# Лекция 2. Постановка задачи оптимального управления

Д.А. Притыкин

14 февраля 2022

## Содержание



## Обозначения

$\mathbf{x} \in \mathbb{R}^n$ – вектор состояния.
$\mathbf{u} \in \mathbb{R}^m$ – вектор управляющих параметров.
$\mathbb{U} \in \mathbb{R}^m$ – множество допустимых управлений.
$t \in [t_0, t_к]$ – интервал времени.
$\mathbb{A}$ – начальное многообразие (множество начальных состояний).
$\mathbb{B}$ – терминальное многообразие (граничные условия на правом конце траектории).
$J = \Phi(\mathbf{x}(t_к), t_к)$ – терминальный функционал задачи оптимального управления.

# 1 Задача оптимального управления

Для корректной постановки задачи необходимы:

## 1.1 Модель системы управления

Модель представляет собой дифференциальное уравнение вида:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t), \tag{1}$$

где

- $\mathbf{x} \in \mathbb{R}^n$ – вектор состояния (вектор фазовых переменных)
- $\mathbf{u} \in \mathbb{U} \in \mathbb{R}^m$ - вектор управляющих параметров, $\mathbb{U}$ - область допустимых управлений.

Если зафиксировать $\mathbf{u}(t) = \tilde{\mathbf{u}}(t)$, то (1) превращается в $\dot{\mathbf{x}} = \tilde{\mathbf{f}}(\mathbf{x}, t)$, и при выборе начальных условий на $\mathbf{x}$: $\mathbf{x}(0) = \mathbf{x}_0$, получаем задачу Коши, для которой мы знаем о существовании её решения - траектории $\tilde{\mathbf{x}}(t)$ и при том единственной.

Так как мы можем выбирать различные управления $\mathbf{u}(t)$, существует много различных траекторий $\mathbf{x}(t)$., удовлетворяющих уравнению (1).

## 1.2 Начальное и терминальное многообразия

Определение 1.1. Начальное и терминальное многообразия – $\mathbb{A}$ и $\mathbb{B}$ – определяют граничные условия для модели (1) на левом и правом концах соответственно. Многообразия $\mathbb{A}$ и $\mathbb{B}$ задаются системами ограничений на переменные вектора состояния и время:

$$\mathbb{A} = \begin{cases} a_1(\mathbf{x}, t) = 0; \\ a_2(\mathbf{x}, t) = 0; \\ \ldots \\ a_\alpha(\mathbf{x}, t) = 0; \end{cases} \qquad \mathbb{B} = \begin{cases} b_1(\mathbf{x}, t) = 0; \\ b_2(\mathbf{x}, t) = 0; \\ \ldots \\ b_\beta(\mathbf{x}, t) = 0; \end{cases}$$

## 1.3 Цель управления

Необходимо знать ответ на вопрос: Зачем мы управляем системой? Целью управления является выбор такой функции $\mathbf{u}(t)$ (или $\mathbf{u}(\mathbf{x}, t)$), которое обеспечивает системе (1) требуемые свойства (например, устойчивость некоторого движения или оптимальность по некоторому показателю качества).

Пример. Можно потребовать, чтобы траектория системы управления $\mathbf{x}(t)$ повторяла некоторую наперед заданную кривую $\mathbf{r}(t)$ в пространстве состояний или была максимально близка к ней.

Постановка задачи оптимального управления подразумевает наличие оптимизируемого показателя качества или функционала.

Определение 1.2. Функционалом $\Phi$ будем называть отображение $\Phi : \mathbb{C} \to \mathbb{R}$.

Здесь и далее, если не оговорено иначе задачи оптимального управления будут включать терминальный функционал (вычисляемый на терминальном многообразии), причём целью задачи будет выбор управления $\mathbf{u}$, доставляющего функционалу

максимум:
$$J = \Phi(\mathbf{x}(t_к), t_к) \longrightarrow \max_{u \in U} \tag{2}$$

# 2 Пример задачи оптимального управления

Рассмотрим простейшую задачу о движениии материальной точки по горизонтальной гладкой прямой под действием горизонтальной управляющей силы. Эта задача, несмотря на её простоту иллюстрирует довольно часто встречающийся закон управления, называемый в англоязычной литературе Bang-bang control.

