

Министерство образования и науки Российской Федерации  
Ярославский государственный университет им. П. Г. Демидова

**В. С. Климов**

**Одномерные  
вариационные  
задачи**

*Учебное пособие*

*Рекомендовано  
Научно-методическим советом университета  
для студентов, обучающихся по специальностям  
Математика и Прикладная математика и информатика*

Ярославль 2011

УДК 51  
ББК В161.8я73  
К49

*Рекомендовано  
Научно-методическим советом университета  
в качестве учебного издания. План 2010/2011 учебного года*

**Рецензенты:**

Смирнов Е. И., доктор педагогических наук, профессор, заведующий кафедрой математического анализа ЯГПУ им. К. Д. Ушинского;  
кафедра прикладной математики и вычислительной техники ЯГТУ

**К49 Климов, В. С. Одномерные вариационные задачи:** учебное пособие / В. С. Климов ; Яросл. гос. ун-т. им. П. Г. Демидова — Ярославль : ЯрГУ, 2011. — 140 с.

ISBN 978-5-8397-0794-8

Пособие «Одномерные вариационные задачи» содержит следующие разделы дисциплины «Вариационное исчисление и методы оптимизации»: гладкие решения одномерных вариационных задач, принцип максимума Понтрягина, дополнения и замечания.

Предназначено для студентов университетов, обучающихся по специальностям 010100.65 Математика и 010200.65 Прикладная математика и информатика (дисциплина «Вариационное исчисление и методы оптимизации», блок ОПД), очной формы обучения. Первая часть пособия может быть полезной и для студентов педагогических университетов.

**ISBN 978-5-8397-0794-8**

УДК 51  
ББК В161.8я73

©Ярославский государственный университет им. П. Г. Демидова, 2010

# Оглавление

<b>Предисловие</b>	<b>5</b>
<b>1 Уравнения Эйлера–Лагранжа</b>	<b>7</b>
1.1 Простейшая вариационная задача . . . . .	7
1.2 Модификации простейшей вариационной задачи . . . . .	21
1.3 Вариационные принципы . . . . .	34
<b>2 Принцип максимума Понтрягина</b>	<b>47</b>
2.1 Постановка задачи оптимального управления . . . . .	47
2.2 Оптимизация линейных систем . . . . .	50
2.3 Леммы . . . . .	62
2.4 Оптимизация нелинейных систем . . . . .	70
2.5 Управляемые процессы с неизвестным временем окончания . . . . .	77
2.6 Принцип максимума и вариационное исчисление . . . . .	82
<b>А Правило множителей Лагранжа</b>	<b>95</b>
<b>В О гладкости решений одномерных вариационных задач</b>	<b>103</b>
<b>С Метод условного градиента</b>	<b>117</b>
С.1 Сходимость метода условного градиента . . . . .	117
С.2 Приложения к задачам оптимального управления . . . . .	126
<b>Послесловие</b>	<b>137</b>
<b>Список литературы</b>	<b>138</b>



# Предисловие

Учебное пособие содержит изложение разделов дисциплины «Вариационное исчисление и методы оптимизации», изучаемых студентами третьего курса университетов специальности 010100.65 Математика и студентами четвёртого курса специальности 010200.65 Прикладная математика и информатика.

Весь материал разбит на две главы и три приложения. Первая глава посвящена классическому вариационному исчислению. Здесь выводятся уравнения Эйлера–Лагранжа для простейшей вариационной задачи, формулируются необходимые и достаточные условия слабого локального минимума. Далее изучаются различные обобщения простейшей вариационной задачи. Завершают главу вариационные принципы механики и физики: принцип Ферма и закон Снеллиуса преломления света, принцип Торичелли и задача о прогибе тяжёлой однородной нити, принцип наименьшего действия и геодезические линии на поверхности.

Во второй главе рассматриваются проблемы оптимального управления. Центральным результатом главы – принцип максимума Понтрягина. Вначале этот принцип доказывается для весьма специального случая: управляемый объект линеен по фазовому переменному, время окончания процесса фиксировано, правый конец траектории свободен, критерий качества управления линейно зависит от правого конца траектории. В этой ситуации принцип максимума устанавливается очень просто, он является и необходимым, и достаточным условием оптимальности, его применение не вызывает затруднений. Вместе с тем, разобравшись с данным частным случаем, читатель будет готов к изучению более сложных вопросов теории линейных управляемых систем: задача на быстродействие и задача терминального управления. Ещё труднее в техническом отношении разделы, связанные с принципом максимума для нелинейных управляемых систем. Предполагаемая подготовка читателя и здесь не выходит за пределы действующих программ, однако становятся желательными определённое терпение и научный энтузиазм.

В значительно большей степени указанные выше положительные качества нужны для усвоения приложений. Содержащийся здесь материал ориентирован прежде всего на читателя, желающего углубить свои познания по экстремальным задачам, найти подходящие темы для курсовых, дипломных и других квалификационных работ.

Принятый порядок изложения не самый экономный с точки зрения затрачи-

ваемого времени. Если поставить экономию времени во главу угла, то было бы логично вначале изложить принцип максимума для нелинейных управляемых систем, затем в качестве следствия получить основные результаты вариационного исчисления. Неоднократно проведённые педагогические эксперименты показали, что способ изложения вариационного исчисления должен во многом повторять историю его развития. Именно подобный принцип – от простого к сложному – автор и пытался реализовать в данном пособии.

Используются следующие обозначения:

$\emptyset$  – пустое множество;

$\forall$  – "для всех" или "для каждого";

$:=$  или  $\stackrel{def}{=}$  – "равно по определению";

◀ – начало доказательства; ▶ – конец доказательства;

$\frac{\partial u}{\partial x}$ ,  $u'_x$ ,  $u_x$  – частная производная функции  $u$  по аргументу  $x$ ;

$X^*$  – сопряжённое к банахову пространству  $X$ ;  $\langle x, x^* \rangle$  – значение линейного функционала  $x^* \in X^*$  на элементе  $x \in X$ ;

$\mathbb{R}^n$  –  $n$ -мерное арифметическое пространство вектор-столбцов  $x = (x_i)$  над полем  $\mathbb{R}$  действительных чисел;

$\mathbb{R}_n$  –  $n$ -мерное арифметическое пространство вектор-строк  $x = (x_i)$  над полем  $\mathbb{R}$  действительных чисел; пространство  $\mathbb{R}_n$  будет отождествляться с пространством  $(\mathbb{R}^n)^*$  линейных на  $\mathbb{R}^n$  функционалов;

$T: \mathbb{R}^n \rightarrow \mathbb{R}_n$  – оператор транспонирования, сопоставляющий вектор-столбцу  $x$  вектор-строку  $x^T$  с теми же компонентами; такое же обозначение используется для оператора транспонирования, действующего из  $\mathbb{R}_n$  в  $\mathbb{R}^n$ ;

$\int_a^b \varphi(t) dt$  – интеграл от функции  $\varphi: [a, b] \rightarrow \mathbb{R}$  по отрезку  $[a, b]$ ;

$\nabla g(x) = \left( \frac{\partial g}{\partial x_i}(x) \right)$  – градиент функции  $g$  в точке  $x$ , иначе говоря, вектор-

столбец, компоненты которого равны частным производным функции  $g$ ;

$g'(x) = (\nabla g(x))^T$  – соответствующая вектор-строка;

В пособии принята автономная (в пределах каждой главы своя) нумерация секций, разбитых на отдельные пункты. Формулы (теоремы, упражнения и т. п.) нумеруются в пределах каждой секции. При ссылках внутри секции указывается лишь номер соответствующей формулы (теоремы и т. п.); в противном случае приводится и номер секции. Например, формула 2.3(1) – это формула (1) из секции 2.3; лемма 1.2.3 – это лемма 3 из секции 1.2.

Многие разделы пособия обсуждались с коллегами из Ярославского государственного университета им. П. Г. Демидова, а также с математиками других вузов. Приношу всем им самую искреннюю признательность. Буду благодарен за указания на возможные ошибки в тексте, ответственность за которые автор полностью берёт на себя.

Климов В. С., доктор физико-математических наук.

# Глава 1

## Уравнения Эйлера–Лагранжа

### 1.1 Простейшая вариационная задача

**1. Постановка простейшей вариационной задачи.** Многие из результатов, относящихся к конечномерной оптимизации, имеют бесконечномерные аналоги. Возникающие здесь новые моменты проявляются при рассмотрении простейшей вариационной задачи. Прежде чем переходить к формализациям, рассмотрим две задачи, возбуждившие в своё время (17-й век) общий интерес среди математиков.

Наиболее известной была **задача о брахистохроне (кривой наискорейшего ската)**, поставленная в 1696 году И. Бернулли и формулируемая следующим образом: *среди всевозможных плоских кривых, лежащих в вертикальной плоскости и соединяющих две точки  $A$  и  $B$ , не расположенные на общей вертикали, найти ту, по которой материальная точка под действием только силы тяжести и без начальной скорости скатилась бы в кратчайшее время.* Для аналитической формулировки задачи примем за ось  $Ox$  горизонтальную прямую, а ось  $Oy$  направим вертикально вниз. Тогда точки  $A$  и  $B$  будут иметь соответственно координаты  $(a, 0)$  и  $(b, y_1)$ . Так как точка движется из  $A$  без начальной скорости, то её скорость  $v$  связана с её ординатой  $y$  соотношением  $v^2 = 2gy$ , где  $g$  – ускорение силы тяжести, или  $v = \sqrt{2gy}$ .

Пусть  $y = y(x)$  есть уравнение кривой, по которой движется точка из  $A$  в  $B$ . Скорость движения точки  $v = \frac{ds}{dt}$ , где  $s$  – путь,  $t$  – время. Поскольку  $ds = \sqrt{1 + y'^2} dx$ , то

$$dt = \frac{ds}{v} = \sqrt{\frac{1 + y'^2}{2gy}} dx.$$

Полное время  $T$  движения из  $A$  в  $B$  выражается интегралом

$$T(y) = \int_a^b \sqrt{\frac{1 + y'^2}{2gy}} dx.$$

Очевидно,  $T(y)$  есть функционал, зависящий от функции  $y$ . Требуется найти функцию  $y(x)$ , для которой  $T$  принимает наименьшее значение: к сравнению допускаются функции  $y(x)$ , определённые на отрезке  $[a, b]$  и удовлетворяющие краевым условиям  $y(a) = 0, y(b) = y_1$ .

Задача о брахистохроне родственна следующей физической задаче иной природы: *в прозрачной среде с переменной оптической плотностью даны две точки  $A$  и  $B$ , требуется определить траекторию луча света, идущего от точки  $A$  к точке  $B$* . Эта задача сводится к задаче на разыскание экстремума на основании принципа Ферма: *из всех кривых, соединяющих точки  $A$  и  $B$ , траектория луча света есть линия, распространяясь вдоль которой свет придёт из  $A$  в  $B$  в кратчайший срок*.

Рассмотрим плоский случай. Примем за плоскость распространения света плоскость  $xOy$ . Пусть  $(x_0, y_0)$  и  $(x_1, y_1)$  — координаты точек  $A$  и  $B$ , а  $y = y(x)$  ( $x_0 \leq x \leq x_1$ ) есть некоторая кривая, соединяющая эти точки. Повторяя рассуждения, проведённые в предыдущем примере, получим: время  $T$  распространения света вдоль кривой  $y = y(x)$  из  $A$  в  $B$  выражается интегралом

$$T(y) = \int_{x_0}^{x_1} \frac{\sqrt{1 + y'^2}}{v[x, y(x)]} dx;$$

тем самым задача определения траектории луча света сводится к нахождению функции  $y$ , для которой функционал  $T(y)$  принимает наименьшее значение.

Простейшая вариационная задача состоит в минимизации интегрального функционала

$$f(x) = \int_{\Delta} L(t, x(t), x'(t)) dt$$

на совокупности функций, определённых на отрезке

$$\Delta = [t_0, t_1], \quad (-\infty < t_0 < t_1 < \infty)$$

и удовлетворяющих краевым условиям  $x(t_0) = x_0, x(t_1) = x_1$ . Она будет записываться следующим образом

$$f(x) = \int_{\Delta} L(t, x(t), x'(t)) dt \rightarrow \min, \quad (1)$$

$$x(t_0) = x_0, \quad x(t_1) = x_1.$$

Задача (1) поставлена недостаточно чётко, поскольку не определён класс функций, на котором минимизируется функционал  $f$ , не указаны предположения относительно функции  $L$ , называемой интегрантом функционала  $f$ . Ниже считаем, что функция  $L$  определена и непрерывна по совокупности переменных



на множестве  $\Delta \times W$ , где  $W$  — открытое подмножество плоскости  $\mathbb{R}^2$  ( $W \neq \emptyset$ , случай  $W = \mathbb{R}^2$  не исключается; при первом чтении можно ограничиться только этим частным случаем). Иногда к функции  $L$  будут предъявляться более жёсткие требования.

Пусть  $m$  — натуральное число. Обозначим через  $C^m(\Delta)$  совокупность  $m$  раз непрерывно дифференцируемых на отрезке  $\Delta$  функций. При естественном определении суммы функций и умножения функции на действительное число класс  $C^m(\Delta)$  образует линейное пространство. Если в  $C^m(\Delta)$  ввести норму

$$\|x\|_m := \sum_{i=0}^m \max_{t \in \Delta} |x^{(i)}(t)|,$$

то  $C^m(\Delta)$  становится банаховым пространством.

Положим

$$X := \{x \in C^1(\Delta), \quad (x(t), x'(t)) \in W \forall t \in \Delta, \quad x(t_0) = x_0, \quad x(t_1) = x_1\}.$$

Для каждой функции  $x$  из множества  $X$  функция  $L(t, x(t), x'(t))$  определена и непрерывна на отрезке  $\Delta$ . Поэтому функционал  $f$  определён на множестве  $X$ . Элемент  $\hat{x}$  из  $X$  называют точкой абсолютного минимума функционала  $f$  на множестве  $X$  (решением задачи (1)), если  $f(\hat{x}) \leq f(x) \forall x \in X$ . Элемент  $\hat{x}$  из  $X$  именуют слабым локальным решением задачи (1), если существует такое  $\delta > 0$ , что  $f(\hat{x}) \leq f(x) \forall x \in X, \|x - \hat{x}\|_1 < \delta$ . Элемент  $\hat{x}$  из  $X$  называют сильным локальным решением задачи (1), если существует такое  $\delta > 0$ , что  $f(\hat{x}) \leq f(x) \forall x \in X, \|x - \hat{x}\|_0 < \delta$ .

Ясно, что каждое решение  $\hat{x}$  задачи (1) реализует сильный локальный минимум задачи (1), а каждое сильное локальное решение этой задачи одновременно является и слабым локальным решением. В пределах этой главы сильные локальные решения не будут рассматриваться, поэтому эпитет „слабое“ опускается.

Основное внимание далее уделяется необходимым условиям локального минимума в задаче (1). С меньшей полнотой обсуждаются достаточные условия локального минимума и совсем мало — условия абсолютного минимума.

**2. Вспомогательные предложения.** Для анализа задачи (1) потребуются вспомогательные предложения, представляющие самостоятельный интерес. Ниже  $C_0^1(\Delta) := \{x \in C^1(\Delta), x(t_0) = x(t_1) = 0\}$  — совокупность непрерывно дифференцируемых на отрезке  $\Delta = [t_0, t_1]$  функций, обращающихся в нуль на границе отрезка  $\Delta$ .

