использовать ожидаемое отклонение от функции цены

$$\xi_1 \leq \frac{J(\bar{u}) - J(\bar{u} + \delta u^*(\eta))}{J_\delta(0) - J_\delta(\delta u^*(\eta))} \leq \xi_2. \quad (4.13)$$

Данный способ не накладывает дополнительных ограничений. При этом для нахождения правильного коэффициента может потребоваться несколько итераций. Изначально коэффициент  $\eta$  инициализируется единицей, затем считается оптимальная поправка, и если условие (4.13) не выполнено, то коэффициент уменьшается  $\eta \leftarrow \gamma \eta$ , где  $\gamma < 1$ .

*Замечание 4.* В случае задачи с заданными ограничениями на управление необходимо комбинировать оба рассмотренных условия. Для численного решения нашей задачи было использовано второе.

## 4.4 Алгоритм

С учетом вышесказанного можно построить следующий алгоритм. Алгоритмы 1, 2 демонстрируют обратный и прямой проход для получения оптимальной поправки. Алгоритм 3 демонстрирует общий алгоритм построения управления.

Продемонстрируем работу алгоритма для классической задачи перехода в целевое состояние  $x^{\text{final}}$  без фазовых ограничений. Для этого будем использовать следующие компоненты функции цены:

$$q^{\text{final}}(x) = \|x - x^{\text{final}}\|^2, \quad q(x) \equiv 0. \quad (4.14)$$

Начальным референсным управлением выберем

$$\bar{u}^k = 0, \quad k = \overline{1, N}. \quad (4.15)$$

---

**Алгоритм 1: Обратный проход**

---

```
function BackwardPass( $\bar{u}$ ,  $\bar{x}$ )  
begin  
     $S_{N+1}, v^{N+1} \leftarrow (4.10)$   
    for  $k \leftarrow N$  to 1 do  
         $S_k, v^k \leftarrow (4.9)$   
    end  
    return  $S, v$   
end
```

---

---

**Алгоритм 2: Прямой проход**

---

```
function ForwardPass( $\bar{u}$ ,  $\bar{x}$ ,  $J_{\text{prev}}$ )  
begin  
     $\eta \leftarrow 1$   
    do  
         $\delta x^0 \leftarrow 0$   
        for  $k \leftarrow 1$  to  $N$  do  
             $\delta u^k, \delta x^{k+1} \leftarrow (4.8), (4.2)$   
             $u^k \leftarrow \bar{u}^k + \delta u^k$   
        end  
         $J \leftarrow (3.5)$   
         $\eta \leftarrow \gamma \eta$   
    while  $\frac{J_{\text{prev}} - J}{J_\delta(0) - J_\delta(\delta u)} \notin [\xi_1, \xi_2]$   
    return  $u, J$   
end
```

---

---

**Алгоритм 3: Синтез управления**

---

```
function Synthesis( $\bar{u}$ )  
begin  
     $J \leftarrow (3.5)$   
    do  
         $J_{\text{prev}} \leftarrow J$   
         $\bar{x} \leftarrow (3.4)$   
         $S, v \leftarrow \text{BackwardPass}(\bar{u}, \bar{x})$   
         $u, J \leftarrow \text{ForwardPass}(S, v, J_{\text{prev}})$   
    while  $|J - J_{\text{prev}}| \geq \varepsilon$   
    return  $u$   
end
```

---

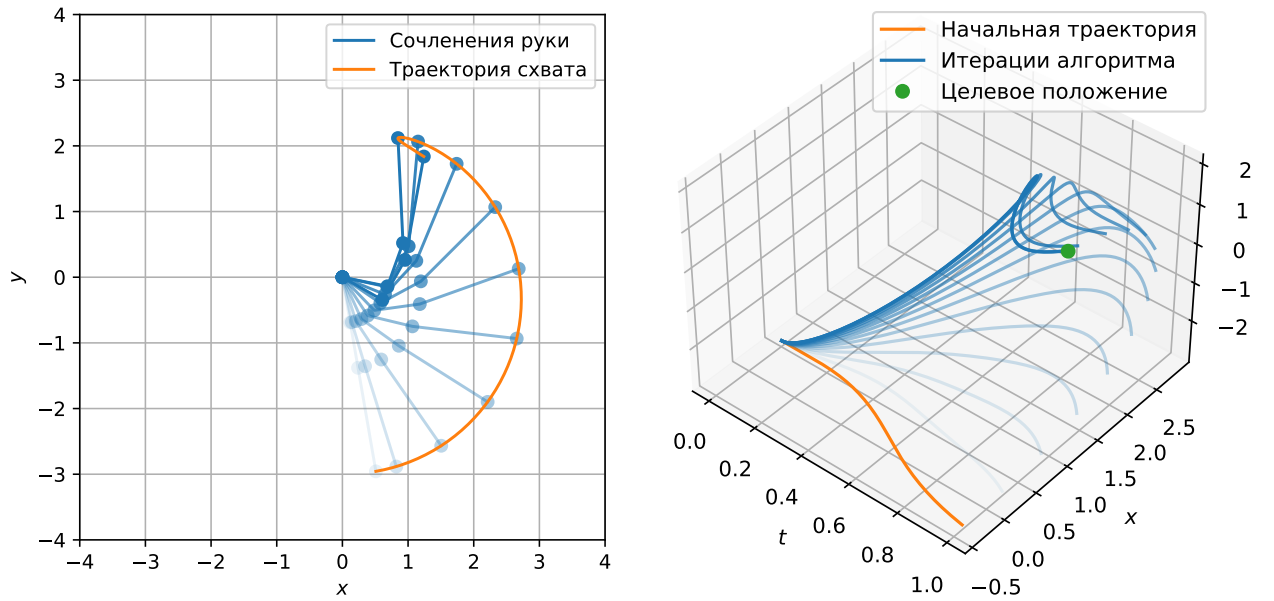


Рис. 3: Решение задачи перехода в целевое состояние (4.14) с начальным референсным управлением (4.15). Слева: поведение системы при полученном управлении. Справа: траектории схвата на каждой итерации алгоритма, более ранние итерации показаны бледнее. Алгоритм сошелся на 14 итерации.