## 2.1 Условие и постановка задачи

Материальная точка массы $m$ может двигаться по гладкой горизонтальной оси под действием управляющей силы $\mathbf{F} : |F| \leq F^*$. Пусть в начальный момент времени $t = 0$ известны координата $x_0$ и скорость $v_0$ точки. Требуется за наименьшее время привести точку в состояние покоя в начале координат.

Приведем задачу к виду задачи теории управления.

Модель задачи составим с помощью 2 закона Ньютона:
$$\ddot{x} = F \quad \Rightarrow \quad \begin{cases} \dot{x} = v \\ \dot{v} = \frac{F}{m} \end{cases} \quad \Rightarrow \quad \begin{cases} \dot{x} = v \\ \dot{v} = u \end{cases},$$
где $|u| \leq u^*$ - управляющий параметр, абсолютная величина которого ограничена значением $u^*$.

Начальное многообразие - фиксированная точка начальных условий $(x_0, v_0)$:
$$\mathbb{A} : \forall x_0, v_0.$$

Терминальное многообразие - начало координат фазовой плоскости:
$$\mathbb{B} : x = 0, v = 0.$$

Функционал задачи - время, за которое система достигает терминального многообразия:
$$\Phi(x(t_к), t_к) = -t_к \longrightarrow \underset{u}{max}$$

## 2.2 Геометрическое решение

В качестве рабочей гипотезы предположим, что искомое управление $u(t)$ будет всегда принадлежать границе области допустимых управлений, то есть либо $u(t) = u^*$, либо $u(t) = -u^*$.

Случай 1 Пусть на некотором интервале $u(t) = u^*$, тогда уравнения движения системы:
$$\begin{cases} \dot{x} &= v \\ \dot{v} &= u^* \end{cases}$$
имеют решение
$$\begin{cases} x\left(t\right) = \frac{u^* t^2}{2} + v_0 t + x_0, \\ v\left(t\right) = u^* t + v_0. \end{cases}$$

Исключая из системы время, получим уравнение семейства траекторий в фазовой плоскости:

$$x = \frac{(v-v_0)^2}{2u^*} + \frac{(v-v_0)v_0}{u^*} + x_0 = \frac{v^2}{2u^*} + \left(x_0 - \frac{v_0^2}{2u^*}\right).$$

Таким образом при $u(t) = u^*$ система движется по одной из семейства парабол, графики которых приведены ниже.