**Лемма 1 (Лемма Дюбуа–Реймона).** Пусть  $a_0(t), a_1(t)$  — непрерывные на отрезке  $\Delta$  функции. Пусть для любой функции  $h$  из  $C_0^1(\Delta)$  выполнено равенство

$$\int_{\Delta} (a_1(t)h'(t) + a_0(t)h(t)) dt = 0. \quad (2)$$

Тогда функция  $a_1(t)$  в каждой точке отрезка  $\Delta$  дифференцируема и  $a_1'(t) = a_0(t) \forall t \in \Delta$ .

◀ Вначале рассмотрим случай, когда  $a_0(t) \equiv 0$ . Подберём константу  $C$  так, чтобы дифференциальное уравнение

$$h'(t) = a_1(t) - C$$

имело решение класса  $C_0^1(\Delta)$ . Это возможно, если

$$\int_{\Delta} (a_1(t) - C) dt = 0, \quad \text{т.е.} \quad C = \frac{1}{t_1 - t_0} \int_{\Delta} a_1(t) dt.$$

Справедливы равенства

$$\int_{\Delta} [a_1(t) - C]^2 dt \stackrel{(I)}{=} \int_{\Delta} (a_1(t) - C)h'(t) dt \stackrel{(II)}{=} \int_{\Delta} a_1(t)h'(t) dt \stackrel{(III)}{=} 0.$$

Действительно, (I) следует из равенства  $h'(t) = a_1(t) - C$ , (II) вытекает из включения  $h \in C_0^1(\Delta)$ , (III) следует из (2). Функция  $(a_1(t) - C)^2$  неотрицательна, непрерывна на отрезке  $\Delta$ , и интеграл от неё равен 0. Это возможно лишь в случае, когда  $a_1(t) \equiv C$ ,  $a_1'(t) \equiv 0 \equiv a_0(t)$ .

В общем случае положим

$$A_0(t) := \int_{t_0}^t a_0(s) ds.$$

Тогда верны равенства

$$\begin{aligned} \int_{\Delta} a_0(t)h(t) dt &= \int_{\Delta} A_0'(t)h(t) dt = - \int_{\Delta} A_0(t)h'(t) dt, \\ \int_{\Delta} (a_1(t)h'(t) + a_0(t)h(t)) dt &= \int_{\Delta} [a_1(t) - A_0(t)]h'(t) dt = 0 \end{aligned}$$

для любой функции  $h$  из  $C_0^1(\Delta)$ . В силу уже доказанного функция  $a_1(t) - A_0(t)$  постоянна, что и приводит к требуемому результату. ►

**Предложение 1 (Правило Лейбница дифференцирования интегралов, зависящих от параметра).** Пусть функция  $\Omega(t, \xi)$  ( $t \in \Delta, \xi \in (\alpha, \beta)$ ) определена и непрерывна на  $\Delta \times (\alpha, \beta)$  вместе с частной производной  $\Omega'_\xi(t, \xi)$ . Тогда функция

$$I(\xi) := \int_{\Delta} \Omega(t, \xi) dt$$

непрерывно дифференцируема на интервале  $(\alpha, \beta)$  и

$$I'(\xi) = \int_{\Delta} \Omega'_{\xi}(t, \xi) dt. \quad (3)$$

Доказательство предложения 1 приведено, например, в [1]. Аналогичные (3) формулы справедливы для производных высших порядков. В качестве примера применения предложения 1 рассмотрим функцию

$$I(\xi) = f(x + \xi h) = \int_{\Delta} L[t, x(t) + \xi h(t), x'(t) + \xi h'(t)] dt, \quad (4)$$

где  $x \in X$ ,  $h \in C_0^1(\Delta)$ . Очевидно, что функция  $I(\xi)$  определена и непрерывна на некоторой окрестности точки 0. При дополнительных предположениях гладкости функции  $L$  можно установить дифференцируемость функции  $I(\xi)$ . Как и в конечномерном случае, число

$$f'(x, h) := \left. \frac{d}{d\xi} f(x + \xi h) \right|_{\xi=0}$$

назовём производной функционала  $f$  в точке  $x$  по направлению  $h$ .

**Лемма 2.** Пусть функция  $L(t, x, u)$  и её частные производные

$$L'_x(t, x, u), \quad L'_u(t, x, u)$$

определены и непрерывны по совокупности переменных на множестве  $\Delta \times W$ . Тогда определяемая равенством (4) функция  $I(\xi)$  дифференцируема в точке 0 и

$$I'(0) = f'(x, h) = \int_{\Delta} [L'_u(t, x(t), x'(t))h'(t) + L'_x(t, x(t), x'(t))h(t)] dt. \quad (5)$$

◀ Для доказательства леммы положим

$$\Omega(t, \xi) = L(t, x(t) + \xi h(t), x'(t) + \xi h'(t))$$

и воспользуемся равенством

$$\Omega'_{\xi}(t, 0) = L'_u(t, x(t), x'(t))h'(t) + L'_x(t, x(t), x'(t))h(t),$$

объединяя которое с формулой (3), приходим к доказываемому утверждению. ▶

**Лемма 3.** Пусть функция  $L(t, x, u)$  и её частные производные второго порядка  $L''_{xx}(t, x, u)$ ,  $L''_{xu}(t, x, u)$ ,  $L''_{uu}(t, x, u)$  определены и непрерывны по совокупности переменных на множестве  $\Delta \times W$ . Тогда функция  $I(\xi)$  дважды дифференцируема в точке 0 и справедливы соотношения

$$I''(0) = \left. \frac{d^2}{d\xi^2} f(x + \xi h) \right|_{\xi=0} = f''(x, h) =$$

$$= \int_{\Delta} [a_{00}(t)h^2(t) + 2a_{01}(t)h(t)h'(t) + a_{22}(t)h'^2(t)] dt, \quad (6)$$

в которых

$$\begin{aligned} a_{00}(t) &= L''_{xx}(t, x(t), x'(t)), & a_{01}(t) &= L''_{xu}(t, x(t), x'(t)), \\ a_{11}(t) &= L''_{uu}(t, x(t), x'(t)). \end{aligned} \quad (7)$$

◀ Доказательство основано на двухкратном применении правила Лейбница дифференцирования интегралов, зависящих от параметра. ▶

Далее используется следующий вариант теоремы о неявной функции.

**Предложение 2.** Пусть  $\Phi(t, u)$  – функция, определенная и непрерывно дифференцируемая по совокупности переменных при  $t \in \Delta, u_0 < u < u_1$ . Пусть  $y \in C(\Delta)$ ,  $u_0 < y(t) < u_1$  ( $t \in \Delta$ ) и

$$\Phi[t, y(t)] = 0, \quad \Phi'_u[t, y(t)] \neq 0 \quad \forall t \in \Delta.$$

Тогда  $y \in C^1(\Delta)$ .

Доказательство этого утверждения можно найти в любом учебнике математического анализа (см. также [1], [5]).

**3. Уравнение Эйлера.** В этом пункте предполагается, что функция  $L : \Delta \times W \rightarrow \mathbb{R}$  и её частные производные  $L'_x, L'_u$  непрерывны по совокупности переменных.

**Теорема 1.** Пусть  $\hat{x}$  – локальное решение задачи (1),

$$\widehat{L}_x(t) := L'_x(t, \hat{x}(t), \hat{x}'(t)), \quad \widehat{L}_u(t) := L'_u(t, \hat{x}(t), \hat{x}'(t)).$$

Тогда функция  $\widehat{L}_u(t)$  дифференцируема всюду на отрезке  $\Delta$  и

$$\frac{d}{dt} \widehat{L}_u(t) = \widehat{L}_x(t). \quad (8)$$

◀ Пусть  $h \in C_0^1(\Delta)$ . При достаточно малом действительном  $\xi$  функция  $\hat{x} + \xi h$  принадлежит классу  $X$ . Так как  $\hat{x}$  – локальное решение задачи (1), то найдётся такое  $\delta > 0$ , что  $f(\hat{x}) \leq f(x) \forall x \in X, \|x - \hat{x}\|_1 < \delta$ . В частности, при малых  $\xi$  справедливо неравенство  $f(\hat{x} + \xi h) \geq f(\hat{x})$ , т.е. 0 есть точка локального минимума функции  $I(\xi) = f(\hat{x} + \xi h)$ . В силу теоремы Ферма  $I'(0) = 0$ . В соединении с формулами (5) это приводит к соотношению

$$f'(\hat{x}, h) = \int_{\Delta} [\widehat{L}_u(t)h'(t) + \widehat{L}_x(t)h(t)] dt = 0 \quad \forall h \in C_0^1(\Delta).$$

Последнее равенство и лемма 1 влекут за собой доказываемый результат. ▶

Теорема 1 означает, что  $\hat{x}$  есть решение дифференциального уравнения

$$\frac{d}{dt}L'_u = L'_x, \quad (9)$$

называемого уравнением Эйлера. В качестве примера рассмотрим функционал

$$g(x) = \int_{\Delta} [a(t)x'^2(t) + b(t)x^2(t) - 2c(t)x(t)] dt, \quad (10)$$

где  $a, b, c$  – непрерывные на отрезке  $\Delta$  функции. Функционал (10) есть сумма квадратичного функционала

$$\mathcal{K}(x) := \int_{\Delta} [a(t)x'^2(t) + b(t)x^2(t)] dt \quad (11)$$

и линейного функционала

$$l(x) := - \int_{\Delta} 2c(t)x(t) dt.$$

В этом случае интегрант  $L$  функционала  $g$  задаётся равенством  $L(t, x, u) = a(t)u^2 + b(t)x^2 - 2c(t)x$ , поэтому

$$L'_x = 2b(t)x - 2c(t), \quad L'_u = 2a(t)u.$$

Соответствующее уравнение Эйлера имеет вид

$$-\frac{d}{dt} \left( a(t) \frac{dx}{dt} \right) + b(t)x(t) = c(t).$$

Оно эквивалентно интегро-дифференциальному уравнению

$$a(t) \frac{dx}{dt} = \int_{t_0}^t [b(s)x(s) - c(s)] ds + const. \quad (12)$$

Если  $a(t) \neq 0 \forall t \in \Delta$  и функция  $a(t)$  всюду непрерывно дифференцируема, то из (12) вытекают включения  $x' \in C^1(\Delta), x \in C^2(\Delta)$ . В этих предположениях уравнение Эйлера записывается в виде

$$-a(t)x''(t) - a'(t)x'(t) + b(t)x(t) = c(t).$$

Предшествующие рассуждения обобщаются на широкий класс интегрантов  $L$ . Решения уравнения Эйлера (9) называются экстремалими функционала  $f$ . В общем случае экстремаль есть функция класса  $C^1(\Delta)$ , удовлетворяющая интегро-дифференциальному уравнению

$$L'_u(t, x(t), x'(t)) = \int_{t_0}^t L'_x(s, x(s), x'(s)) ds + const. \quad (13)$$

Экстремаль  $x^*$  функционала  $f$ , удовлетворяющую граничным условиям  $x(t_0) = x_0, x(t_1) = x_1$ , иногда называют критической точкой, а число  $f(x^*)$  – критическим значением  $f$ .

**Теорема 2.** Пусть  $L : \Delta \times W \rightarrow \mathbb{R}$  дважды непрерывно дифференцируемая функция,  $x(t)$  – экстремаль функционала  $f$  и

$$L''_{uu}(t, x(t), x'(t)) \neq 0 \forall t \in \Delta. \quad (14)$$

Тогда  $x \in C^2(\Delta)$ .

◀ Введём в рассмотрение функцию

$$\Phi(t, u) := L(t, x(t), u) - \int_{t_0}^t L_x(s, x(s), x'(s)) ds - const.$$

Тогда функция  $\Phi$  непрерывно дифференцируема по переменным  $t, u$  в естественной области определения. Функция  $y(t) = x'(t)$  непрерывна на отрезке  $\Delta$  и удовлетворяет уравнению  $\Phi(t, y(t)) = 0 \forall t \in \Delta$ . Очевидно, что

$$\Phi'_u(t, y(t)) = L''_{uu}(t, x(t), x'(t)) \neq 0.$$

Согласно предложению 2 функция  $y$  принадлежит классу  $C^1(\Delta)$ . Это эквивалентно включению  $x \in C^2(\Delta)$ . ▶

Теорема 2 принадлежит Гильберту. Неравенство (14) называют условием Лежандра. Теорема Гильберта выражает повышенную гладкость экстремалей простейшей вариационной задачи. Формальное дифференцирование (13) по  $t$  приводит к уравнению  $L''_{ut} + L''_{ux}x' + L''_{uu}x'' = L'_x$ . Если  $L''_{uu} \neq 0$ , то уравнение Эйлера эквивалентно дифференциальному уравнению второго порядка

$$x'' = \frac{L_x - L''_{ux}x' - L''_{ut}}{L''_{uu}}.$$

В некоторых случаях порядок уравнения Эйлера может быть понижен.

1°. Функция  $L$  не зависит от  $u$ . Уравнение Эйлера представляет соотношение  $L'_x(t, x) = 0$ , т.е. не является дифференциальным уравнением.

2°. Функция  $L$  не зависит от  $x$ . Тогда  $L'_x = 0$  и уравнение Эйлера допускает понижение порядка

$$L'_u(t, x'(t)) = C_1. \quad (15)$$

Соотношение (15) называют интегралом импульса. В задачах механики такого рода соотношения выражают закон сохранения импульса. Функция  $L'_u$  не обязана быть постоянной; она постоянна лишь вдоль экстремалей.

3°. Функция  $L$  не зависит от  $t$ . Тогда для любой экстремали  $x(t)$  выполняется соотношение

$$x'L'_u(x, x') - L(x, x') = C_1. \quad (16)$$

Для доказательства (16) достаточно взять производную от левой части (16). Имеем

$$x''L'_u(x, x') + x' \frac{d}{dt} L'_u(x, x') - L'_x(x, x')x' - L'_u(x, x')x'' = 0.$$

Равенство (16) называют интегралом энергии.

**Замечание.** Выше всюду рассматривался функционал  $f$ , определённый на функциях  $x(t)$  переменного  $t$ , интерпретируемого как время. В ряде задач независимое переменное имеет характер величины, распределённой в пространстве. Например, это так для упоминавшейся выше задачи о брахистохроне. Поэтому представляется естественным обозначить независимое переменное не через  $t$ , а через  $x$ . Простейшая вариационная задача часто записывается в следующей форме:

$$J(y) = \int_a^b L(x, y(x), y'(x)) dx \rightarrow \min, \quad y(a) = A, y(b) = B. \quad (17)$$

Для задачи (17) очевидным образом вводятся понятия решения и локального решения. Уравнение Эйлера имеет вид

$$\frac{d}{dx} L'_z = L'_y. \quad (18)$$

Его решения называют экстремальными функционала  $J$ . Если функция  $L$  не зависит от одного из переменных, то уравнение Эйлера допускает понижение порядка. В частности, имеют место естественные аналоги равенств (15), (16)

$$L'_z(x, y'(x)) = C_1, \quad \text{соответственно,} \quad y' L'_z(y, y') - L(y, y') = C_1,$$

за которыми мы сохраним уже не столь оправданные названия интеграла импульса и интеграла энергии.