## 5 Синтез начального референсного управления

### 5.1 Описание метода

Скорость сходимости метода, приведенного в Разделе 4, зависит от выбора начального референсного управления  $\bar{u}$ . Считается, что при выборе начального управления можно положиться на мнение эксперта в предметной области, который предложит траекторию системы  $\bar{x}$ , близкую к оптимальной. В этом случае перед применением метода необходимо решить задачу оптимального преследования, то есть найти управление  $\bar{u}$ , приводящее систему к заданной траектории  $\bar{x}$ .

Однако в отсутствии эксперта необходимо предложить метод для построения начальной референсной траектории. Выпишем основные критерии, которым должен удовлетворять метод:

1. Метод должен строить управление *быстро* — желательно, чтобы алгоритм имел линейную асимптотику;
2. Получившаяся референсная траектория должна быть близкой к оптимальной;
3. Получившаяся референсная траектория должна быть возможной для рассматриваемой задачи. Данное условие важно, если в задаче присутствуют ограничения на управление.

В данном разделе предложен метод удовлетворяющий критериям выше. Выпишем его основные шаги:

1. Необходимо аналитически найти состояние системы, которое минимизирует терминальное условие

$$x^{\text{final}} \in \text{Argmin } q^{\text{final}}(x);$$

2. Привести систему к линейной и поставить для нее задачу минимизации интегрально-квадратичного функционала для перехода в состояние  $x^{\text{final}}$ ;
3. Построить линейно-квадратичный регулятор для полученной задачи.

В случае, если мы можем аналитически найти несколько минимизаторов терминального условия, можно провести перебор: построить начальную траекторию для каждого минимизатора, а затем выбрать управление с минимальной величиной функционала качества  $J$ .

Тем самым мы получим управление, которое минимизирует терминальное условие, но ничего не говорит об энергетическом и фазовом условиях. Тем не менее такой подход будет работать лучше, чем выбор случайного управления. Сравнение скорости сходимости для задачи (4.14) для нулевого начального управления и начального управления, предложенного в данном разделе, приведено на Рис. 5.

*Замечание 5.* Как мы увидим далее, полученная референсная траектория минимизирует угловые ускорения сочленений руки. Таким образом результирующая траектория будет самой плавной из возможных. С физической точки зрения такая траектория кажется допустимой для руки, что позволяет использовать её и в задачах с наличием ограничений на управление.

## 5.2 Синтез управления

Приведём систему (3.4) к линейному виду заменой управления на

$$\hat{u} = M^{-1}(x_1)[\tau - L(x_1, x_2)]. \quad (5.1)$$

Тогда в фазовом пространстве  $\hat{x} = [x_1, x_2]^T \in \mathbb{R}^6$  задача Коши примет вид

$$\begin{cases} \hat{x}^{k+1} = \hat{A}\hat{x}^k + \hat{B}\hat{u}^k, & k = \overline{1, N}, \\ \hat{x}^1 = x^{\text{start}}, \end{cases} \quad (5.2)$$

где

$$\hat{A} = \left( \begin{array}{c|c} I & \Delta t I \\ \hline O & I \end{array} \right), \quad \hat{B} = \left( \begin{array}{c} O \\ \Delta t I \end{array} \right).$$

Будем считать, что для исходной задачи задачи (3.4)-(3.5) известно некоторое состояние  $x^{\text{final}}$ , минимизирующее терминальное условие  $q^{\text{final}}$ , то есть

$$x^{\text{final}} \in \text{Argmin } q^{\text{final}}(x).$$

В таком случае для задачи Коши (5.2) поставим задачу минимизации следу-

ющего функционала

$$J = \|\hat{x}^{N+1} - x^{\text{final}}\|^2 + \hat{w}_1 \sum_{k=1}^N \langle \hat{x}^k, \hat{Q}\hat{x}^k \rangle + \hat{w}_2 \sum_{k=1}^N \|\hat{u}^k\|^2 \longrightarrow \min. \quad (5.3)$$

Здесь матрица  $\hat{Q} = \hat{Q}^T$  выбирается для исключения возможного *проворачивания* сочленений относительно друг друга следующим образом

$$\hat{Q} = \left( \begin{array}{ccc|c} 2 & -1 & 0 & O \\ -1 & 2 & -1 & O \\ 0 & -1 & 1 & O \\ \hline & & & O \end{array} \right). \quad (5.4)$$

Иными словами матрица  $\hat{Q}$  является матрицей следующей квадратичной формы

$$\langle \hat{x}, \hat{Q}\hat{x} \rangle = \theta_1^2 + (\theta_2 - \theta_1)^2 + (\theta_3 - \theta_2)^2.$$

Данное фазовое условие штрафует траекторию в случае большого относительного отклонения между углами сочленений.

После решения задачи (5.2)-(5.3) мы можем восстановить соответствующее управление исходной задачи. Пусть  $\hat{u}^*$ ,  $\hat{x}^*$  — оптимальное управление и соответствующая ему оптимальная траектория задачи (5.2)-(5.3). Тогда соответствующее управление для исходной задачи  $u$ , можно получить по формуле

$$u^k = \frac{\hat{\tau}^k - \hat{\tau}^{k-1}}{\Delta t}, \quad (5.5)$$

где

$$\begin{cases} \hat{\tau}^k = M(\hat{x}_1^{k*})\hat{u}^{k*} + L(\hat{x}_1^{k*}, \hat{x}_2^{k*}), \text{ при } k = \overline{1, N}, \\ \hat{\tau}^0 = \tau^{\text{start}}. \end{cases}$$

Построим гамильтониан задачи (5.2)-(5.3)

$$\hat{H}_k = \langle \hat{x}^k, \hat{w}_1 \hat{Q} \hat{x}^k \rangle + \langle \hat{u}^k, \hat{w}_2 \hat{u}^k \rangle + (\hat{\lambda}^{k+1})^T [\hat{A} \hat{x}^k + \hat{B} \hat{u}^k]. \quad (5.6)$$