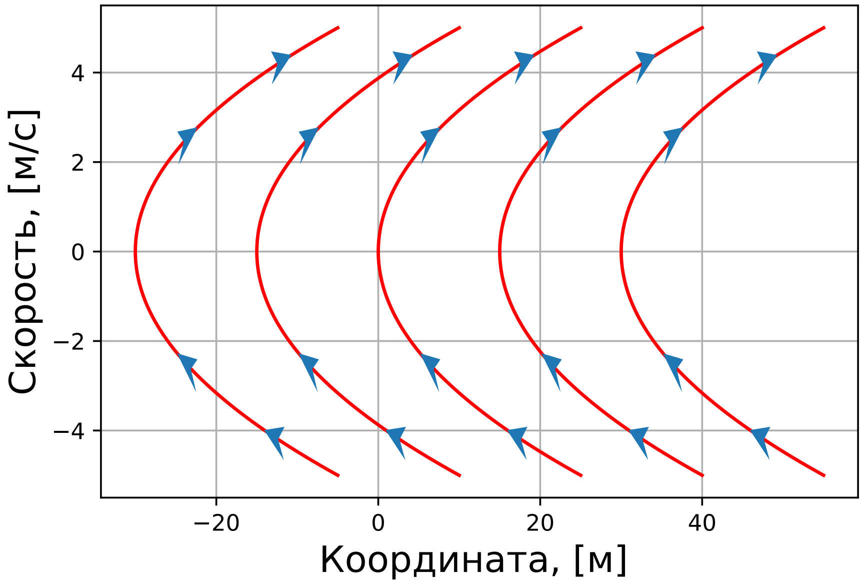


Рис. 1: Фазовые траектории при $u(t) = u^*$

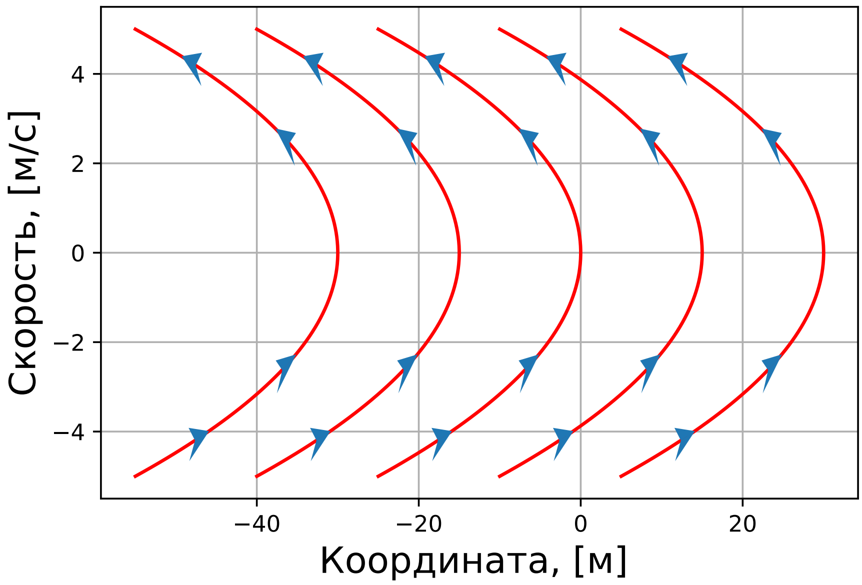


Рис. 2: Фазовые траектории при $u(t) = -u^*$

Случай 2 Пусть на некотором интервале $u(t) = -u^*$, тогда решение уравнений движения системы запишется так же как и для Случая 1, однако на фазовой плоскости их изображение будет представлено семейством парабол, ветви которых направлены в обратную сторону (поскольку коэффициент при $v^2$ меняет знак).

Таким образом, движение системы происходит по траекториям одного из двух семесйств парабол. Параболы, ветви которых направлены вправо, соответствуют u = $+u^*$. Параболы, ветви которых направлены влево, соответствуют u = -$u^*$.