**4. Достаточные условия абсолютного минимума.** Пусть  $\hat{x}$  – экстремаль функционала  $f$  и  $\hat{x} \in X$ . Тогда число 0 есть критическая точка функции  $I(\xi) = f(\hat{x} + \xi h)$  для любой функции  $h$  из  $C_0^1(\Delta)$ . Если  $W$  – выпуклое подмножество плоскости  $\mathbb{R}^2$ , то  $X$  – выпуклое подмножество пространства  $C^1(\Delta)$ . Назовём функционал  $f$  выпуклым, если при любых  $u, v$  из  $X$  и любого числа  $\lambda$  из  $(0, 1)$  имеет место неравенство

$$f[(1 - \lambda)u + \lambda v] \leq (1 - \lambda)f(u) + \lambda f(v).$$

**Лемма 4.** Если  $f : X \rightarrow \mathbb{R}$  – выпуклый функционал,  $\hat{x}$  – экстремаль функционала  $f$  и  $\hat{x} \in X$ , то  $\hat{x}$  – решение задачи (1), т.е.  $f(\hat{x}) \leq f(x)$   $\forall x \in X$ .

◀ Пусть  $x \in X$ ,  $h = x - \hat{x}$ ,  $I(\xi) := f(\hat{x} + \xi h)$ . Из выпуклости функционала  $f$  вытекает выпуклость функции  $I(\xi)$ . Так как  $\hat{x}$  – экстремаль функционала  $f$ , то  $I'(0) = f'(\hat{x}, h) = 0$ . Следовательно, 0 – точка абсолютного минимума функции  $I$ . В частности,  $I(1) \geq I(0)$ , что эквивалентно неравенству  $f(x) \geq f(\hat{x})$ . ▶

**Теорема 3.** Пусть функции  $L, L'_x, L'_u$  непрерывны по совокупности переменных и при любых  $t$  из  $\Delta$  функция  $L(t, \cdot, \cdot) : W \rightarrow \mathbb{R}$  выпукла. Если  $\hat{x}$  – экстремаль функционала  $f$  и  $\hat{x} \in X$ , то  $\hat{x}$  – решение задачи (1).

◀ Выпуклость функции  $L(t, \cdot, \cdot) : W \rightarrow \mathbb{R}$  влечёт за собой выпуклость функционала  $f : X \rightarrow \mathbb{R}$ . Теперь теорема 3 следует из леммы 4. ▶

Применим теорему 3 к определяемому равенством (10) интегральному функционалу. В данном случае  $L(t, x, u) = a(t)u^2 + b(t)x^2 - 2c(t)x$ ,  $W = \mathbb{R}^2$ . Функция  $L(t, \cdot, \cdot)$  выпукла по совокупности переменных  $x, u$ , если и только если  $a(t) \geq 0, b(t) \geq 0$ .

Более точный (но менее обозримый) результат получается, если вместо теоремы 3 использовать лемму 4. В рассматриваемом случае  $I(\xi) = g(x + \xi h)$ , ( $x \in X, h \in C_0^1(\Delta)$ ) есть квадратичная функция скалярного переменного  $\xi$  и  $I''(\xi) = 2\mathcal{K}(h)$ , где  $\mathcal{K}$  – определяемый равенством (11) интегральный функционал. Если  $\mathcal{K}(h) \geq 0 \forall h \in C_0^1(\Delta)$ , то  $I''(\xi) \geq 0$ ,  $I$  – выпуклая функция,  $g : X \rightarrow \mathbb{R}$  – выпуклый функционал. Если  $\hat{x}$  – экстремаль функционала  $g$  и  $\hat{x} \in X$ , то согласно лемме 4  $g(\hat{x}) \leq g(x) \forall x \in X$ . Условие неотрицательной определённости функционала  $\mathcal{K}$  может выполняться и для отрицательной функции  $b(t)$ . Например, функционал

$$\mathcal{K}_0(h) = \int_0^T (h'^2(t) - h^2(t)) dt$$

неотрицательно определён, если  $0 < T \leq \pi$  (см. следующий пункт). Вместе с тем справедлива

**Лемма 5.** Если квадратичный функционал  $\mathcal{K}$  неотрицательно определён, то  $a(t) \geq 0 \forall t \in \Delta$ .

◀ В предположении противного найдутся числа  $m > 0, \tau \in \mathbb{R}, \delta > 0$ , что  $[\tau - \delta, \tau + \delta] \subset (t_0, t_1)$  и  $a(t) < -m \forall t \in [\tau - \delta, \tau + \delta]$ . Введём в рассмотрение функцию  $\varphi$ , определённую на действительной прямой равенством

$$\varphi(t) = \begin{cases} \exp(\frac{t^2}{t^2-1}), & \text{если } |t| < 1 \\ 0, & \text{если } |t| \geq 1. \end{cases}$$

Функция  $\varphi$  бесконечно дифференцируема на всей прямой, положительна на интервале  $(-1, 1)$  и равна нулю вне него. При достаточно больших натуральных числах  $n$  функции  $h_n(t) = \varphi(n(t - \tau))$  принадлежат пространству  $C_0^1(\Delta)$ , при этом  $h'_n(t) = n\varphi'(n(t - \tau)), |h_n(t)| \leq 1$ . Справедливы соотношения

$$\begin{aligned} \mathcal{K}(h_n) &= \int_{n|t-\tau|<1} (a(t)h_n'^2(t) + b(t)h_n^2(t)) dt \leq \\ &\leq \int_{n|t-\tau|<1} (a(t)n^2\varphi'^2(n(t - \tau)) dt + c_0 \leq -mn^2 \int_1^1 \varphi'^2 dt + c_0, \end{aligned}$$



постоянная  $c_0$  не зависит от  $n$ . Ясно, что  $\mathcal{K}(h_n) < 0$  при достаточно больших  $n$ . Полученное противоречие влечёт за собой неравенство  $a(t) \geq 0$ . ►

**5. Положительная определённость интегрального квадратичного функционала.** В этом пункте приводятся условия положительной определённости квадратичного функционала (11). Ниже  $a \in C^1(\Delta)$ ,  $a(t) > 0 \forall t \in \Delta$ ,  $b \in C(\Delta)$ . Существенную роль далее играет дифференциальное уравнение второго порядка

$$-(a(t)h')' + b(t)h = 0. \quad (19)$$

Его можно интерпретировать как уравнение Эйлера для функционала  $\mathcal{K}$ .

**Теорема 4. (Теорема Лежандра–Якоби).** Пусть  $u_1(t)$  есть решение уравнения (19), удовлетворяющее начальным условиям  $u_1(t_0) = 0$ ,  $u_1'(t_0) = 1$ . Если  $u_1(t) > 0 \forall t \in (t_0, t_1]$ , то найдётся такое  $\varepsilon_0 > 0$ , что

$$\mathcal{K}(h) \geq \varepsilon_0 \int_{\Delta} (h'^2(t) + h^2(t)) dt \quad \forall h \in C_0^1(\Delta). \quad (20)$$

◀ Первый этап. Докажем существование положительного на отрезке  $\Delta$  решения  $u(t)$  уравнения (1). Пусть  $u_2(t)$  – решение уравнения (19), удовлетворяющее начальным условиям  $u_2(t_0) = 1$ ,  $u_2'(t_0) = 0$ . Тогда функция  $u(t) = u_1(t) + \delta u_2(t)$  при малых  $\delta > 0$  является искомой.

Второй этап. Функция  $w = -au'/u$  удовлетворяет равенству

$$a(t)(b(t) + w') = w^2(t).$$

Для доказательства достаточно учесть, что функция  $u$  удовлетворяет дифференциальному уравнению (19).

Третий этап. Для любой функции  $h$  из  $C_0^1(\Delta)$  справедливо равенство

$$\int_{\Delta} (w(t)h'^2(t))'(t) dt = w(t)h^2(t)|_{t_0}^{t_1} = 0.$$

Следовательно, имеем

$$\mathcal{K}(h) = \int_{\Delta} [a(t)h'^2(t) + 2w(t)h(t)h'(t) + (b(t) + w'(t))h^2(t)] dt.$$

Поскольку  $a(t) > 0$ ,  $a(t)(b(t) + w') = w^2(t)$ , то

$$a(t)h'^2(t) + 2w(t)h(t)h'(t) + (b(t) + w')h^2(t) \geq 0,$$

поэтому  $\mathcal{K}(h) \geq 0 \forall h \in C_0^1(\Delta)$ .

Четвёртый этап. Заменим  $a(t), b(t)$  функциями  $a(t) - \varepsilon, b(t) - \varepsilon$ . При малых положительных  $\varepsilon$  функции  $a(t) - \varepsilon, b(t) - \varepsilon$  удовлетворяют тем же требованиям, поэтому соответствующий функционал

$$\mathcal{K}_{\varepsilon}(h) := \int_{\Delta} [(a(t) - \varepsilon)h'^2(t) + (b(t) - \varepsilon)h^2(t)] dt$$

неотрицателен:  $\mathcal{K}_\varepsilon(h) \geq 0 \forall h \in C_0^1(\Delta)$ . Это эквивалентно неравенству (20). Теорема 4 полностью доказана. ►

**Замечание 1.** Справедливо обратное к теореме 4 утверждение: если имеет место оценка (20) с  $\varepsilon_0 > 0$ , то  $u_1(t) > 0 \forall t \in (t_0, t_1]$ .

**Замечание 2.** Оценка (20) с  $\varepsilon = 0$  эквивалентна неравенству  $u_1(t) > 0, \forall t \in (t_0, t_1)$ .

Доказательства приведённых в замечаниях 1, 2 утверждений можно найти в [5] - [10], [21].

**Следствие.** Пусть  $\Delta = [0, T], a(t) \equiv 1, b(t) \equiv -1$ . Тогда положительная определённость квадратичного функционала  $\mathcal{K}$  на пространстве  $C_0^1(\Delta)$  эквивалентна оценке  $0 < T < \pi$ ; при  $T = \pi$  данный функционал неотрицателен; при  $T > \pi$  рассматриваемый функционал знакопеременен.

**6. Необходимые условия локального минимума второго порядка для простейшей вариационной задачи.** Вернёмся к задаче (1). В этом и следующем пунктах предполагается, что функция  $L : \Delta \times W \rightarrow \mathbb{R}$  и её частные производные  $L'_x, L'_u, L''_{xx}, L''_{xu}, L''_{uu}$  определены и непрерывны по совокупности переменных. Пусть  $\hat{x}(t)$  – локальное решение задачи (1). Положим

$$\begin{aligned}\hat{L}_{00}(t) &= L''_{xx}(t, \hat{x}(t), \hat{x}'(t)), & \hat{L}_{01}(t) &= L''_{xu}(t, \hat{x}(t), \hat{x}'(t)), \\ \hat{L}_{11}(t) &= L''_{uu}(t, \hat{x}(t), \hat{x}'(t)).\end{aligned}$$

Введём в рассмотрение интегральный квадратичный функционал

$$\hat{\mathcal{K}}(h) := \int_{\Delta} \left[ \hat{L}_{00} h^2(t) + 2\hat{L}_{01}(t) h(t) h'(t) + \hat{L}_{11}(t) h'^2(t) \right] dt.$$

**Теорема 5.** Если  $\hat{x}$  – локальное решение задачи (1), то  $\hat{\mathcal{K}}$  – неотрицательный квадратичный функционал на  $C_0^1(\Delta)$ , т.е.  $\hat{\mathcal{K}}(h) \geq 0 \forall h \in C_0^1(\Delta)$ .

► Пусть  $h \in C_0^1(\Delta), I(\xi) = f(\hat{x} + \xi h)$ . Тогда 0 есть точка локального минимума функции  $I$ , поэтому  $I'(0) = 0, I''(0) \geq 0$ . Остаётся заметить, что  $I''(0) = \hat{\mathcal{K}}(h)$ . ►

Полезно заметить, что если  $\hat{L}_{01}(t)$  – непрерывно дифференцируемая функция, то

$$\int_{\Delta} 2\hat{L}_{01}(t) h(t) h'(t) dt = \int_{\Delta} \hat{L}_{01}(t) dh^2(t) = - \int_{\Delta} h^2(t) (\hat{L}_{01}(t))' dt,$$

поэтому

$$\hat{\mathcal{K}}(h) = \int_{\Delta} \left( a(t) h'^2(t) + b(t) h^2(t) \right) dt,$$

где  $a(t) = \hat{L}_{11}(t), b(t) = \hat{L}_{00}(t) - (\hat{L}_{01}(t))'$ , т.е. функционал  $\hat{\mathcal{K}}$  приводится к стандартному виду. Из теоремы 5 и леммы 5 вытекает

**Следствие.** Если  $\hat{x}$  – локальное решение задачи (1), то

$$\hat{L}_{11}(t) = L''_{uu}(t, \hat{x}(t), \hat{x}'(t)) \geq 0 -$$

необходимое условие Лежандра локального минимума для простейшей вариационной задачи.

**7. Достаточные условия локального минимума второго порядка для простейшей вариационной задачи.** Пусть  $\hat{x}$  – экстремаль функционала  $f$ ,  $\hat{x} \in C^3(\Delta) \cap X$ , функция  $L$  трижды непрерывно дифференцируема по своим аргументам,  $h \in C_0^1(\Delta)$ ,  $\|h\| \leq \rho$ . В силу формулы Тейлора справедливо равенство

$$\begin{aligned} L(t, \hat{x}(t) + h(t), \hat{x}'(t) + h'(t)) - L(t, \hat{x}(t), \hat{x}'(t)) = \\ = \hat{L}_x(t)h(t) + \hat{L}_u(t)h'(t) + \\ + \frac{1}{2} \left( \hat{L}_{00}(t)h^2(t) + 2\hat{L}_{01}(t)h(t)h'(t) + \hat{L}_{11}(t)h'(t)^2(t) \right) + \mathcal{R}, \end{aligned} \quad (21)$$

где  $\mathcal{R} = o(\rho^2)$ . Из равенства (21) путём интегрирования получаем соотношение

$$f(\hat{x} + h) - f(\hat{x}) = \frac{1}{2} \hat{\mathcal{K}}(h) + \omega(h), \quad (22)$$

в котором

$$|\omega(h)| \leq \varepsilon(\rho) \int_{\Delta} \left( h'^2(t) + h^2(t) \right) dt, \quad \text{причём } \lim_{\rho \rightarrow 0} \varepsilon(\rho) = 0. \quad (23)$$

**Теорема 6.** Пусть функционал  $\hat{\mathcal{K}}$  положительно определён, т.е. справедлива оценка (20) с  $\hat{\mathcal{K}} = \mathcal{K}$ ,  $\varepsilon_0 > 0$ . Тогда  $\hat{x}$  – локальное решение задачи (1).