Оптимальное управление  $\hat{u}^*$  должно удовлетворять необходимому условию оптимальности:

$$\left. \frac{\partial \hat{H}_k}{\partial \hat{u}^k} \right|_{\hat{u}^k = \hat{u}^{k*}} = \hat{w}_2 \hat{u}^{k*} + \hat{B}^T \hat{\lambda}^{k+1} = 0,$$

что дает следующее выражение для управления

$$\hat{u}^{k*} = -\frac{1}{\hat{w}_2} \hat{B}^T \hat{\lambda}^{k+1}. \quad (5.7)$$

И уравнение (5.2) можно переписать в следующем виде:

$$\hat{x}^{k+1} = \hat{A} \hat{x}^k - \frac{1}{\hat{w}_2} \hat{B} \hat{B}^T \hat{\lambda}^{k+1}. \quad (5.8)$$

При этом имеет силу следующая сопряженная система:

$$\begin{cases} \hat{\lambda}^k = \hat{w}_1 \hat{Q} \hat{x}^k + \hat{A}^T \hat{\lambda}^{k+1}, \text{ при } k = \overline{1, N} \\ \hat{\lambda}^{N+1} = \hat{x}^{N+1} - x^{\text{final}} \end{cases} \quad (5.9)$$

**Теорема 3.** *Оптимальное управление  $\hat{u}^*$  задачи (5.2)-(5.3) задается формулой*

$$\hat{u}^{k*} = \hat{L}_k \hat{x}^k + \hat{d}^k, \quad (5.10)$$

где

$$\begin{aligned} \hat{L}_k &= -\frac{1}{\hat{w}_2} \hat{B}^T S_{k+1} \left( I + \frac{1}{\hat{w}_2} \hat{B} \hat{B}^T S_{k+1} \right)^{-1} \hat{A}, \\ \hat{d}^k &= -\frac{1}{\hat{w}_2} \hat{B}^T \left( I - S_{k+1} \left( I + \frac{1}{\hat{w}_2} \hat{B} \hat{B}^T S_{k+1} \right)^{-1} \hat{B} \hat{B}^T \right) v^{k+1}. \end{aligned}$$

Причем переменные  $S_k$ ,  $v^k$  могут быть найдены в обратном времени из соотношений

$$\begin{aligned} S_k &= \hat{w}_1 \hat{Q} + \hat{A}^T S_{k+1} \left( I + \frac{1}{\hat{w}_2} \hat{B} \hat{B}^T S_{k+1} \right)^{-1} \hat{A}, \\ v_k &= \hat{A}^T \left( I - \frac{1}{\hat{w}_2} S_{k+1} \left( I + \frac{1}{\hat{w}_2} \hat{B} \hat{B}^T S_{k+1} \right)^{-1} \hat{B} \hat{B}^T \right) v^{k+1} \end{aligned} \quad (5.11)$$

с граничными условиями

$$\begin{aligned} S_{N+1} &= I, \\ v^{N+1} &= -x^{\text{final}}. \end{aligned} \quad (5.12)$$

**Д о к а з а т е л ь с т в о.** Аналогично доказательству теоремы (1) будем искать решение сопряженной системы в аффинном виде

$$\hat{\lambda}^k = S_k \hat{x}^k + v^k. \quad (5.13)$$



Из сопряженной системы (5.9) получаем граничные условия:

$$\hat{\lambda}^{N+1} = \hat{x}^{N+1} - x^{\text{final}} \implies S_{N+1} = I, v^{N+1} = -x^{\text{final}}.$$

Подставив выражение (5.13) в уравнение (5.8), получим

$$\hat{x}^{k+1} = \hat{A}\hat{x}^k - \frac{1}{\hat{w}_2}\hat{B}\hat{B}^T(S_{k+1}\hat{x}^{k+1} + v^{k+1}),$$

откуда выражаем

$$\hat{x}^{k+1} = \left( \underbrace{I + \frac{1}{\hat{w}_2}\hat{B}\hat{B}^T S_{k+1}}_{K_k} \right)^{-1} \left( \hat{A}\hat{x}^k - \frac{1}{\hat{w}_2}\hat{B}\hat{B}^T v^{k+1} \right).$$

Теперь подставим получившееся выражение в (5.9):

$$\begin{aligned} \hat{\lambda}^k &= S_k \hat{x}^k + v^k = \hat{w}_1 \hat{Q} \hat{x}^k + \hat{A}^T (S_{k+1} \hat{x}^{k+1} + v^{k+1}) = \\ &= \hat{w}_1 \hat{Q} \hat{x}^k + \hat{A}^T S_{k+1} K_k^{-1} \left( \hat{A} \hat{x}^k - \frac{1}{\hat{w}_2} \hat{B} \hat{B}^T v^{k+1} \right) + \hat{A}^T v^{k+1} = \\ &= \left( \hat{w}_1 \hat{Q} + \hat{A}^T S_{k+1} K_k^{-1} \hat{A} \right) \hat{x}^k + \hat{A}^T \left( I - \frac{1}{\hat{w}_2} S_{k+1} K_k^{-1} \hat{B} \hat{B}^T \right) v^{k+1}. \end{aligned}$$

Откуда получаем искомые соотношения:

$$\begin{aligned} S_k &= \hat{w}_1 \hat{Q} + \hat{A}^T S_{k+1} \left( I + \frac{1}{\hat{w}_2} \hat{B} \hat{B}^T S_{k+1} \right)^{-1} \hat{A}, \\ v_k &= \hat{A}^T \left( I - \frac{1}{\hat{w}_2} S_{k+1} \left( I + \frac{1}{\hat{w}_2} \hat{B} \hat{B}^T S_{k+1} \right)^{-1} \hat{B} \hat{B}^T \right) v^{k+1}. \end{aligned}$$

Теперь выражение для оптимального управления (5.10) получается прямой подстановкой получившихся соотношений в выражение (5.7). ■

### 5.3 Алгоритм

Алгоритм 4 резюмирует метод, предложенный в данном разделе. На Рис. 4 представлен результат работы основного алгоритма из Раздела 4 для задачи (4.14) с построенным начальным управлением по предложенному алгоритму.