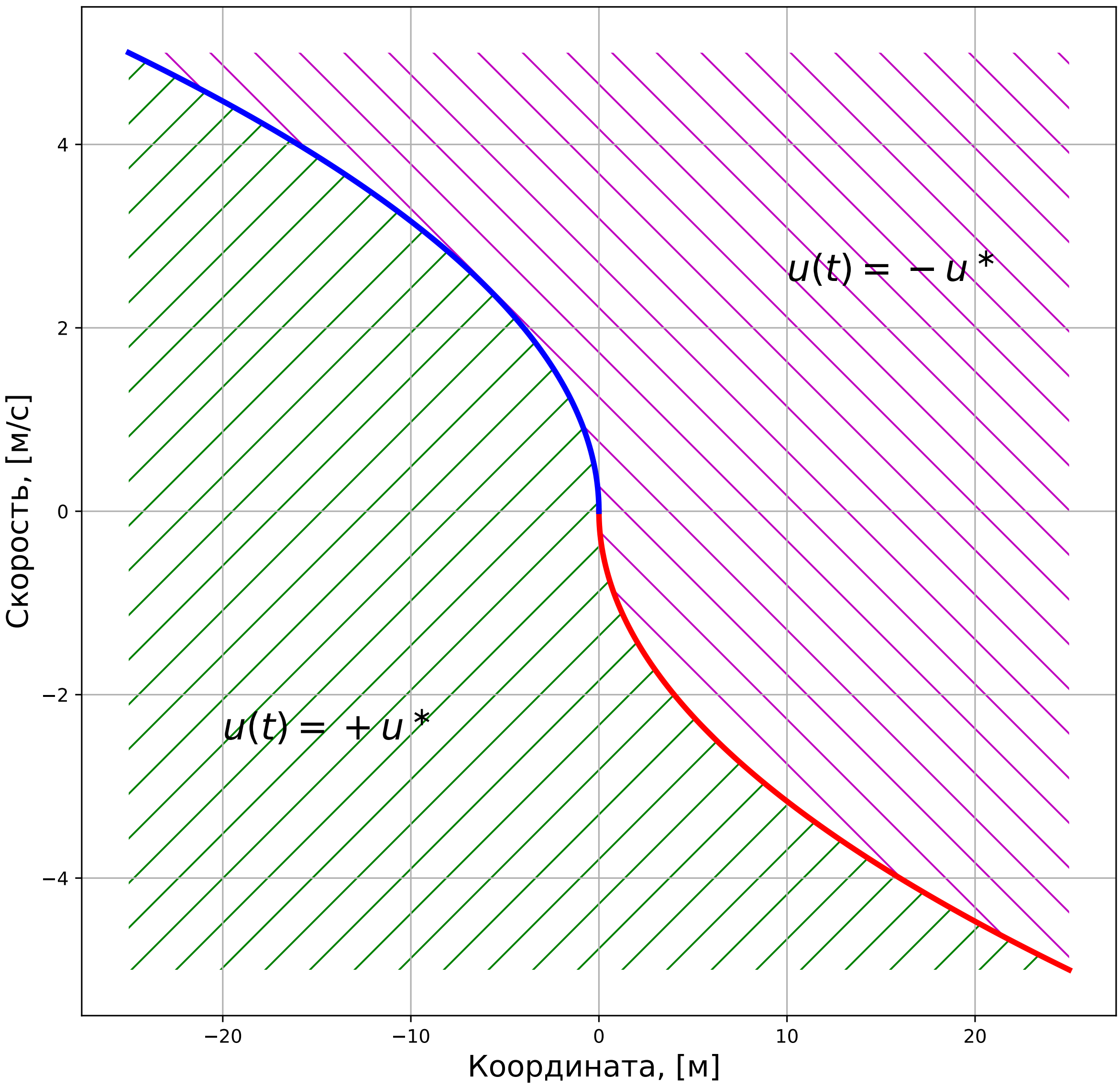


Рис. 3: Разбиение фазовой плоскости на области выбора начального управления

Если нам повезёт, точка начальных условий сразу попадёт на параболу, ведущую в начало координат (синяя ветвь с отрицательным управлением $x = -\frac{v^2}{2u^*}$ или красная ветвь с положительным управлением $x = \frac{v^2}{2u^*}$).

В противном случае, мы всегда можем выбрать начальную траекторию, пересекающуюся с одной из ветвей, ведущих в начало координат и переключить управление в точке пересечения.

Объединение синей и красной ветвей делит всю фазовую плоскость на две области. В верхней области любая точка начальных условий требует выбора $u(t) = -u^*$ на первом участке траектории, в нижней области - $u(t) = u^*$. При пересечении границы между областями, знак управления следует изменить, что приведёт к движению по ветви параболы, ведущей в начало координат.

Таким образом, закон управления строится объединением двух участков, на которых управление идёт по границе разрешённых значений и при смене участка меняет знак. Такой алгоритм управления в англоязычной литературе получил название "bang-bang-control"

> Замечание. Отметим, что для задач оптимального управления с терминальным функционалом, не зависящим от управляющих параметров, и моделью линейной по u, "bang-bang"управление довольно распространённый вид решения.