◀ Фиксируем число  $\rho_0$  так, что  $\varepsilon(\rho) < \varepsilon_0/4$  при  $\rho < \rho_0$ . Объединяя (22), (23), получаем, что при  $\|h\|_1 < \rho < \rho_0$  имеют место соотношения

$$\begin{aligned} f(\hat{x} + h) - f(\hat{x}) &= \frac{1}{2} \hat{\mathcal{K}}(h) + \omega(h) \geq \\ &\geq \left( \frac{\varepsilon_0}{2} - \varepsilon(\rho) \right) \int_{\Delta} \left( h'^2(t) + h^2(t) \right) dt \geq \varepsilon_0/4 \int_{\Delta} \left( h'^2(t) + h^2(t) \right) dt. \end{aligned}$$

Следовательно,  $f(\hat{x} + h) \geq f(\hat{x}) \forall h \in C_0^1(\Delta)$ ,  $\|h\| < \rho_0$ . ▶

Из теорем 6, 4 вытекает

**Следствие.** Пусть  $u_1(t)$  есть решение дифференциального уравнения

$$-(\hat{L}_{11}(t)h'(t))' + (\hat{L}_{00}(t) - (\hat{L}_{01}(t))')h(t) = 0,$$

удовлетворяющее начальным условиям  $h(t_0) = 0$ ,  $h'(t_0) = 1$ . Пусть

$$\hat{L}_{11}(t) > 0 \quad \forall t \in \Delta, \quad u_1(t) > 0 \quad \forall t \in (t_0, t_1].$$

Тогда  $\hat{x}(t)$  есть локальное решение задачи (1).

Сформулируем теперь систему условий, достаточных для того, чтобы функция  $\hat{x}$  из  $X$  была локальным решением задачи (1). Для трижды непрерывно дифференцируемого по совокупности переменных интегранта  $L(t, x, u)$  ( $t \in \Delta$ ,  $(x, u) \in W \subset \mathbb{R}^2$ ) эта совокупность условий состоит в следующем.

1. Функция  $x = \hat{x}(t)$  является экстремалью, т.е. удовлетворяет уравнению Эйлера

$$\frac{d}{dt}L'_u = L'_x.$$

2. Выполнено усиленное условие Лежандра

$$L''_{uu}(t, \hat{x}(t), \hat{x}'(t)) > 0 \quad \forall t \in \Delta.$$

3. Фигурирующая в следствии теоремы 6 функция  $u_1(t)$  положительна на промежутке  $(t_0, t_1]$  (усиленное условие Якоби).

Наиболее сложным для проверки представляется условие 3. Остановимся на аналитическом способе его проверки. Обозначим через  $x(t, \alpha)$  решение задачи Коши

$$\frac{d}{dt}L'_u = L'_x, \quad x(t_0) = x_0, \quad x'(t_0) = \alpha.$$

Из теории дифференциальных уравнений (см., например, [5]) известно, что функция  $x(t, \alpha)$  непрерывно дифференцируема по параметру  $\alpha$ ; справедливо равенство

$$u_1(t) = \left. \frac{\partial x(t, \alpha)}{\partial \alpha} \right|_{\alpha=\hat{x}'(t_0)}. \quad (24)$$

Формула (24) позволяет найти функцию  $u_1(t)$ , а вместе с тем доставляет способ проверки усиленного условия Якоби. Реализация этого подхода требует знания функции  $x(t, \alpha)$ .

Геометрически условие Якоби означает, что достаточно близкие к  $\hat{x}(t)$  экстремали  $x(t)$  функционала  $f$ , выходящие из точки  $x_0$ , не имеют общих точек с  $\hat{x}(t)$ , т.е.  $x(t_0) = \hat{x}(t_0) = x_0$ , но  $x(t) \neq \hat{x}(t)$  для  $t$  из  $(t_0, t_1]$ . Сформулированные выше достаточные условия локального минимума близки к необходимым. Более подробное изложение данного круга вопросов можно найти в [1], [6]-[11], [14], [15], [21].

Предшествующие рассуждения приводят к пессимистическому выводу: простейшая вариационная задача совсем не проста! Действительно, развитый выше подход позволяет лишь свести вариационную задачу к не менее сложной двухточечной краевой задаче для дифференциального уравнения второго порядка. В ряде случаев соответствующая краевая задача может быть решена; не слишком обширная коллекция примеров такого рода анализируется в разделе 3. Её можно пополнить, обратясь к руководствам по вариационному исчислению (см., например, [5]-[11], [14], [15], [21], [23], [28]-[30]).

Вместе с тем отмеченная выше связь между вариационными и краевыми задачами оказалась полезной для дифференциальных уравнений. Она повлияла и

на постановку многих проблем дифференциальных уравнений, и на разработку методов их решений.

## 1.2 Модификации простейшей вариационной задачи

**1. Постановка изопериметрической задачи.** В геометрии изопериметрической задачей называют следующую: среди замкнутых плоских кривых данной длины найти линию, ограничивающую наибольшую площадь. Решением этой задачи является окружность. Оно было известно в глубокой древности, однако аккуратное обоснование, удовлетворяющее современным требованиям строгости, проведено лишь в 19-м веке. Примыкает к изопериметрической и знаменитая задача Дидоны: среди кривых, опирающихся на данный отрезок и имеющих заданную длину, найти линию, ограничивающую вместе с отрезком фигуру наибольшей площади.

Изопериметрической в вариационном исчислении называют задачу, записываемую следующим образом

$$f_0(x) = \int_{\Delta} L_0(t, x(t), x'(t)) dt \rightarrow \min, \quad (1)$$

$$f_i(x) = \int_{\Delta} L_i(t, x(t), x'(t)) dt = C_i \quad (i = 1, \dots, m). \quad (2)$$

$$x(t_0) = x_0, \quad x(t_1) = x_1. \quad (3)$$

В этом и следующем пунктах  $\Delta = [t_0, t_1]$ ,  $-\infty < t_0 < t_1 < \infty$ , функции  $L_i(t, x, u)$  ( $i = 0, 1, \dots, m$ ) — определены и непрерывны вместе с частными производными по  $x$  и  $u$  на  $\Delta \times W$ ; здесь  $W$  — открытое подмножество плоскости  $\mathbb{R}^2$ . Положим

$$X := \{x \in C^1(\Delta), \quad x(t_0) = x_0, \quad x(t_1) = x_1, \quad (x(t), x'(t)) \in W \forall t \in \Delta\}.$$

С учётом принятых обозначений задача (1)-(3) может быть записана в стандартной форме

$$f_0(x) \rightarrow \min, \quad f_i(x) - C_i = 0 \quad (i = 1, \dots, m), \quad x \in X. \quad (4)$$

Естественным образом определяются понятия решения и локального решения задачи (1)-(3). Например,  $\hat{x}$  есть локальное решение задачи (1)-(3), если  $\hat{x} \in X$ ,  $f_i(\hat{x}) = C_i$  ( $i = 1, \dots, m$ ) и существует такое  $\delta > 0$ , что для всех  $x$ , удовлетворяющих неравенству  $\|x - \hat{x}\|_1 < \delta$  и ограничениям  $x \in X$ ,  $f_i(x) = C_i$  ( $i = 1, \dots, m$ ), имеет место неравенство  $f_0(x) \geq f_0(\hat{x})$ .

## 2. Правило множителей Лагранжа для изопериметрической задачи.

**Теорема 1.** Пусть  $\hat{x}$  – локальное решение задачи (1). Тогда существует такой ненулевой набор чисел  $\hat{\lambda}_0, \hat{\lambda}_1, \dots, \hat{\lambda}_m$ , что  $\hat{x}$  есть экстремаль функционала  $f = \hat{\lambda}_0 f_0 + \hat{\lambda}_1 f_1 + \dots + \hat{\lambda}_m f_m(x)$ .

◀ Рассмотрим случай  $m = 1$ . Если  $\hat{x}$  – экстремаль функционала  $f_1$ , то всё доказано; достаточно положить  $\hat{\lambda}_0 = 0, \hat{\lambda}_1 = 1$ .

Пусть  $\hat{x}$  не является экстремалью функционала  $f_1$ . Тогда существует функция  $h_1$  из  $C_0^1(\Delta)$ , для которой

$$f'_1(\hat{x}, h_1) = \int_{\Delta} [\hat{L}_{1x}(t)h_1(t) + \hat{L}_{1u}(t)h'_1(t)] dt \neq 0 \quad (5)$$

(здесь

$$\hat{L}_{1x}(t) = \frac{\partial L_1}{\partial x}(t, \hat{x}(t), \hat{x}'(t)), \quad \hat{L}_{1u}(t) = \frac{\partial L_1}{\partial u}(t, \hat{x}(t), \hat{x}'(t)).$$

Фиксируем произвольную функцию  $h_2$  из  $C_0^1(\Delta)$  и введём в рассмотрение две числовые функции

$$\varphi_0(\xi_1, \xi_2) = f_0(\hat{x} + \xi_1 h_1 + \xi_2 h_2),$$

$$\varphi_1(\xi_1, \xi_2) = f_1(\hat{x} + \xi_1 h_1 + \xi_2 h_2) - C_1.$$

Поскольку  $\hat{x}$  – локальное решение задачи (1)-(3), то точка  $0=(0,0)$  из  $\mathbb{R}^2$  есть локальное решение задачи

$$\varphi_0(\xi_1, \xi_2) \rightarrow \min, \quad \varphi_1(\xi_1, \xi_2) = 0. \quad (6)$$

К задаче (6) можно применить правило множителей Лагранжа. Действительно, функции  $\varphi_0, \varphi_1$  непрерывно дифференцируемы; справедливы равенства

$$\frac{\partial \varphi_i}{\partial \xi_j}(0,0) = f'_i(\hat{x}, h_j) \quad (i = 0, 1; j = 1, 2). \quad (7)$$

Из (5),(7) вытекает соотношение  $\frac{\partial \varphi_1}{\partial \xi_1}(0,0) \neq 0$ , поэтому для задачи (6) выполнены условия регулярности. Существует такая константа  $\lambda$ , что точка  $(0,0)$  является критической для функции  $\varphi(\xi_1, \xi_2) = \varphi_0(\xi_1, \xi_2) + \lambda \varphi_1(\xi_1, \xi_2)$ , т.е. справедливы соотношения

$$\frac{\partial \varphi_0}{\partial \xi_1}(0,0) + \lambda \frac{\partial \varphi_1}{\partial \xi_1}(0,0) = 0, \quad \frac{\partial \varphi_0}{\partial \xi_2}(0,0) + \lambda \frac{\partial \varphi_1}{\partial \xi_2}(0,0) = 0.$$

Первое равенство однозначно определяет  $\lambda$ , второе соотношение означает, что для любого выбора элемента  $h_2$  из  $C_0^1(\Delta)$  верно равенство  $(f_0 + \lambda f_1)'(\hat{x}, h_2) = 0$ . Это влечёт за собой доказываемое утверждение в случае  $m = 1$ . Общий случай ( $m \geq 1$ ) рассматривается аналогично. ▶

**3. Простейшая вариационная задача в случае вектор-функций.** Рассмотрим обобщение простейшей вариационной задачи в плане замены скалярных функций векторными. Пусть  $n$  – натуральное число,  $\mathbb{R}^n$  – евклидово пространство размерности  $n$ . Обозначим через  $C^1(\Delta, \mathbb{R}^n)$  совокупность непрерывно дифференцируемых отображений отрезка  $\Delta = [t_0, t_1]$  в  $\mathbb{R}^n$ . Для  $x$  из  $C^1(\Delta, \mathbb{R}^n)$  положим

$$\|x\| = \max_{t \in \Delta} |x(t)| + \max_{t \in \Delta} |x'(t)|,$$

где  $|v|$  – длина вектора  $v$  из  $\mathbb{R}^n$ .

Векторное обобщение простейшей вариационной задачи записывается стандартным образом

$$f(x) = \int_{\Delta} L(t, x(t), x'(t)) dt \rightarrow \min, \quad x(t_0) = x_0, x(t_1) = x_1. \quad (8)$$

Здесь  $L(t, x, u) = L(t, x_1, \dots, x_n, u_1, \dots, u_n)$  – функция, непрерывная по совокупности переменных вместе с частными производными  $\frac{\partial L}{\partial x_i}, \frac{\partial L}{\partial u_i}$  ( $i = 1, \dots, n$ ) ( $t \in \Delta, (x, u) \in W \subset \mathbb{R}^{2n}$ ),  $W$  – открытое подмножество  $\mathbb{R}^{2n}$ ;  $x_0, x_1$  – фиксированные элементы  $\mathbb{R}^n$ . Естественным образом вводятся понятия решения и локального решения задачи (8).

Все результаты предшествующего параграфа переносятся на задачу (8). В частности, локальное решение  $\hat{x} = (\hat{x}_1 \dots \hat{x}_n)^T$  задачи (8) удовлетворяет системе дифференциальных уравнений

$$\frac{d}{dt} L'_{u_i}(t, x(t), x'(t)) = L_{x_i}(t, x(t), x'(t)) \quad (i = 1, \dots, n), \quad (9)$$

называемой системой уравнений Эйлера–Лагранжа. Для вывода (9) достаточно в задаче (8) фиксировать функции  $\hat{x}_1(t), \dots, \hat{x}_{i-1}(t), \hat{x}_{i+1}(t), \dots, \hat{x}_n(t)$ , а менять лишь  $i$ -ую компоненту  $x_i(t)$  вектор-функции  $x(t)$ . В итоге относительно  $x_i(t)$  возникает простейшая вариационная задача, решение которой удовлетворяет уравнению Эйлера. Это и приводит к системе уравнений (9).

Решения системы (9) называют экстремалиями функционала  $f$ . Если функция  $L(t, x, u)$  при любом  $t$  из  $\Delta$  выпукла по совокупности переменных  $(x, u) \in W$  и  $W$  – выпуклое подмножество  $\mathbb{R}^{2n}$ , то необходимое условие локального минимума (уравнения Эйлера–Лагранжа) становится достаточным условием абсолютного минимума. Требование выпуклости интегранта можно ослабить. Проведённый в предшествующем параграфе анализ уравнения Эйлера может быть повторен и для векторного случая. Имеют место аналоги интегралов импульса и энергии. Например, если интегрант  $L$  не зависит от переменной  $u_k$  при некотором  $k$ , то справедливо равенство

$$L'_{u_k}(t, x(t), x'(t)) = C_1 - \text{интеграл импульса.}$$

Если функция  $L$  не зависит от переменного  $t$ , то

$$\sum_{i=1}^n x'_i(t) L'_{u_i}(x(t), x'(t)) - L(x(t), x'(t)) = C_1 - \text{интеграл энергии.}$$

**4. Вариационные задачи для функционалов, содержащих производные высших порядков.** Рассмотрим следующее обобщение простейшей вариационной задачи

$$f(x) = \int_{\Delta} L(t, x(t), x'(t), \dots, x^{(m)}(t)) dt \rightarrow \min, \quad (10)$$

$$x^{(j-1)}(t_0) = x_{0j}, \quad x^{(j-1)}(t_1) = x_{1j} \quad (j = 1, \dots, m). \quad (11)$$

Здесь  $\Delta = [t_0, t_1]$  – фиксированный отрезок действительной прямой,  $-\infty < t_0 < t_1 < \infty$ ,  $L(t, u_0, u_1, \dots, u_m)$  – функция, определённая на прямом произведении  $\Delta \times W$  отрезка  $\Delta$  и открытого подмножества  $W$  пространства  $\mathbb{R}^{m+1}$ ,  $x_{0j}$ ,  $x_{1j}$  ( $j = 1, \dots, m$ ) – фиксированные числа. Будем считать, что функция  $L$  и её частные производные  $L'_{u_i}$  ( $i = 0, 1, \dots, m$ ) определены и непрерывны по совокупности переменных на  $\Delta \times W$ . Положим

$$X := \{x \in C^m(\Delta), x^{(j-1)}(t_0) = x_{0j}, x^{(j-1)}(t_1) = x_{1j} \quad (j = 1, \dots, m),$$

$$(x(t), x'(t), \dots, x^{(m)}(t)) \in W \quad \forall t \in \Delta\},$$

$$C_0^m(\Delta) := \{x \in C^m(\Delta), x^{(j-1)}(t_0) = x^{(j-1)}(t_1) = 0, \quad j = 1, \dots, m\}.$$

Задача (10), (11) может быть записана в стандартной форме

$$f(x) \rightarrow \min, \quad x \in X. \quad (12)$$

Это позволяет определить понятия глобального и локального решения задачи (10), (11). Например,  $\hat{x}$  – локальное решение задачи (10), (11), если найдётся такое  $\delta > 0$ , что

$$f(x) \geq f(\hat{x}) \quad \forall x \in X, \quad \|x - \hat{x}\|_m < \delta.$$

Основное необходимое условие локального минимума в рассматриваемой задаче – дифференциальное уравнение Эйлера–Пуассона порядка  $2m$ . Для его вывода нам потребуется несколько вспомогательных утверждений.