---

**Алгоритм 4:** Поиск начальной траектории

---

**function** *InitialControl*( $x^{\text{final}}$ )**begin**

/\* Обратный проход

\*/

 $S_{N+1}, v^{N+1} \leftarrow (5.12)$     **for**  $k \leftarrow N$  **to** 1 **do**         $S_k, v^k \leftarrow (5.11)$     **end**

/\* Прямой проход

\*/

 $\hat{x}^0 \leftarrow \theta^{\text{start}}, \dot{\theta}^{\text{start}}$     **for**  $k \leftarrow 1$  **to**  $N$  **do**         $\hat{u}^k, \hat{x}^{k+1} \leftarrow (5.10), (5.2)$     **end**

/\* Конвертация управления

\*/

**for**  $k \leftarrow 1$  **to**  $N$  **do**         $u^k \leftarrow (5.5)$     **end**    **return**  $u$ **end**

---

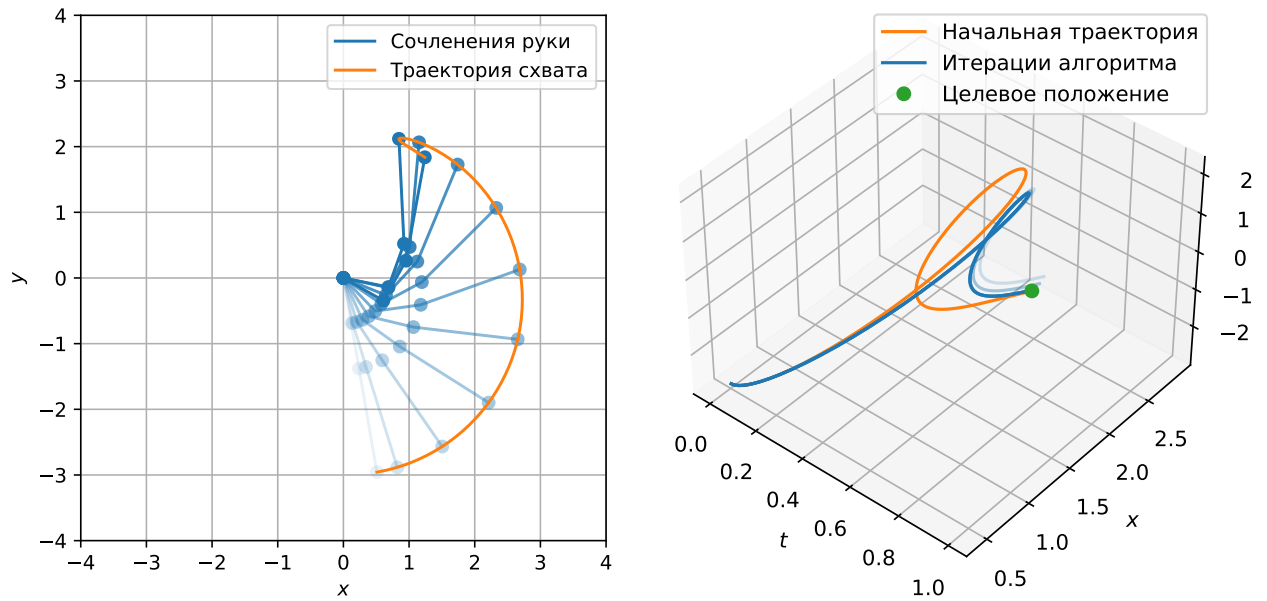


Рис. 4: Оптимальная траектория и траектории схвата на различных итерациях алгоритма при решении задачи (4.14) с начальным управлением, построенным методом из данного раздела. Алгоритм сошелся на 4 итерации.

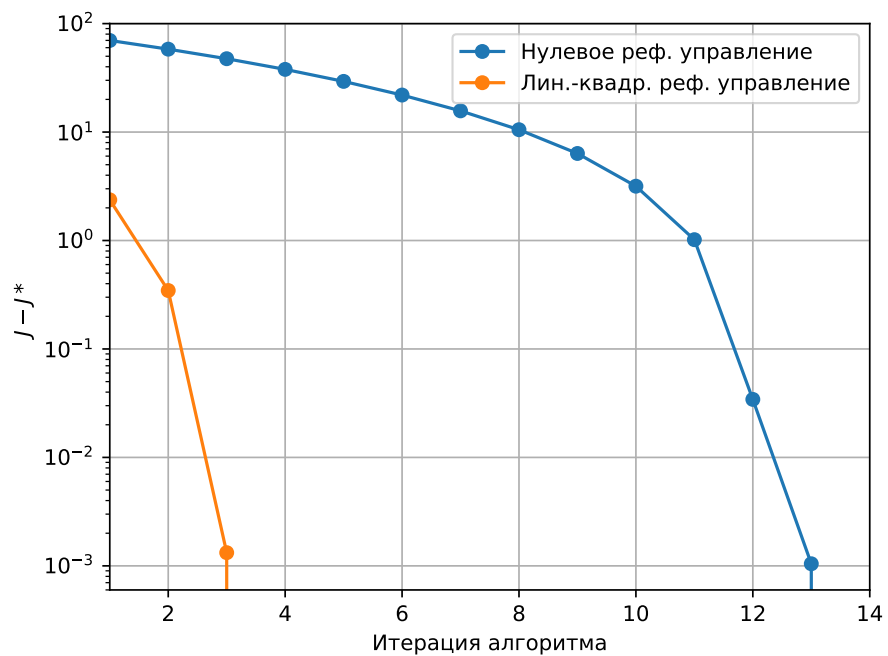


Рис. 5: Сравнение скорости сходимости для задачи (4.14) в зависимости от выбора начального референсного управления.

## 6 Примеры

### 6.1 Целевое положение схвата

В данном разделе мы приведем примеры работы алгоритма для классических задач биологического движения.