## 2.3 Пример

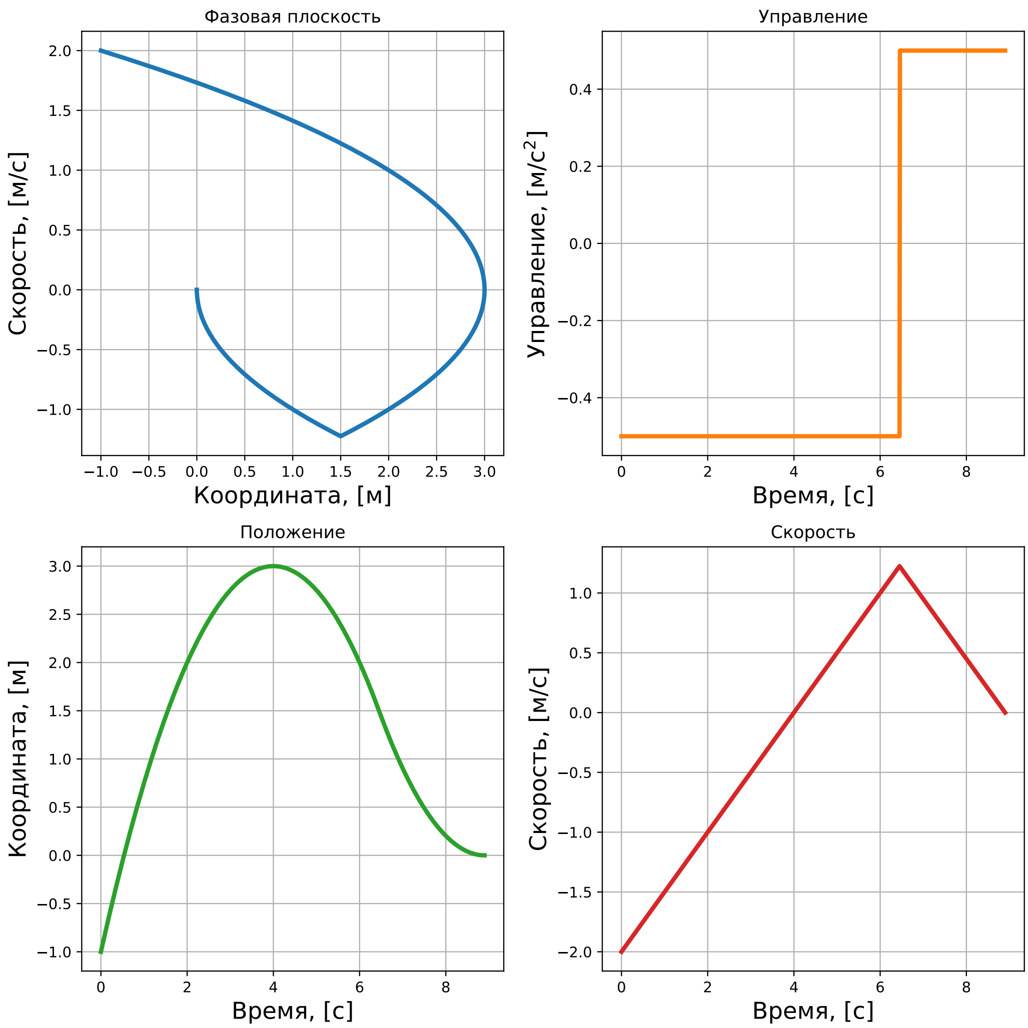


Рис. 4: Пример динамики управляемой системы

Проиллюстрируем изложенное выше геометрическое решение поставленной задачи. Для этого проинтегрируем систему (1) с начальными условиями $x_0 = -1$ м, $v_0 =$

2 м/с. В качестве ограничения на управляющее воздействие примем $u^* = 0.5$ м/с$^2$.

Графики, демонстриующие движение системы и закон управления приведены на рисунке 2.3.

## 2.4 Сравнение с другим законом управления

Для сравнения, попробуем применить к задаче ещё один распространенный подход к построению закона управления - управление по ошибке.