**Лемма 1 (Лемма Лагранжа).** Пусть  $a(t)$  – непрерывная на отрезке  $\Delta$  функция. Если для любой функции  $h$  класса  $C_0^m(\Delta)$  справедливо равенство

$$\int_{\Delta} a(t) h(t) dt = 0, \quad (13)$$

то  $a(t) \equiv 0$ .



◀ В предположении противного найдётся точка  $\theta$  из интервала  $(t_0, t_1)$ , для которой  $a(\theta) \neq 0$ . Пусть для определённости  $a(\theta) = 2\alpha > 0$ . В силу непрерывности функции  $a(t)$  найдётся такое  $\varepsilon > 0$ , что справедливо неравенство  $a(t) > \alpha$  для всех чисел  $t$  из  $[\theta - \varepsilon, \theta + \varepsilon] \subset (t_0, t_1)$ . Введём в рассмотрение функцию

$$h(t) = \begin{cases} (t - \theta - \varepsilon)^{m+1}(t - \theta + \varepsilon)^{m+1}, & \text{если } t \in [\theta - \varepsilon, \theta + \varepsilon] \\ 0, & \text{если } t \notin [\theta - \varepsilon, \theta + \varepsilon]. \end{cases}$$

Функция  $h$  принадлежит классу  $C_0^m(\Delta)$ , поэтому для неё выполняется равенство (13). Вместе с тем

$$\begin{aligned} \int_{\Delta} a(t)h(t) dt &= \int_{\theta-\varepsilon}^{\theta+\varepsilon} a(t)(t - \theta - \varepsilon)^{m+1}(t - \theta + \varepsilon)^{m+1} dt > \\ &> \alpha \int_{\theta-\varepsilon}^{\theta+\varepsilon} (t - \theta - \varepsilon)^{m+1}(t - \theta + \varepsilon)^{m+1} dt > 0. \end{aligned}$$

Полученное противоречие доказывает лемму. ▶

**Лемма 2.** Пусть  $a_k \in C^k(\Delta)$  ( $k = 0, 1, \dots, m$ ) и для любой функции  $h$  класса  $C_0^m(\Delta)$  выполнено равенство

$$\int_{\Delta} \sum_{k=0}^m a_k(t)h^{(k)}(t) dt = 0.$$

Тогда

$$\sum_{k=0}^m (-1)^k \frac{d^k}{dt^k} a_k(t) = 0.$$

◀ Пусть  $1 \leq k \leq m$ ,  $h \in C_0^m(\Delta)$ . Используя интегрирование по частям и включение  $h \in C_0^m(\Delta)$ , получаем последовательно

$$\int_{\Delta} a_k(t)h^{(k)}(t) dt = - \int_{\Delta} (a_k(t))'(t)h^{(k-1)}(t) dt = \dots = (-1)^k \int_{\Delta} (a_k(t))^{(k)}h(t) dt.$$

Суммируя подобные равенства, приходим к соотношению

$$\int_{\Delta} \sum_{k=0}^m a_k(t)h^{(k)}(t) dt = \int_{\Delta} \sum_{k=0}^m (-1)^k (a_k(t))^{(k)}h(t) dt,$$

верному для произвольной функции  $h$  из  $C_0^m(\Delta)$ . Последнее равенство согласно лемме Лагранжа влечёт за собой доказываемое утверждение. ▶

**Лемма 3.** Пусть  $\hat{x}$  – локальное решение задачи (10), (11),

$$\hat{a}_k(t) := \frac{\partial L}{\partial u_k}(t, \hat{x}(t), \hat{x}'(t), \dots, \hat{x}^{(m)}(t)) \quad (k = 0, 1, \dots, m).$$

Тогда для любой функции  $h$  из  $C_0^m(\Delta)$  справедливо равенство

$$\int_{\Delta} \sum_{k=0}^m \hat{a}_k(t) h^{(k)}(t) dt = 0,$$

называемое уравнением Эйлера–Пуассона в интегральной форме.

◀ Пусть  $h \in C_0^m(\Delta)$ . Введём в рассмотрение функцию  $I(\xi) := f(\hat{x} + \xi h)$ . Функция  $I(\xi)$  определена и дифференцируема на некоторой окрестности 0. Так как  $\hat{x}$  – локальное решение задачи (10), (11), то  $I(\xi)$  достигает локального минимума при  $\xi = 0$ . В силу теоремы Ферма  $I'(0) = 0$ . Для завершения доказательства достаточно заметить, что

$$I'(0) = \int_{\Delta} \sum_{k=0}^m \hat{a}_k(t) h^{(k)}(t) dt.$$

Последнее равенство вытекает из правила Лейбница дифференцирования интегралов, зависящих от параметра. ▶

**Теорема 2.** Если в условиях леммы 3  $\hat{x}$  принадлежит  $C^{2m}(\Delta)$ , а функции  $L'_{u_k}(t, u_0, u_1, \dots, u_m)$   $k$  раз непрерывно дифференцируемы по совокупности переменных, то  $\hat{a}_k \in C^k(\Delta)$  ( $k = 0, 1, \dots, m$ ) и функция  $\hat{x}$  удовлетворяет дифференциальному уравнению

$$\sum_{k=0}^m (-1)^k \frac{d^k}{dt^k} L'_{u_k}(t, x(t), x'(t), \dots, x^{(m)}(t)) = 0, \quad (14)$$

называемому уравнением Эйлера–Пуассона в дифференциальной форме.

◀ Теорема вытекает из лемм 2, 3. ▶

Поскольку порядок дифференциального уравнения (14) равен  $2m$ , его общее решение  $x$  зависит от  $2m$  произвольных постоянных:  $x = x(t, C_1, C_2, \dots, C_{2m})$ . Для нахождения этих постоянных можно воспользоваться граничными условиями (11). В качестве примера рассмотрим вариационную задачу

$$f(x) = \int_0^1 x''^2(t) dt \rightarrow \min,$$

$$x(0) = x'(0) = x'(1) = 0, x(1) = 1.$$

В рассматриваемом случае  $L(t, u_0, u_1, u_2) = u_2^2$ , поэтому  $L'_{u_0} = L'_{u_1} = 0$ ,  $L'_{u_2} = 2u_2$ . Уравнение (14) имеет вид  $2x^{(4)} = 0$ . Его общее решение

$$x(t) = C_0 + C_1 t + C_2 t^2 + C_3 t^3$$

– произвольный многочлен третьей степени. Постоянные  $C_i$  находим из граничных условий:

$$x(0) = C_0 = 0, x'(0) = C_1 = 0, x'(1) = 2C_2 + 3C_3 = 0, x(1) = C_2 + C_3 = 1.$$

Отсюда получаем  $C_0 = C_1 = 0, C_2 = 3, C_3 = -2$ . Искомая экстремаль  $x(t) = 3t^2 - 2t^3$ . В рассматриваемом примере функционал  $f$  выпукл, поэтому необходимое условие локального минимума является достаточным условием абсолютного минимума.

Возвращаясь к общему случаю задачи (10), (11), заметим, что для неё справедливы аналоги результатов, установленных для простейшей вариационной задачи. Вопрос о существовании и гладкости обобщённых решений задачи (10), (11) обсуждается в дополнении В.

**5. Вариационные задачи на условный экстремум.** Рассматривается задача о минимизации интегрального функционала

$$f(x) = \int_{\Delta} L(t, x(t), x'(t)) dt,$$

определённого на вектор-функциях  $x$  класса  $C^1(\Delta, \mathbb{R}^n)$ , удовлетворяющих краевым условиям

$$x(t_0) = x_0, \quad x(t_1) = x_1 \quad (15)$$

и условиям связи вида

$$\varphi_i(t, x(t)) = 0, \quad i = 1, \dots, m, \quad m < n. \quad (16)$$

Предполагается, что функции  $L : \Delta \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  и  $\varphi_i : \Delta \times \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $i = 1, \dots, m$  дважды непрерывно дифференцируемы по всем переменным; уравнения связи (16) независимы, т.е.

$$\text{rang} \begin{pmatrix} \frac{\partial \varphi_1}{\partial x_1} & \dots & \frac{\partial \varphi_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \varphi_m}{\partial x_1} & \dots & \frac{\partial \varphi_m}{\partial x_n} \end{pmatrix} = m;$$

связи (16) согласованы с краевыми условиями (15):

$$\varphi_i(t_0, x_0) = 0, \quad \varphi_i(t_1, x_1) = 0 \quad (i = 1, \dots, m).$$

Поставленная задача относится к задачам на условный экстремум, поскольку, кроме краевых условий (15), на искомую вектор-функцию наложены дополнительные условия связи (16). Краткости ради исследуемую задачу будем обозначать символом (3) и называть задачей на условный экстремум. Нас будут интересовать необходимые условия локального минимума в этой задаче.

**Теорема 3 (необходимые условия локального минимума в задаче (3)).** Пусть  $\hat{x}(t)$  – локальное решение задачи (3) класса  $C^2(\Delta, \mathbb{R}^n)$ . Тогда существуют такие непрерывные на отрезке  $\Delta$  функции  $\lambda_1(t), \dots, \lambda_m(t)$ , что  $\hat{x}(t)$  удовлетворяет системе дифференциальных уравнений

$$L'_{x_j}(t, x(t), x'(t)) - \frac{d}{dt} L'_{u_j}(t, x(t), x'(t)) + \sum_{i=1}^m \lambda_i(t) \frac{\partial g_i}{\partial x_j}(t, x(t)) = 0. \quad (17)$$

Доказательство теоремы 3 приводится в следующем пункте. Если ввести обозначение

$$\bar{L}(t, x, u, \lambda(t)) = L(t, x, u) + \sum_{i=1}^m \lambda_i(t) g_i(t, x),$$

где  $\bar{L}(t, x, u, \lambda(t))$  называется функцией Лагранжа, а функции  $\lambda_i$  ( $i = 1, \dots, m$ ) – множителями Лагранжа,  $\lambda(t) = (\lambda_1(t) \dots \lambda_m(t))^T$ , то система (17) совпадает с системой уравнений Эйлера–Лагранжа

$$\bar{L}'_{x_j} = \frac{d}{dt} \bar{L}'_{u_j}, \quad (j = 1, \dots, n)$$

составленной для функционала

$$\bar{f}(x) = \int_{\Delta} \bar{L}(t, x(t), x'(t), \lambda(t)) dt.$$

Теорема 3 сводит решение задачи (3) к исследованию экстремалей функционала  $\bar{f}$  при отсутствии уравнений связи. Приведённый способ аналогичен методу множителей Лагранжа для решения задачи математического программирования. В рассматриваемом случае множители Лагранжа  $\lambda_i$  уже не скаляры, а функции переменного  $t \in \Delta$ .

**6. Доказательство теоремы 3** проведём в случае  $m = 1$ . Оно опирается на ряд вспомогательных утверждений. Пусть  $b(t)$  – гладкая на отрезке  $\Delta$  вектор-функция со значениями в пространстве  $\mathbb{R}^n$  (кратко  $b \in C^1(\Delta, \mathbb{R}^n)$ ) и  $b(t) \neq 0 \forall t \in \Delta$ . Введём в рассмотрение множество

$$C_0^1(\Delta, b^\perp) := \{h \in C_0^1(\Delta, \mathbb{R}^n), \quad (h(t), b(t)) = 0 \forall t \in \Delta\}.$$

**Лемма 4.** Пусть  $a \in C(\Delta, \mathbb{R}^n)$ . Если

$$\int_{\Delta} (a(t), h(t)) dt = 0 \quad \forall h \in C_0^1(\Delta, b^\perp),$$

то  $a(t) = \mu(t)b(t)$ , где  $\mu : \Delta \rightarrow \mathbb{R}$  – непрерывная на отрезке  $\Delta$  функция.

◀ Пусть  $h \in C_0^1(\Delta, b^\perp)$ ,  $\varphi \in C^1(\Delta)$ . Тогда  $\varphi h \in C_0^1(\Delta, b^\perp)$  и в силу условий леммы справедливо равенство

$$\int_{\Delta} \varphi(t)(a(t), h(t)) dt = 0.$$

Так как  $\varphi$  – произвольная функция класса  $C^1(\Delta)$ , то согласно лемме Лагранжа  $(a(t), h(t)) = 0$  для каждой функции  $h$  из  $C_0^1(\Delta, b^\perp)$ .

Фиксируем число  $\tau$  из интервала  $(t_0, t_1)$  и вектор  $v$  из  $\mathbb{R}^n$ , для которого  $(v, b(\tau)) = 0$ . Построим такую вектор-функцию  $h$  из  $C_0^1(0, b^\perp)$ , что  $h(\tau) = v$ ; способ построения предлагается придумать читателю в качестве самостоятельного упражнения. В силу уже доказанного  $(a(t), h(t)) = 0$ , в частности,

$$(a(\tau), h(\tau)) = (a(\tau), v) = 0.$$

Таким образом,  $(a(\tau), v) = 0$  для каждого вектора  $v$ , ортогонального вектору  $b(\tau)$ . Отсюда следует равенство  $a(\tau) = \mu(\tau)b(\tau)$ . Определяемая таким образом функция  $\mu(\tau)$  непрерывна на интервале  $(t_0, t_1)$ . По непрерывности равенство  $a(t) = \mu(t)b(t)$  продолжается на весь отрезок  $\Delta$ . ►

Положим

$$b_i(t) = \frac{\partial \varphi_1}{\partial x_i}(t, \hat{x}(t)) \quad (i = 1, \dots, n)$$

и введём в рассмотрение вектор-функцию  $b(t) = (b_1(t) \dots b_n(t))^T$ . В силу принятых выше предположений  $b(t)$  – гладкая вектор-функция на отрезке  $\Delta$  и  $b(t) \neq 0 \forall t \in \Delta$ . Класс  $C_0^1(\Delta, b^\perp)$ , определяемый этой вектор-функцией, заведомо непуст.