*Замечание 6.* В каждом из примеров, для избежания проворачивания сочленений, в интегральную часть функционала качества дополнительно добавлена поправка

$$10^{-5} \cdot \langle x^k, \hat{Q}x^k \rangle.$$

Пусть целью нашего движения является достижение схватом заранее определенного положения

$$e^{\text{final}} \in \mathcal{B}_0 \left( \sum_{i=1}^3 l_i \right).$$

Таким образом получаем следующее терминальное условие функционала качества:

$$q^{\text{final}}(x) = \|e^3(x) - e^{\text{final}}\|^2 \quad (6.1)$$

Теперь аналитически найдем минимизатор терминального условия, требующийся для построения начального референсного управления. Угол третьего сочленения возьмем из следующего соотношения:

$$\text{tg } \theta_3 = \begin{cases} \frac{e_2^{\text{final}}}{e_1^{\text{final}}}, & \text{при } e_1^{\text{final}} \neq 0, \\ 0, & \text{иначе.} \end{cases}$$

Данная форма  $\theta_3$  гарантирует, что  $e^2(x) \in \mathcal{B}_0(l_1 + l_2)$ .

Оставшиеся углы найдем, как решение следующей системы:

$$\begin{cases} l_1 \cos \theta_1 + l_2 \cos \theta_2 = e_1^{\text{final}} - l_3 \cos \theta_3, \\ l_1 \sin \theta_1 + l_2 \sin \theta_2 = e_2^{\text{final}} - l_3 \sin \theta_3. \end{cases}$$

Данная система имеет два решения на рассматриваемом интервале  $\theta_1, \theta_2 \in [-\pi, \pi]$ , которые соответствуют возможному положению руки. В работе строились начальные траектории для обоих вариантов, затем для итеративного алгоритма выбиралась траектория с наименьшим значением функционала качества.

Остальные граничные параметры, не участвующие в терминальном условии фиксируем в нуле:

$$\dot{\theta}^{\text{final}} = 0.$$

Рис. 6 демонстрирует траекторию руки при построенном управлении, а также траектории схвата при управлениях, полученных на различных итерациях алгоритма.

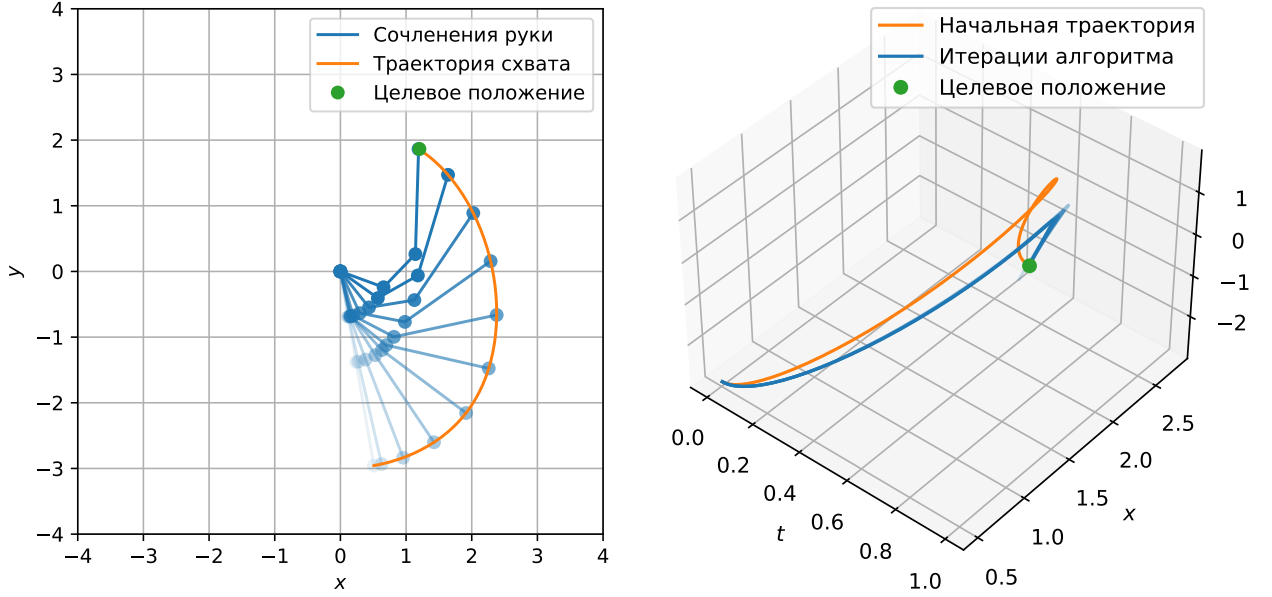


Рис. 6: Траектория системы при оптимальном управлении и итеративные траектории схвата для задачи целевого положения схвата. Алгоритм сошёлся на 5 итерации.

## 6.2 Целевое положение и скорость схвата

Пусть целью нашего движения является достижение схватом заранее определённого положения и заранее определённой скорости

$$e^{\text{final}} \in \mathcal{B}_0 \left( \sum_{i=1}^3 l_i \right), \quad \dot{e}^{\text{final}} \in \mathbb{R}^2.$$

Таким образом получаем следующие компоненты функционала качества:

$$q^{\text{final}}(x) = \|e^3(x) - e^{\text{final}}\|^2 + w_3 \|\dot{e}^3(x) - \dot{e}^{\text{final}}\|^2, \quad (6.2)$$

где  $w_3 < 1$  — вес терминального критерия скорости.

Минимизатор терминального условия по  $\theta$  находим из соотношений, представленных в предыдущем подразделе. Для поиска минимизатора по  $\dot{\theta}$  фиксируем

$$\dot{\theta}^{\text{final}} = 0,$$

и находим оставшиеся угловые скорости находим как решение следующей системы:

$$\begin{cases} -l_2 \sin \theta_2 \dot{\theta}_2 - l_3 \sin \theta_3 \dot{\theta}_3 = \dot{e}_1^{\text{final}}, \\ l_2 \cos \theta_2 \dot{\theta}_2 + l_3 \cos \theta_3 \dot{\theta}_3 = \dot{e}_2^{\text{final}}, \end{cases}$$

Рис. 7 демонстрирует траекторию руки при построенном управлении, а также траектории схвата при управлениях, полученных на различных итерациях алгоритма.