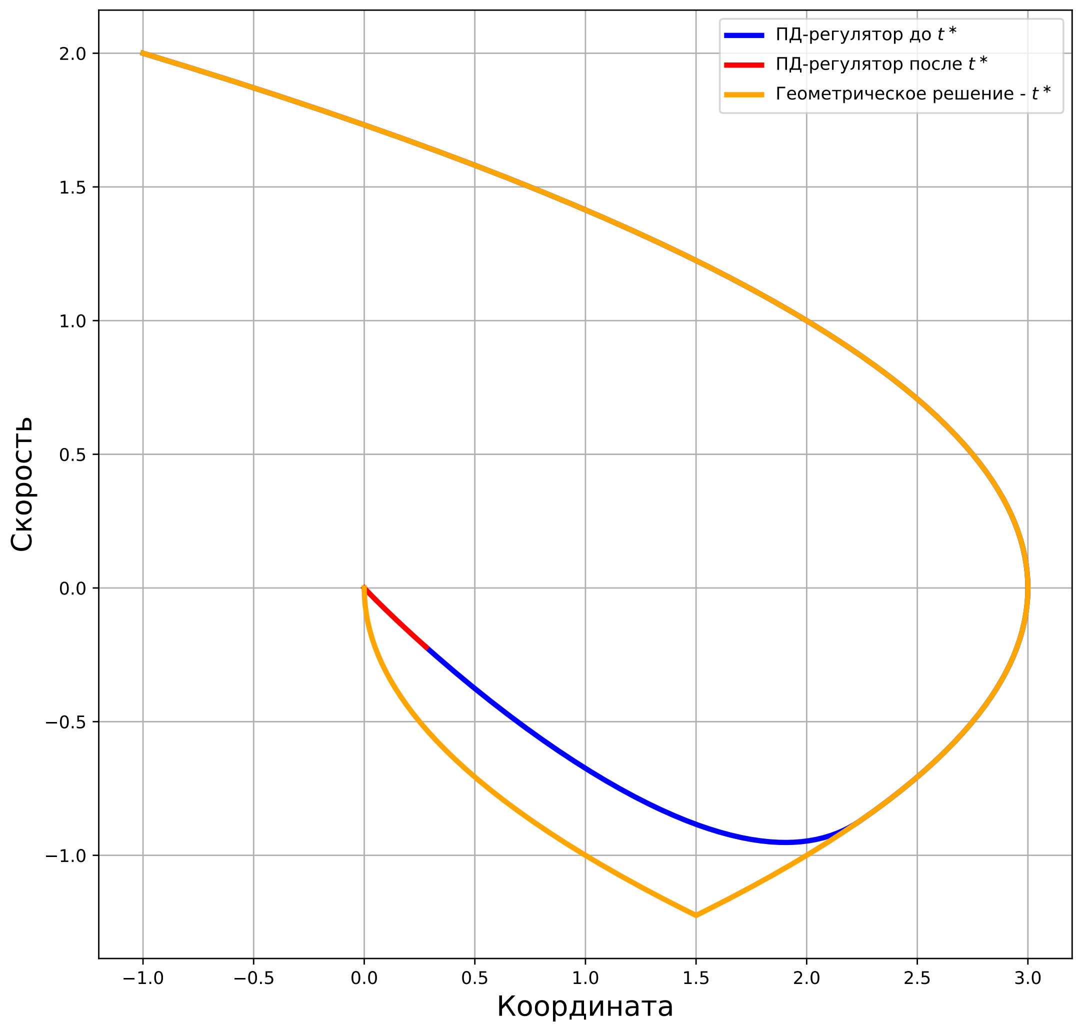


Рис. 5: Сравнение двух законов управления

Будем вычислять управление в каждый момент времени как сумму двух слагаемых:

- слагаемое пропорциональное ошибке положения точки $-k_p \cdot x$, где $k_p$ - некоторый настраиваемый коэффициент;
- слагаемое пропорциональное ошибке скорости $-k_d \cdot v$, где $k_d$ - некоторый настраиваемый коэффициент.

Таким образом, закон управления (так называемый ПД-регулятор, П - управление пропорциональное ошибке положения, Д - управление пропорциональное производной ошибки положения):

$$u(t) = -k_p \cdot x - k_d \cdot v. \tag{3}$$

Для примера, проиллюстрированного на рисунке ниже значения коэффициентов выбраны равными: $k_p = 1$, $k_d = 2$.

На рисунке 2.4 жёлтой линией показана фазовая траектория, соответствующая найденному аналитическому управлению, а составной (синей и красной) линией показана траектория соотвествующая управлению (3). Синяя линия заканчивается в тот момент времени $t^*$, когда оптимальная по быстродействию траектория достигает начала координат. Траектория, соответствующая управлению (3) продолжается (красная линия), чтобы продемонстрировать, что закон управления (3) также приводит систему на терминальное многообразие.