**Лемма 5.** Пусть  $h \in C_0^1(\Delta, b^\perp)$ . Тогда существует такая функция  $\eta = \eta(t, \xi)$  ( $t \in \Delta$ ,  $|\xi| < \delta$ ,  $\delta$  – некоторое положительное число), что

$$\varphi_1(t, \hat{x}(t) + \xi h(t) + \eta b(t)) = 0; \quad (18)$$

функции  $\eta(t, \xi), \eta'_\xi(t, \xi)$  непрерывны по совокупности переменных на  $\Delta \times (-\delta, \delta)$  и

$$\frac{\partial \eta}{\partial \xi}(t, 0) = 0. \quad (19)$$

◀ Положим

$$\Phi(t, \xi, \eta) = \varphi_1(t, \hat{x}(t) + \xi h(t) + \eta b(t)).$$

Тогда  $\Phi(t, 0, 0) = 0$ ; функция  $\Phi$  непрерывно дифференцируема по переменным  $t, \eta, \xi$  при  $t \in \Delta$ ,  $(\xi, \eta) \in U$ ,  $U$  – некоторая окрестность точки  $(0, 0)$ ;

$$\Phi'_\xi(t, 0, 0) = 0, \Phi'_\eta(t, 0, 0) \neq 0. \quad (20)$$

Согласно теоремам о неявных функциях (см., например, [1]) существует функция  $\eta = \eta(t, \xi)$ , удовлетворяющая условиям

1)  $\Phi(t, \xi, \eta(t, \xi)) = 0$ ;

2) функции  $\eta(t, \xi)$ ,  $\eta'_\xi(t, \xi)$  непрерывны по совокупности переменных на  $\Delta \times (-\delta, \delta)$  при некотором  $\delta > 0$ , справедливо правило дифференцирования неявной функции  $\eta(t, \xi)$ , а именно

$$\frac{\partial \eta}{\partial \xi}(t, \xi) = -\frac{\Phi'_\xi(t, \xi, \eta)}{\Phi'_\eta(t, \xi, \eta)}. \quad (21)$$

Функция  $\eta(t, \xi)$  является искомой. Действительно, она и её частная производная  $\eta'_\xi(t, \xi)$  непрерывны по совокупности переменных на множестве  $\Delta \times (-\delta, \delta)$ ,  $\delta$  – некоторое положительное число; равенство (19) вытекает из соотношений (20), (21). ►

Введём в рассмотрение функции

$$a_{0i}(t) := L'_{x_i}(t, \hat{x}(t), \hat{x}'(t)), \quad a_{1i}(t) := L'_{u_i}(t, \hat{x}(t), \hat{x}'(t)) \quad (i = 1, \dots, n)$$

и вектор-функции

$$a_0(t) := (a_{01}(t), \dots, a_{0n}(t))^T, \quad a_1(t) := (a_{11}(t), \dots, a_{1n}(t))^T.$$

Поскольку  $\hat{x} \in C^2(\Delta, \mathbb{R}^n)$ , а функция  $L(t, x, u)$  дважды непрерывно дифференцируема, то вектор-функции  $a_0(t)$ ,  $a_1(t)$  непрерывно дифференцируемы на отрезке  $\Delta$ . В частности, для любой вектор-функции  $h$  класса  $C_0^1(\Delta, \mathbb{R}^n)$  имеет место равенство

$$\int_{\Delta} (a_1(t), h'(t)) dt = (a_1(t), h(t))|_{t_0}^{t_1} - \int_{\Delta} (a'_1(t), h(t)) dt = - \int_{\Delta} (a'_1(t), h(t)) dt.$$

Из этого равенства вытекает формула

$$\int_{\Delta} [(a_0(t), h(t)) + (a_1(t), h'(t))] dt = \int_{\Delta} [(a_0(t) - a'_1(t), h(t))] dt. \quad (22)$$

**Доказательство теоремы 3.** Пусть  $h \in C_0^1(\Delta, b^\perp)$ ,  $\eta(t, \xi)$  – функция, фигурирующая в лемме 5. Вектор-функция  $\hat{x}(t) + \xi h(t) + \eta(t, \xi)b(t)$  удовлетворяет ограничениям задачи (3). Поскольку  $\hat{x}$  – локальное решение этой задачи, то  $f(\hat{x} + \xi h + \eta(\cdot, \xi)b) \geq f(\hat{x})$  при достаточно малых  $\xi$ . Иначе говоря,  $\xi = 0$  есть точка локального минимума функции  $I(\xi) := f(\hat{x} + \xi h + \eta(\cdot, \xi)b)$ . В силу теоремы Ферма имеем  $I'(0) = 0$ . Распишем последнее равенство. Используя правило Лейбница дифференцирования интегралов, зависящих от параметра, и формулы (19), (22), приходим к соотношениям

$$I'(0) = \int_{\Delta} [(a_0(t), h(t)) + (a_1(t), h'(t))] dt = \int_{\Delta} ((a_0(t) - a'_1(t), h(t))) dt = 0.$$

В силу леммы 4 найдётся такая непрерывная на отрезке  $\Delta$  функция  $\mu(t)$ , что  $a_0(t) - a_1'(t) = \mu(t)b(t)$ . Отсюда и следует (17) с  $\lambda_1(t) = -\mu(t)$ .

В случае  $m = 1$  теорема доказана. При  $m > 1$  доказательство проводится по той же схеме; читатель может найти его в [8], с. 105-106. ►

Наряду с методом множителей Лагранжа для решения задачи (3) применяются метод исключения. Условия связи (16) используют для исключения  $m$  неизвестных координат вектор-функции  $x(t)$ . Рассмотрим, например, приведённую в [14] задачу о нахождении экстремума функционала

$$f(x) = \int_0^{\frac{\pi}{2}} \left[ x_1^2(t) + x_2^2(t) - x_1'^2(t) - x_2'^2(t) \right] dt$$

на классе вектор-функций  $x(t) = (x_1(t), x_2(t))^T$ , удовлетворяющих граничным условиям:

$$x_1(0) = 1, \quad x_2(0) = -1, \quad x_1\left(\frac{\pi}{2}\right) = 1, \quad x_2\left(\frac{\pi}{2}\right) = 1$$

и уравнению связи  $x_1 - x_2 - 2 \cos t = 0$ .

Вряд ли для решения этой задачи целесообразно применять правило множителей Лагранжа; впрочем, данный подход также реализуем (см. [14], с. 514-516). Проще всего воспользоваться вытекающим из уравнения связи равенством  $x_1(t) = x_2(t) + 2 \cos t$  и свести исследуемую проблему к задаче минимизации интегрального функционала, зависящей от одной неизвестной функции  $x_2(t)$ . При реализации намеченного подхода относительно неизвестной функции возникает уравнение Эйлера.

**7. Задача Больца.** В простейшей вариационной задаче

$$f(x) = \int_{\Delta} L(t, x(t), x'(t)) dt \rightarrow \min, \quad x(t_0) = x_0, \quad x(t_1) = x_1$$

числа  $x_0, x_1$  предполагаются фиксированными. Отказ от этого предположения приводит к задаче Больца

$$f(x) = \int_{\Delta} L(t, x(t), x'(t)) dt + g(x(t_0), x(t_1)) \rightarrow \min. \quad (23)$$

Минимизируемый функционал наряду с интегральным слагаемым содержит терминальную часть —  $g(x(t_0), x(t_1))$ . Задача (1) может быть записана в стандартной форме

$$f(x) \rightarrow \min, \quad x \in X := C^1(\Delta).$$

Это позволяет обычным образом ввести понятия решения и локального решения задачи (23).

Считаем, что функции  $L(t, x, u)$ ,  $L'_x(t, x, u)$ ,  $L'_u(t, x, u)$  определены и непрерывны по совокупности переменных на  $\Delta \times \mathbb{R}^2$ , а функция  $g(z_0, z_1)$  определена

и непрерывно дифференцируема на плоскости  $\mathbb{R}^2$ . В наших предположениях функционал  $f : X \rightarrow \mathbb{R}$  всюду дифференцируем, а если говорить более точно, при любых  $x$  и  $h$  из  $X$  существует производная  $f'(x, h)$  функционала  $f$  в точке  $x$  по направлению  $h$ :

$$\begin{aligned} f'(x, h) &= \lim_{\xi \rightarrow 0} \frac{f(x + \xi h) - f(x)}{\xi} = \frac{d}{d\xi} f(x + \xi h) \Big|_{\xi=0} = \\ &= \int_{\Delta} [L'_x(t, x(t), x'(t))h(t) + L'_u(t, x(t), x'(t))h'(t)] dt + \\ &\quad + g'_{z_0}(x(t_0), x(t_1))h(t_0) + g'_{z_1}(x(t_0), x(t_1))h(t_1). \end{aligned} \quad (24)$$

Формула (24) вытекает из правила Лейбница дифференцирования интегралов, зависящих от параметра.

Назовём элемент  $x$  из  $X$  критической точкой функционала  $f$ , если  $f'(x, h) = 0$  для любого направления  $h$  из  $X$ .

**Лемма 6.** *Элемент  $x$  из  $X$  есть критическая точка функционала  $f$  в том и только в том случае, если выполнены условия:*

1) *функция  $x(t)$  является решением уравнения Эйлера*

$$\frac{d}{dt} L'_u = L'_x;$$

2) *справедливы соотношения*

$$\left( \frac{\partial L}{\partial u}(t, x(t), x'(t)) \right) \Big|_{t=t_1} + g'_{z_1}(x(t_0), x(t_1)) = 0, \quad (25)$$

$$\left( \frac{\partial L}{\partial u}(t, x(t), x'(t)) \right) \Big|_{t=t_0} = g'_{z_0}(x(t_0), x(t_1)). \quad (26)$$

◀ Пусть  $x$  – критическая точка функционала  $f$ . Тогда  $f'(x, h) = 0 \forall h$  из  $X$ . В частности, данное равенство имеет место для всех  $h$  из  $C_0^1(\Delta)$ . Согласно формуле (24) отсюда вытекает равенство

$$\int_{\Delta} [L'_x(t, x(t), x'(t))h(t) + L'_u(t, x(t), x'(t))h'(t)] dt = 0 \forall h \in C_0^1(\Delta).$$

Теперь выполнение первого условия следует из леммы Дюбуа–Реймона.

Для проверки соотношений (25), (26) снова воспользуемся формулой (24), но уже для произвольной функции  $h$  из  $X$ . Интегрирование по частям влечёт равенство

$$\int_{\Delta} L'_u(t, x(t), x'(t))h'(t) dt = L'_u(t, x(t), x'(t))h(t) \Big|_{t_0}^{t_1} -$$



$$- \int_{\Delta} h(t) \frac{d}{dt} L'_u(t, x(t), x'(t)) dt,$$

объединяя которое с формулой (24) и уравнением Эйлера, получаем, что

$$\begin{aligned} f'(x, h) = & (L'_u(t_1, x(t_1), x'(t_1)) + g'_{z_1}(x(t_0), x(t_1))) h(t_1) + \\ & + (-L'_u(t_0, x(t_0), x'(t_0)) + g'_{z_0}(x(t_0), x(t_1))) h(t_0). \end{aligned} \quad (27)$$

Поскольку  $f'(x, h) = 0$  для произвольной функции  $h$  из  $X$ , то из (27) следуют равенства (25), (26).

Верно и обратное утверждение. Если выполнены условия 1), 2), то  $x$  – критическая точка функционала  $f$ . Для доказательства достаточно воспользоваться проведёнными выше выкладками, но уже в обратном порядке. ►

Непосредственным следствием леммы 6 является

**Теорема 4.** Пусть  $\hat{x}$  – локальное решение задачи (23). Тогда  $\hat{x}$  есть решение уравнения Эйлера, удовлетворяющее граничным условиям (25), (26).

Граничные условия (25), (26) называют естественными. В исходной вариационной задаче (23) эти условия отсутствуют, однако они возникают при её анализе. В качестве примера рассмотрим функционал

$$f(x) = \int_{\Delta} (a(t)x^2(t) + b(t)x^2(t) - 2c(t)x(t)) dt + k_0x^2(t_0) + k_1x^2(t_1).$$

Здесь  $a(t), b(t), c(t)$  – непрерывные на отрезке  $\Delta$  функции,  $k_0, k_1$  – действительные постоянные. В этом случае

$$L(t, x, u) = a(t)u^2 + b(t)x^2 - 2c(t)x, \quad g(z_0, z_1) = k_0z_0^2 + k_1z_1^2.$$

Естественные краевые условия имеют вид

$$a(t_1)x'(t_1) + k_1x(t_1) = 0, \quad a(t_0)x'(t_0) = k_0x(t_0).$$

В математической физике в подобной ситуации говорят о смешанном краевом условии ( $k_0^2 + k_1^2 > 0$ ) и задаче Неймана при  $k_0 = k_1 = 0$ .

В теореме 4 содержатся необходимые условия локального минимума первого порядка для задачи (23). Они могут быть дополнены достаточными условиями минимума. Например, если функции  $L(t, \cdot), g$  выпуклы на  $\mathbb{R}^2$ , то необходимые условия локального минимума, даваемые теоремой (4), являются достаточными для абсолютного минимума. Двукратная непрерывная дифференцируемость функций  $L, g$  гарантирует двукратную дифференцируемость функционала  $f$  и позволяет сформулировать условия минимума второго порядка, аналогичные изложенным в предшествующем параграфе применительно к простейшей вариационной задаче.

## 1.3 Вариационные принципы

**1. Принцип Ферма и закон Снеллиуса.** Хорошо известно, что в однородной среде свет распространяется по прямой линии. Если же среда неоднородна, то траектория света может быть и криволинейной; например, если имеются две однородные среды, занимающие две полуплоскости, то при переходе через плоскость раздела происходит преломление света. Соответствующий закон преломления был найден голландским учёным Снеллиусом.

Математическое объяснение закона преломления было дано Ферма. Для его формулировки введём понятие оптической длины пути. Пусть  $v = v(x, y, z)$  – скорость распространения луча света в точке  $(x, y, z)$ ,  $\Gamma$  – траектория луча света,  $ds$  – элемент длины кривой  $\Gamma$ . Тогда по определению скорости  $ds/dt = v(x, y, z)$ . Отсюда время прохождения траектории  $\Gamma$  равно криволинейному интегралу первого рода

$$T = \oint_{\Gamma} \frac{ds}{v(x, y, z)};$$

число  $T$  называют оптической длиной пути. Принцип Ферма гласит: *из всех траекторий, соединяющих точки  $A, B$ , свет выбирает ту, оптическая длина которой минимальна.* Принцип Ферма сводит задачу оптики к следующей вариационной задаче:

$$T = \oint_{\Gamma} \frac{ds}{v(x, y, z)} \rightarrow \min; \quad (1)$$

минимизация происходит по всем кривым, соединяющим точки  $A, B$ .