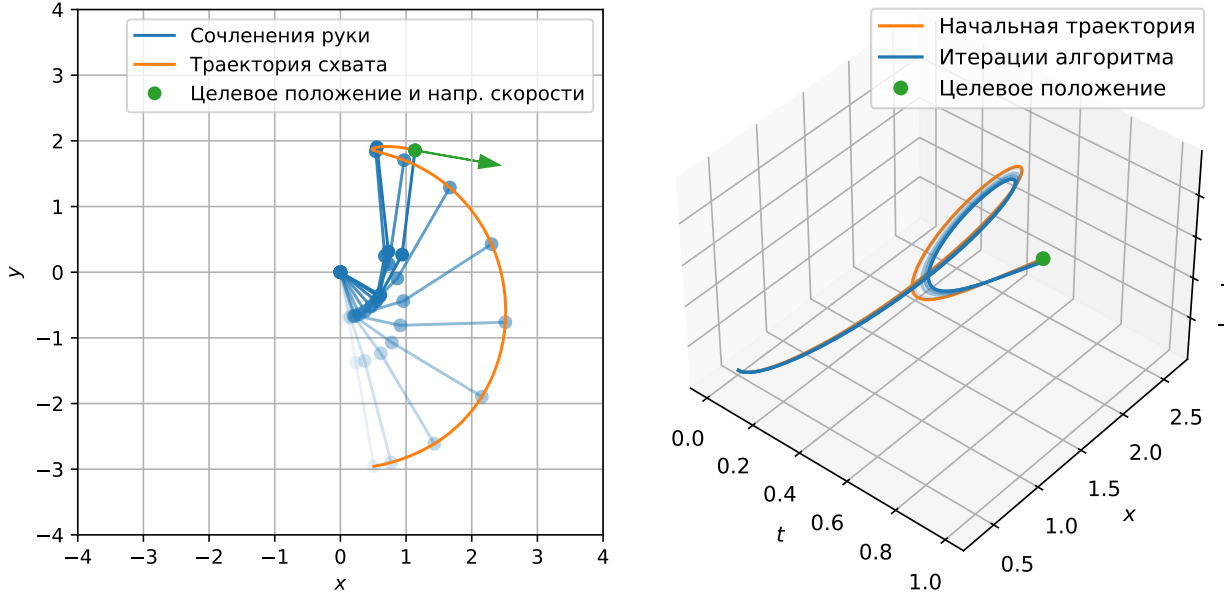


Рис. 7: Траектория системы при оптимальном управлении и итеративные траектории схвата для задачи целевого положения и скорости схвата. Алгоритм сошёлся на 7 итерации.

### 6.3 Обход препятствия

Учёт фазовых ограничений в интегральной части функционала качества  $J$ , представленный в работе, позволяет лишь приближенно описать условия вида

$$g_i(x) \leq 0,$$

которые возникают естественным образом в задаче обхода препятствия. Для этого фазовое условие  $q$  выбирается таким образом, чтобы штрафовать за приближение траектории к препятствию.

*Замечание 7.* Для формального решения задачи с подобными условиями, необходимо пользоваться методами расширенного лагранжиана [1], которые предполагают решение серии задач типа (3.4)-(3.5). Это приводит к ухудшению асимптотики алгоритмов и тем самым существенному увеличению времени работы программного решения.

Пусть задано некоторое точечное препятствие с центром  $e^{\text{obstacle}}$  и радиусом  $r_{\text{obstacle}}$ . Тогда зададим интегральное условие в виде:

$$q(x) = \left( \|e^3(x) - e^{\text{obstacle}}\|^2 - r_{\text{obstacle}} \right)^{-2}. \quad (6.3)$$

Рис. 8 и Рис. 9 демонстрируют траекторию руки при построенном управлении, а также траектории схвата при управлениях, полученных на различных итерациях алгоритма для решения задачи целевого состояния (4.14) с фазовым ограничением (6.3) при различных весах на фазовое условие.

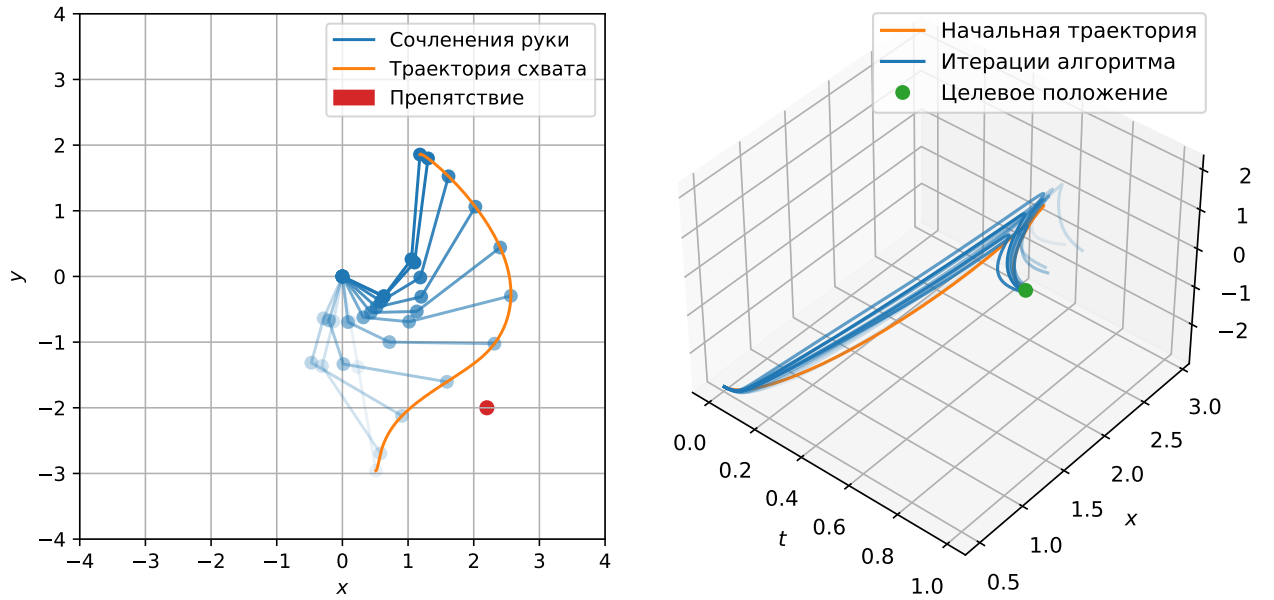


Рис. 8: Траектория системы при оптимальном управлении и итеративные траектории схвата для задачи обхода препятствия. Вес фазового условия  $w_1 = 10^{-1}$ . Алгоритм сошёлся на 10 итерации.