Применим принцип Ферма к задаче Снеллиуса о законе преломления. Пусть в плоскости  $\Pi$  даны две точки  $A, B$ , расположенные по разные стороны от прямой  $l$ , разделяющей две среды. Требуется на прямой  $l$  найти точку  $C$  такую, что время преодоления пути  $ACB$  было минимальным при условии, что скорость света в содержащей точку  $A$  полуплоскости  $\Pi_+$  равна  $v_1$ , а скорость в другой полуплоскости (содержащей точку  $B$ ) равна  $v_2$ . Для формализации задачи направим ось  $OX$  по прямой  $l$ , разделяющей две среды, а ось  $OY$  проведём перпендикулярно прямой  $l$  через точку  $A$ . Пусть координаты точек  $A, B$  оказались такими  $A = (0, a), B = (d, -b)$  ( $a > 0, b > 0, d > 0$ ). Возьмём на оси  $OX$  точку  $C$  с координатами  $(x, 0)$ . Тогда оптическая длина пути  $ACB$  равна

$$T = \frac{|AC|}{v_1} + \frac{|CB|}{v_2} = \frac{\sqrt{a^2 + x^2}}{v_1} + \frac{\sqrt{b^2 + (d - x)^2}}{v_2}.$$

В итоге возникла следующая экстремальная задача без ограничений

$$f(x) := \frac{\sqrt{a^2 + x^2}}{v_1} + \frac{\sqrt{b^2 + (d - x)^2}}{v_2} \rightarrow \min. \quad (2)$$

Минимизируемая функция дважды дифференцируема и строго выпукла. Она имеет единственную критическую точку, определяемую из уравнения

$$f'(x) = \frac{x}{v_1 \sqrt{a^2 + x^2}} - \frac{d-x}{v_2 \sqrt{b^2 + (d-x)^2}} = 0. \quad (3)$$

В силу выпуклости функции  $f$  решение  $\hat{x}$  уравнения (3) есть решение экстремальной задачи (2). Равенство

$$\frac{\hat{x}}{v_1 \sqrt{a^2 + \hat{x}^2}} = \frac{d - \hat{x}}{v_2 \sqrt{b^2 + (d - \hat{x})^2}}$$

эквивалентно соотношению

$$\frac{\sin \alpha_1}{v_1} = \frac{\sin \alpha_2}{v_2}, \quad (4)$$

выражающему закон Снеллиуса: *отношение синусов угла падения и угла отражения равно отношению скоростей в первой и второй среде.*

**2. Задача Лопиталья.** Пусть скорость распространения света в плоскости  $xOy$  равна  $v(x, y)$ . Найдём световой луч, соединяющий точки  $(x_0, y_0), (x_1, y_1)$ . Будем считать, что  $x_0 < x_1$  и траектория распространения светового луча задаётся явным уравнением  $y = y(x)$  ( $x_0 \leq x \leq x_1$ ). В силу принципа Ферма поставленная задача сводится к следующей простейшей вариационной задаче

$$J(y) := \int_{x_0}^{x_1} \frac{\sqrt{1 + y'^2(x)}}{v(x, y(x))} dx \rightarrow \min, \quad (5)$$

$$y(x_0) = y_0, \quad y(x_1) = y_1.$$

Если функция  $v(x, y)$  зависит лишь от одного из переменных  $x, y$ , то задача (5) существенно упрощается. Рассмотрим вначале случай, когда  $v$  зависит лишь от  $y$  (задача Лопиталья). В этом случае интегрант задачи (5)

$$L(x, y, z) = \frac{\sqrt{1 + z^2}}{v(y)}$$

не зависит от  $x$ , поэтому имеет место интеграл энергии  $y' L'_z - L = C_1$ , что в рассматриваемом случае эквивалентно соотношению

$$\sqrt{1 + y'^2} v(y) = C_1 \quad (\text{закон Снеллиуса}).$$

Отсюда вытекают равенства

$$1 + \left( \frac{dy}{dx} \right)^2 = \frac{C_1^2}{v^2(y)}, \quad \frac{dy}{dx} = \frac{\sqrt{C_1^2 - v^2(y)}}{v(y)}.$$

В итоге получено дифференциальное уравнение первого порядка с разделяющимися переменными. Таким образом, имеем

$$x + C_2 = \int \frac{v(y)}{\sqrt{C_1^2 - v^2(y)}} dy = \Phi(y, C_1). \quad (6)$$

Для определения неизвестных констант  $C_1, C_2$  следует воспользоваться краевыми условиями  $y(x_0) = y_0, y(x_1) = y_1$ . После этого найденное решение  $y(x; C_1, C_2)$  краевой задачи для уравнения Эйлера необходимо исследовать на оптимальность, используя, например, достаточные условия минимума.

Если функция  $v$  зависит лишь от  $x$ , то функция  $L(x, y, z) = \frac{\sqrt{1+z^2}}{v(x)}$  при любом  $x$  выпукла по совокупности переменных  $y, z$ , что влечёт за собой выпуклость функционала  $J$  и совпадение необходимых и достаточных условий минимума. В этом случае имеет место интеграл импульса  $L'_z = C_1$ , что приводит к равенствам

$$\frac{y'}{\sqrt{1+y'^2}} = C_1 v(x), \quad y' = \frac{C_1 v(x)}{\sqrt{1 - C_1^2 v^2(x)}}.$$

Отсюда находится общее решение уравнения Эйлера  $y(x) = \Phi(x, C_1) + C_2$ ; здесь  $\Phi(x, C_1)$  – некоторая первообразная функции

$$\frac{C_1 v(x)}{\sqrt{1 - C_1^2 v^2(x)}}.$$

Константы  $C_1, C_2$  определяются из граничных условий задачи (5). Соответствующее решение  $y(x; C_1, C_2)$  краевой задачи для уравнения Эйлера в данном случае является решением задачи (5).

**3. Частные случаи задачи Лопиталя.** Пусть  $v(y) = ky^m$ , где  $y > 0, k$  – положительная постоянная. Не нарушая общности, можно считать  $k = 1$ ; на возражения физиков ответим, что равенства  $k = 1$  можно добиться, изменяя единицы измерения.

1) Наиболее прост случай  $m = 0$ . В этом случае траектории луча света есть прямые — факт очевидный и без вариационного исчисления.

2) Подробнее остановимся на случае  $m = 1$ . В этом случае  $v(y) = y$ ; из равенства (6) следует соотношение

$$x + C_2 = \int \frac{y}{\sqrt{C_1^2 - y^2}} dy.$$

Поэтому экстремальными рассматриваемой задачи являются окружности, определяемые уравнением  $(x + C_2)^2 + y^2 = C_1^2$ . Центры этих окружностей расположены на оси абсцисс. Искомой будет та экстремаль, которая проходит через заданные точки  $(x_0, y_0), (x_1, y_1)$ , где  $y_0 > 0, y_1 > 0$ . Задача имеет единственное решение, так как через любые две точки, лежащие в верхней полуплоскости, проходит

одна и только одна полуокружность с центром на оси абсцисс. Экстремали, проходящие через точку  $(x_0, y_0)$ , попарно не пересекаются. На основании признака Якоби в его геометрической форме (см. п. 1.7) отсюда вытекает, что экстремаль, проходящая через точки  $(x_0, y_0)$ ,  $(x_1, y_1)$ , реализует локальный минимум функционала  $J$ .

С рассматриваемой задачей связана модель плоскости Лобачевского, предложенная французским математиком А. Пуанкаре. Можно показать, что часть  $AD$  полуокружности  $\Gamma$ , один из концов которой лежит на оси абсцисс, имеет бесконечную оптическую длину. Точки оси  $OX$  будем называть бесконечно удалёнными. Будем считать полуокружности с центрами на оси абсцисс прямыми, оптические длины таких полуокружностей – их длинами, углами между такими прямыми – углы между касательными в точке их пересечения. Получим плоскую геометрию, в которой сохраняются многие положения обычной геометрии. Например, через две точки можно провести одну и только одну прямую. Параллельными будем считать две прямые, имеющие общую бесконечно удалённую точку (т.е. две полуокружности, касающиеся друг друга в точке, лежащей на оси абсцисс). Тогда через данную точку, не лежащую на прямой  $\Gamma$ , можно провести две прямые  $\Gamma_1$  и  $\Gamma_2$ , параллельные  $\Gamma$ .

3) При  $m > 0$  ситуация во многом аналогична возникающей при  $m = 1$ . Рассмотрим случай  $m = 1/2$ , встречавшийся ранее в задаче о брахистохроне. Из закона Снеллиуса следует равенство  $y(1 + y'^2) = C_1$ . Проинтегрируем это уравнение в параметрической форме, полагая

$$y' = \operatorname{tg} \frac{\varphi}{2}, \quad (\varphi \neq \pm\pi).$$

Отсюда получаем

$$y = C_1 \cos^2 \frac{\varphi}{2} \quad (C_1 > 0).$$

Поскольку

$$dy = -C_1 \cos \frac{\varphi}{2} \sin \frac{\varphi}{2} d\varphi = y' dx = \operatorname{tg} \frac{\varphi}{2} dx,$$

то

$$dx = -C_1 \cos^2 \frac{\varphi}{2} d\varphi = -\frac{C_1}{2} (1 + \cos \varphi) d\varphi.$$

Интегрируя, приходим к соотношению

$$x = -\frac{C_1}{2} (\varphi + \sin \varphi) + C'.$$

Последнее уравнение преобразуем, переходя от параметра  $\varphi$  к параметру  $t$  с помощью равенства  $\varphi = \pi - t$ . Тогда

$$x = -\frac{C_1}{2} (\pi - t + \sin(\pi - t)) + C',$$

или окончательно

$$x = \frac{C_1}{2} (t - \sin t) + C_2.$$

Вспоминая выражение для  $y$ , получаем

$$y = C_1 \cos^2 \frac{\varphi}{2} = \frac{C_1}{2}(1 + \cos \varphi) = \frac{C_1}{2}(1 - \cos t).$$

Таким образом, экстремальными в задаче о брахистохроне являются **циклоиды**, параметрические уравнения которых имеют вид

$$x = \frac{C_1}{2}(t - \sin t) + C_2; \quad y = \frac{C_1}{2}(1 - \cos t),$$

причём геометрический смысл  $C_1$  – диаметр окружности, катящейся без скольжения по прямой  $OX$ , а  $C_2$  – величина сдвига вдоль оси  $OX$  в том или ином направлении. Полученные циклоиды имеют точки возврата на оси  $OX$  при  $t = 0, \pm 2\pi, \dots$ , где  $y'$  обращается в бесконечность. Следовательно, задача о брахистохроне решений класса  $C^1$  не имеет.

4) При  $m = -1$  задача Лопиталья с  $v(y) = y^{-1}$  эквивалентна задаче о поверхности вращения наименьшей площади, ставящейся следующим образом: среди множества гладких кривых, проходящих через две фиксированные точки  $A$  и  $B$ , найти такую кривую, чтобы поверхность вращения этой кривой вокруг прямой  $l$ , лежащей в её плоскости, имела наименьшую площадь.

Формализуем задачу. Будем считать, что прямая  $l$  совпадает с осью абсцисс в координатной плоскости  $XOY$ , а рассматриваемые кривые являются графиками положительных функций  $y$  класса  $C^1(\Delta)$ ,  $\Delta = [x_0, x_1]$ ,  $-\infty < x_0 < x_1 < \infty$ , причём  $y(x_0) = y_0$ ,  $y(x_1) = y_1$ ; здесь  $(x_0, y_0)$  и  $(x_1, y_1)$  – координаты точек  $A$  и  $B$  соответственно. Поставленная геометрическая задача сводится к следующей простейшей вариационной задаче

$$S(y) = 2\pi \int_{\Delta} y(x) \sqrt{1 + y'^2(x)} dx \rightarrow \min,$$

$$y(x_0) = y_0, \quad y(x_1) = y_1.$$

Здесь  $S(y)$  – площадь соответствующей поверхности вращения. Из закона Снеллиуса следует соотношение

$$y = C_1 \sqrt{1 + y'^2}.$$

Для интегрирования полученного уравнения введём новое переменное  $\varphi$  с помощью равенства

$$y' = \operatorname{sh} \varphi,$$

из которого, пользуясь свойствами гиперболических функций, получаем

$$1 + y'^2 = \operatorname{ch}^2 \varphi.$$

Следовательно,  $y = C_1 \operatorname{ch} \varphi$  и  $dy = C_1 \operatorname{sh} \varphi d\varphi$ . Деля обе части последнего равенства на  $dx$  и сравнивая с соотношением, вводящим новое переменное, получаем

простое дифференциальное уравнение  $dx = C_1 d\varphi$ , из которого находим сначала  $x = C_1\varphi + C_2$ , а затем и

$$\varphi = \frac{x - C_2}{C_1}, \quad C_1 \neq 0.$$

Подставляя найденное значение  $\varphi$  в полученное ранее соотношение  $y = C_1 \operatorname{ch} \varphi$ , получаем двухпараметрическое семейство цепных линий

$$y = C_1 \operatorname{ch} \frac{x - C_2}{C_1},$$

с осями симметрии, параллельными оси ординат и задаваемыми уравнениями  $x = C_2$ . Вершины цепных линий располагаются в точках  $(C_2, C_1)$ .

Далее следует выяснить, можно ли через точки  $A, B$  провести кривую линии рассматриваемого семейства. Обсудим этот вопрос в случае  $x_0 = -1$ ,  $x_1 = 1$ ,  $y_0 = y_1 = h > 0$ . В этой ситуации вопрос сводится к задаче о разрешимости уравнения

$$C \operatorname{ch} \frac{1}{C} = h$$

относительно  $C$ . Несложные вычисления показывают, что функция  $\psi(\xi) := \xi \operatorname{ch} \frac{1}{\xi}$  строго выпукла и бесконечно дифференцируема на луче  $(0, \infty)$ . Она достигает абсолютного минимума в единственной точке  $\xi_0$  и  $\psi(\xi_0) \approx 1,51$ . Если  $h < \psi(\xi_0)$ , то экстремали, соединяющей точки  $A, B$ , нет. При  $h = \psi(\xi_0)$  будет одна экстремаль, соединяющая точки  $A, B$ , а при  $h > \psi(\xi_0)$  таких экстремалей будет две. В этом последнем случае мы получаем два значения  $C$ : меньшему из них соответствует более низкая цепная линия, а большему – более высокая. Для нижней цепной линии условие Якоби не выполнено, а для верхней оно выполнено. Поэтому нижняя цепная линия не даёт минимума, а верхняя даёт сильный минимум (соответствующие выкладки можно найти в [7]).

**4. Принцип Торичелли и задача о прогибе тяжелой однородной нити.** Определим форму однородной, гибкой, но нерастяжимой нити длины  $l$ , подвешенной своими концами в точках  $A = (x_0, y_0)$  и  $B = (x_1, y_1)$ , ( $x_0 < x_1$ ), если на неё действует только сила тяжести. Основу решения составляет принцип Торичелли: *тяжёлая система материальных точек с идеальными связями находится в равновесии только при том условии, что высота её центра масс имеет стационарное значение.*

Если нить задаётся как график функции  $y = y(x)$ ,  $x \in \Delta = [x_0, x_1]$ , то координаты  $(x_c, y_c)$  её центра масс определяются равенствами

$$x_c = \frac{1}{l} \int_{\Delta} x \sqrt{1 + y'^2(x)} dx, \quad y_c = \frac{1}{l} \int_{\Delta} y(x) \sqrt{1 + y'^2(x)} dx.$$

Поэтому задача о прогибе нити сводится к нахождению критической точки функционала

$$J_0(y) = \int_{\Delta} y(x) \sqrt{1 + y'^2(x)} dx$$

при условиях связи

$$J_1(y) = \int_{\Delta} \sqrt{1 + y'^2(x)} dx = l, \quad y(x_0) = y_0, y(x_1) = y_1.$$

Решение поставленной задачи есть экстремаль функционала

$$J(y) = J_0(y) + \lambda J_1(y) = \int_{\Delta} (y(x) + \lambda) \sqrt{1 + y'^2(x)} dx$$

при некотором  $\lambda$ . Положим  $z(x) = y(x) + \lambda$ . Тогда всё сводится к поиску экстремалей функционала

$$I(z) = \int_{\Delta} z(x) \sqrt{1 + z'^2(x)} dx.$$

В силу результатов предшествующего пункта, относящихся к случаю  $v(y) = y^{-1}$ , семейство экстремалей этого функционала имеет вид

$$z(x; C_1, C_2) = C_1 \operatorname{ch} \frac{x - C_2}{C_1}.$$

Следовательно,

$$y(x) = C_3 + C_1 \operatorname{ch} \frac{x - C_2}{C_1}.$$

Постоянные  $C_1, C_2, C_3$  находятся из условий связи. Исследование возникающей здесь системы уравнений составляло (и будет составлять) тему многих курсовых, дипломных и других квалификационных работ.