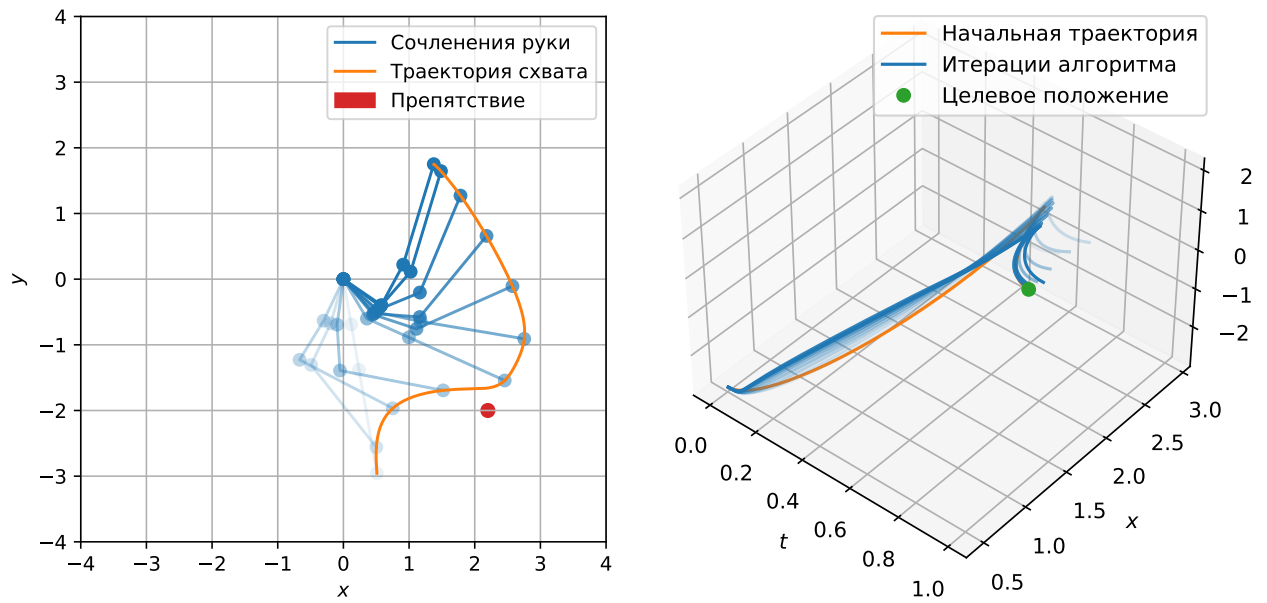


Рис. 9: Траектория системы при оптимальном управлении и итеративные траектории схвата для задачи обхода препятствия. Вес фазового условия  $w_1 = 2 \cdot 10^{-2}$ . Алгоритм сошёлся на 8 итерации.



## 7 Заключение

В рамках данной работы был предложен метод управления биологической системой, соответствующий реальному поведению человека при осуществлении целевого движения.

С целью возможности применения метода было проведено следующее:

1. Предложена математическая модель планарной руки человека, держащей предмет, и выведено уравнение динамики для данной системы;
2. Постулирован принцип оптимальности, дающий возможность применения методов оптимального управления для построения траекторий движения, соответствующих траекториям человека;
3. Рассмотрены предложенные в литературе формализации энергетического критерия оптимальности, минимизация которого присуща биологическому движению;
4. В соответствии с полученным уравнением динамики и выбранным критерием оптимальности поставлена задача оптимального целевого управления системой в непрерывной и дискретной формах.

Для решения поставленной задачи были рассмотрены известные базовые методы решения нелинейных задач оптимального управления, и на основе метода итеративного линейно-квадратичного регулятора был разработан метод построения оптимального управления для дискретной постановки задачи. Особенностью метода, предложенного в данной работе, является способ регуляризации итеративной поправки на управление, позволяющий улучшить сходимость базового метода для рассматриваемой задачи. Рассмотрены альтернативные варианты регуляризации поправки и возможность их применения для задач с ограничениями на управление. Для демонстрации работы метода было реализовано программное решение.

С целью уменьшения числа итераций предложенного алгоритма и, соответственно, времени работы программного решения был предложен способ построения начального референсного управления как линейно-квадратичного регулятора приведенной к линейному виду задачи для минимизации терминального критерия исходной задачи. Данный способ быстро работает и не

опирается на мнение эксперта в области биологической моторики при построении начальной референсной траектории. На примере было показано, что он действительно снижает необходимое число итераций основного алгоритма.

Наконец для возможности сравнения предложенного метода с имеющимися в литературе в качестве примера были рассмотрены конкретные постановки задачи, являющиеся классическими задачами биомеханического движения:

1. Переход в целевое состояние;
2. Переход в целевое положение схвата;
3. Переход в целевое положение схвата с заданной скоростью;
4. Задача обхода препятствия.

Для рассмотренных классических задач:

1. Представлена формализация в подходящем для предложенного метода виде, а также способ нахождения минимизатора терминального критерия для построения начальной референсной траектории;
2. При помощи программного решения построено оптимальное управление, и приведена соответствующая ему траектория системы;
3. Приведено число итераций, потребовавшихся программному решению для построения оптимального управления.

Таким образом, работа содержит полный анализ движения руки человека, держащей предмет: от построения математической модели и постановки задачи до разработки метода её решения и его имплементации в виде программного решения. Основными научными результатами работы являются метод построения оптимального управления и метод построения начальной референсной траектории, выраженными в Теореме 1, Теореме 2 и предложенном способе регуляризации оптимальной поправки (Подраздел 4.3).

## Список литературы

- [1] Birgin, E. G. and J. M. Martínez: *Practical augmented Lagrangian methods*. Springer US, Boston, MA, 2009.
- [2] Breteler, M., R. Meulenbroek, and S. Gielen: *An evaluation of the minimum-jerk and minimum torque-change principles at the path, trajectory, and movement-cost levels*. Motor control, 6:69–83, February 2002.
- [3] Harris, C. M. and D. M. Wolpert: *Signal-dependent noise determines motor planning*. Nature, 394, August 1998.
- [4] Hogan, N.: *An organizing principle for a class of voluntary movement*. Journal of Neuroscience, 4(11):2745–2754, 1984.
- [5] Li, W. and E. Todorov: *Iterative linear quadratic regulator design for non-linear biological movement systems*. In *ICINCO (1)*, pages 222–229. Cite-seer, 2004.
- [6] Madduri, M. M., S. A. Burden, and A. L. Orsborn: *Biosignal-based co-adaptive user-machine interfaces for motor control*. Current Opinion in Biomedical Engineering, 27, 2023.
- [7] Manchester, Z. and S. Kuindersma: *Derivative-free trajectory optimization with unscented dynamic programming*. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 3642–3647, 2016.
- [8] Maroger, I., O. Stasse, and B. Watier: *Inverse optimal control to model human trajectories during locomotion*. Computer Methods in Biomechanics and Biomedical Engineering, 25(5):499–511, 2022.
- [9] Mayne, D.: *A second-order gradient method for determining optimal trajectories of non-linear discrete-time systems*. International Journal of Control, 3(1):85–95, 1966.
- [10] Mitrovic, D., S. Klanke, and S. Vijayakumar: *Adaptive Optimal Feedback Control with Learned Internal Dynamics Models*, pages 65–84. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

- [11] Murray, D. and S. Yakowitz: *Differential dynamic programming and newton's method for discrete optimal control problems*. Journal of Optimization Theory and Applications, 43(3):395–414, 1984.
- [12] Sideris, A. and J. E. Bobrow: *An efficient sequential linear quadratic algorithm for solving nonlinear optimal control problems*. Transactions on Automatic Control, 50(12):2043—2047, 2005.
- [13] Todorov, E. and M. Jordan: *Optimal feedback control as a theory of motor coordination*. Nature Neuroscience, 5(11):1226–1235, 2002.
- [14] Todorov, E. and W. Li: *A generalized iterative lqg method for locally-optimal feedback control of constrained nonlinear stochastic systems*. In *American Control Conference, 2005.*, pages 300–306, 2005.
- [15] Uno, Y., M. Kawato, and R. Suzuki: *Formation and control of optimal trajectory in human multijoint arm movement - minimum torque-change model*. Biological Cybernetics, 61(2):89–101, 1989.
- [16] Wang, J. M., D. J. Fleet, and A. Hertzmann: *Gaussian process dynamical models for human motion*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(2):283–298, 2008.
- [17] Егоров, К. Ю. и И. В. Востриков: *Математическое моделирование движения руки и поведенческих движений*. В «Ломоносовские чтения — 2023», Москва, апрель 2023.
- [18] Журавлёв, В. Ф.: *Основы теоретической механики*. М: Физматлит, Москва, 2001.
- [19] Колубин, С. А.: *Динамика робототехнических систем*. Редакционно-издательский отдел Университета ИТМО, Санкт-Петербург, 2017.
- [20] Понтрягин, Л. С., В. Г. Болтянский, Р. В. Гамкредидзе, и Е. Ф. Мищенко: *Математическая теория оптимальных процессов*. М: Наука, Москва, 4-ая, стереотипная редакция, 1983.
- [21] Филиппов, А. Ф.: *Дифференциальные уравнения с разрывной правой частью*. М: Наука, Москва, 1985.

# Приложения

## Приложение А Таблица параметров

Для возможности сопоставления результатов в работе использовались общие параметры при построении численных решений задач. Ниже приведена таблица параметров, которые использовались для построения графиков в случае, если в подписи к соответствующему рисунку не оговорено обратное.

Символ	Краткое описание	Значение
--------	------------------	----------

Параметры модели		
$l$	Длины сочленений	[0,7, 0,7, 1,6]
$m$	Массы сочленений	[0,8, 0,8, 1,2]
$g$	Ускорение свободного падения	9,8

Параметры задачи		
$t_{\text{start}}$	Время начала движения	0
$t_{\text{final}}$	Время окончания движения	1
$\Delta t$	Шаг дискретизации	$10^{-3}$
$\theta^{\text{start}}$	Начальные углы	$[-1,4, -1,4, -1,4]$
$\dot{\theta}^{\text{start}}$	Начальные угловые скорости	[0, 0, 0]
$\tau^{\text{start}}$	Начальные моменты силы	[0, 0, 0]
$w_2$	Вес энергетического критерия	$10^{-2}$

Параметры основного алгоритма		
$\varepsilon$	Критерий останова	$10^{-2}$
$\mu$	Константа регуляризации матрицы $K$	$10^{-8}$
$\xi_1, \xi_2$	Интервал регуляризации поправки	$10^{-1}, 10$
$\gamma$	Мультипликатор регуляризатора $\eta$	0,5

Параметры алгоритма поиска начальной траектории		
$\hat{w}_1$	Вес фазового критерия	$10^{-2}$
$\hat{w}_2$	Вес критерия минимизации ускорения	$10^{-2}$

Задача «Целевое состояние» (Разделы 4 и 5)		
$\theta^{\text{final}}$	Целевые углы	$[-0,5, 1,1, 1,4]$
$\dot{\theta}^{\text{final}}$	Целевые угловые скорости	$[-5, -5, -5]$
$\tau^{\text{final}}$	Целевые моменты силы	$[0, 0, 0]$

Задача «Целевое положение схвата» (Подраздел 6.1)		
$e^{\text{final}}$	Целевое положение	$[1,2, 1,9]$

Задача «Целевая скорость схвата» (Подраздел 6.2)		
$e^{\text{final}}$	Целевое положение схвата	$[1,2, 1,9]$
$\dot{e}^{\text{final}}$	Целевая скорость схвата	$[-2,9, 0,5]$
$w_3$	Вес критерия скорости	0,5

Задача «Обход препятствия» (Подраздел 6.3)		
$e^{\text{obstacle}}$	Положение препятствия	$[2,2, -2,0]$
$r^{\text{obstacle}}$	Радиус препятствия	0,1