Рассмотрим близкую задачу о равновесии стержня. Пусть стержень длины  $l$  заделан в точках  $A = (x_0, y_0)$  и  $B = (x_1, y_1)$ . Из теории упругости известно, что потенциальная энергия стержня в деформированном состоянии пропорциональна интегралу, взятому вдоль стержня, от квадрата его кривизны. Примем за независимую переменную длину стержня  $s$ , отсчитываемую от точки  $A$ , и обозначим через  $\theta(s)$ ,  $s \in \Delta = [0, l]$  угол, образованный касательной к стержню с осью  $OX$ . Кривизна равна  $\theta'(s)$ , а потенциальная энергия деформированного стержня (с точностью до постоянного множителя) выражается интегралом

$$J_0(\theta) := \int_{\Delta} \theta'^2(s) ds.$$

Справедливы равенства

$$\frac{dx}{ds} = \cos \theta, \quad \frac{dy}{ds} = \sin \theta,$$



приводящие к следующим условиям связи

$$J_1(\theta) := \int_{\Delta} \cos \theta(s) ds = x_1 - x_0, \quad J_2(\theta) := \int_{\Delta} \sin \theta(s) ds = y_1 - y_0. \quad (7)$$

Кроме того, заделанность стержня эквивалентна граничным условиям

$$\theta(0) = a_0, \quad \theta(l) = a_1, \quad (8)$$

где  $a_0, a_1$  – некоторые фиксированные числа. Воспользуемся принципом стационарности потенциальной энергии: *система материальных точек находится в равновесии тогда и только тогда, когда её потенциальная энергия имеет стационарное значение.*

В соответствии с этим принципом характеризующая положение равновесия стержня функция  $\theta(s)$ ,  $s \in \Delta$  является экстремалью функционала Лагранжа

$$J(\theta) = J_0(\theta) + \lambda_1 J_1(\theta) + \lambda_2 J_2(\theta) = \int_{\Delta} \left[ \theta'^2(s) + \lambda \cos \theta(s) + \lambda_2 \sin \theta(s) \right] ds,$$

где  $\lambda_1, \lambda_2$  – множители Лагранжа (проверка условий регулярности условий связи (7) предоставляется читателю). Интегрант функционала  $J$  не содержит независимой переменной  $s$ , а потому можно сразу выписать первый интеграл уравнения Эйлера:

$$\theta'^2 = C + \lambda_1 \cos \theta + \lambda_2 \sin \theta. \quad (9)$$

Введём новые постоянные

$$h = C + \sqrt{\lambda_1^2 + \lambda_2^2}, \quad k^2 = \frac{2\sqrt{\lambda_1^2 + \lambda_2^2}}{C + \sqrt{\lambda_1^2 + \lambda_2^2}}$$

и вместо  $\theta$  введём новую переменную

$$\varphi = \frac{\theta - \theta_0}{2}, \quad \text{где} \quad \theta_0 = \arctg \frac{\lambda_2}{\lambda_1}.$$

Теперь (9) приводится к виду

$$\frac{d\varphi}{ds} = \frac{\sqrt{h}}{2} \sqrt{1 - k^2 \sin^2 \varphi},$$

откуда получаем

$$s = \frac{2}{\sqrt{h}} \int \frac{d\varphi}{\sqrt{1 - k^2 \sin^2 \varphi}} + s_0.$$

Постоянные  $h, k^2, \theta_0, s_0$  должны определяться из условий (7), (8).

Декартовы координаты точек стержня находятся так:

$$dx = \cos \theta ds = \cos(2\varphi + \theta_0) ds, \quad dy = \sin \theta ds = \sin(2\varphi + \theta_0) ds,$$

или в силу того, что

$$ds = \frac{2d\varphi}{\sqrt{h}\sqrt{1-k^2\sin^2\varphi}},$$

получим

$$dx = \frac{2\cos(2\varphi + \theta_0)}{\sqrt{h(1-k^2\sin^2\varphi)}} d\varphi, \quad dy = \frac{2\sin(2\varphi + \theta_0)}{\sqrt{h(1-k^2\sin^2\varphi)}} d\varphi,$$

откуда  $x$  и  $y$  определяются с помощью квадратур.

**5. Принцип стационарного действия.** Рассмотрим систему  $n$  материальных точек с массами  $m_k$  и координатами  $x_k, y_k, z_k$  ( $k = 1, \dots, n$ ). Предположим, что движение системы подчинено связям

$$\varphi_j(t, x, y, z) = 0 \quad (j = 1, \dots), \quad \text{где } m < n, \quad (10)$$

и происходит под действием сил  $F_k(X_k, Y_k, Z_k)$  ( $k = 1, \dots, n$ ), имеющих потенциал (силовую функцию)  $U = U(x, y, z, t)$ :

$$X_k = \frac{\partial U}{\partial x_k}, \quad Y_k = \frac{\partial U}{\partial y_k}, \quad Z_k = \frac{\partial U}{\partial z_k} \quad (k = 1, \dots, n).$$

Кинетическая энергия этой системы равна

$$T = \frac{1}{2} \sum_{k=1}^n m_k (\dot{x}_k^2 + \dot{y}_k^2 + \dot{z}_k^2).$$

Пусть из некоторого положения  $A$ , соответствующего моменту времени  $t = t_0$ , эта система переместилась к моменту времени  $t = t_1$  в некоторое положение  $B$ . Из всех возможных перемещений системы из  $A$  в  $B$  выбирается класс допустимых движений, а именно тех движений, которые совместимы с условиями связи (10) и в заданный промежуток времени  $\Delta = [t_0, t_1]$  переводят систему из  $A$  в  $B$ .

Обозначим кинетический потенциал (функцию Лагранжа) через  $L = T + U$ . Каждому допустимому движению  $x(t), y(t), z(t)$  ( $t \in \Delta$ ) сопоставим интеграл

$$W = \int_{\Delta} L dt, \quad (11)$$

называемый интегралом действия. Его можно рассматривать как функционал, определённый на допустимых действиях. **Принцип стационарного действия** можно сформулировать следующим образом: *кривая  $x(t), y(t), z(t)$ , определяющая закон движения системы, согласуемый со связями (10), является экстремалью для интеграла действия при дополнительных условиях – уравнениях связи (10).*

Это означает, что она должна быть решением системы уравнений Эйлера, построенной для интегранта

$$\bar{L} = L + \sum_{j=1}^m \lambda_j \varphi_j.$$

Подставляя функцию  $\bar{L}$  в систему уравнений Эйлера

$$\frac{d}{dt} \bar{L}'_{x_i} = \bar{L}'_{x_i}, \quad \frac{d}{dt} \bar{L}'_{y_i} = \bar{L}'_{y_i}, \quad \frac{d}{dt} \bar{L}'_{z_i} = \bar{L}'_{z_i}, \quad (k = 1, \dots, n),$$

получим соотношения:

$$m_k \ddot{x}_k = \frac{\partial U}{\partial x_k} + \sum_{j=1}^m \lambda_j \frac{\partial \varphi_j}{\partial x_k};$$

$$m_k \ddot{y}_k = \frac{\partial U}{\partial y_k} + \sum_{j=1}^m \lambda_j \frac{\partial \varphi_j}{\partial y_k};$$

$$m_k \ddot{z}_k = \frac{\partial U}{\partial z_k} + \sum_{j=1}^m \lambda_j \frac{\partial \varphi_j}{\partial z_k},$$

являющиеся уравнениями Лагранжа первого рода механической системы.

Отметим частный, но весьма важный случай. Пусть имеется одна материальная точка единичной массы ( $m = 1$ ), одно условие связи  $\varphi(x, y, z) = 0$  и отсутствуют внешние силы ( $F = 0$ ). В этом случае уравнения Лагранжа первого рода существенно упрощаются и принимают вид

$$\frac{d^2 x}{dt^2} = \lambda(t) \frac{\partial \varphi}{\partial x}, \quad \frac{d^2 y}{dt^2} = \lambda(t) \frac{\partial \varphi}{\partial y}, \quad \frac{d^2 z}{dt^2} = \lambda(t) \frac{\partial \varphi}{\partial z}.$$

Если  $\vec{r}(t) = (x(t), y(t), z(t))$  – радиус-вектор рассматриваемой материальной точки, то справедливы соотношения

$$\varphi(\vec{r}) = 0, \quad \frac{d^2 \vec{r}}{dt^2} = \lambda(t) \nabla \varphi(r). \quad (12)$$

Дифференцируя условие связи  $\varphi(\vec{r}(t)) = 0$  по  $t$ , приходим к равенству

$$\left( \nabla \varphi(\vec{r}(t)), \frac{d\vec{r}(t)}{dt} \right) = 0.$$

Это равенство и (12) приводят к соотношению

$$\frac{d}{dt} \left| \frac{d\vec{r}(t)}{dt} \right|^2 = 2 \left( \vec{r}(t), \frac{d\vec{r}(t)}{dt} \right) \equiv 0,$$

означающему, что скорость  $v(t) = |\vec{r}'(t)|$  движения точки постоянна. Если  $s$  – натуральный параметр кривой  $\vec{r} = \vec{r}(t)$ , то  $v = s'_t$ . Справедливо равенство

$$\frac{d^2\vec{r}}{dt^2} = \frac{d^2\vec{r}}{ds^2}v^2 + \frac{d\vec{r}}{ds}v'_t = \frac{d^2\vec{r}}{ds^2}v^2 = \lambda(t)\nabla\varphi(\vec{r}),$$

имеющее простой геометрический смысл: главная нормаль к кривой  $\vec{r} = \vec{r}(t)$  совпадает с нормалью к поверхности  $\varphi(\vec{r}) = 0$ . Отсюда вытекает, что кривая  $\vec{r} = \vec{r}(t)$  является геодезической. Суммируя, приходим к весьма интересному и с точки зрения механики, и с точки зрения математики выводу: скорость перемещения материальной точки по поверхности постоянна, а траектория движения будет геодезической.

Коснёмся уравнений Лагранжа второго рода. Для их вывода тоже можно воспользоваться необходимыми условиями для соответствующей вариационной задачи.

Рассмотрим произвольную механическую систему в промежуток времени  $\Delta = [t_0, t_1]$ . Пусть состояние системы определяется обобщёнными координатами  $q_i$  и обобщёнными скоростями  $\dot{q}_i$ ,  $i = 1, \dots, n$ . Тогда функция Лагранжа будет определяться временем  $t$  и этими величинами:  $L = L(t, q_i, \dot{q}_i)$ , а интеграл действия  $W$  – равенством (11). **Принцип стационарного действия** в этом случае формулируется следующим образом: *кривая  $q_i = q_i(t)$  ( $i = 1, \dots, n$ ), описывающая реальное движение механической системы есть экстремаль для интеграла действия, то есть для неё должны иметь место уравнения Эйлера*

$$\frac{d}{dt}L_{\dot{q}_i} = L_{q_i}, \quad (i = 1, \dots, n),$$

совпадающие в рассматриваемом случае с уравнениями Лагранжа второго рода.

В заключение главы позволю себе привести достаточно обширную цитату из замечательной книги К. Ланцоша «Вариационные принципы механики» [23]. Вот что пишет автор этого исключительного по красоте произведения: «в задачах о движении существенны лишь *стационарные значения* некоторых определённых интегралов. Поэтому имеется значительное различие между вариационным исчислением – ветвью чистой математики, и его приложением к задачам механики – с другой. С точки зрения чистой математики задача о нахождении стационарных значений не представляет большого интереса. После установления критерия для стационарных точек идут дальше и ищут дополнительные критерии для истинных экстремумов. Для вариационных принципов механики, однако, эти последние исследования представляют интерес только при решении задач устойчивости, когда ищется действительный минимум потенциальной энергии. Задачи же по определению движения не связаны со специальными условиями существования таких минимумов».

Точка зрения Ланцоша во многом спорна. Уже к середине двадцатого века трудами многих математиков было создано вариационное исчисление в целом,

та самая часть вариационного исчисления, в которой основное внимание уделяется поиску стационарных (не обязательно минимальных или максимальных) значений функционалов. Механики, физики и вообще все учёные проявляли и проявляют значительный интерес к оптимизационным задачам. Вместе с тем учебников по глобальному вариационному исчислению, доступных массовому читателю, так и нет. В научной литературе вопросы вариационного исчисления в целом достаточно широко обсуждались. Ограничусь лишь ссылкой на имеющую большую библиографию книгу [16].



## Глава 2

# Принцип максимума Понтрягина

### 2.1 Постановка задачи оптимального управления

**1. Уравнения движения объекта.** Пусть состояние некоторого объекта в моменты времени  $t \in [t_0, t_1]$  характеризуется  $n$  величинами  $x_1(t), \dots, x_n(t)$ , которые назовём фазовыми переменными и будем считать компонентами вектор-функции  $x(t) = \{x_1(t), \dots, x_n(t)\}^T$  (здесь и далее символ  $T$  означает операцию транспонирования). Предположим, что объект может изменять своё состояние под влиянием некоторых воздействий  $u_1(t), \dots, u_m(t)$ , называемых в дальнейшем управлениями. Вектор-функцию  $u(t) = (u_1(t) \dots u_m(t))^T$  назовём управлением. Размерности  $n$  и  $m$  фазового и управляющего векторов могут быть различными.

Пару вектор-функций  $x(t), u(t)$  будем называть управляемым процессом. В качестве примера управляемого процесса рассмотрим движение автомобиля в плоскости  $OX_1X_2$ . Состояние автомобиля можно охарактеризовать двумя координатами  $x_1, x_2$  и компонентами скорости  $v_1, v_2$ . Эти величины изменяются в зависимости от силы двигателя  $F = (F_1, F_2)$ . В рассматриваемом примере имеются шесть связанных между собой параметров:  $x_1, x_2, v_1, v_2, F_1, F_2$ . Силу двигателя можно принять за управление, параметры  $x_1, x_2, v_1, v_2$  играют роль фазовых переменных.

Ниже предполагается, что скорость изменения фазовых переменных определяется значениями в тот же момент фазового вектора и управления. Именно, пусть поведение объекта описывается системой дифференциальных уравнений

$$\frac{dx_i}{dt} = f_i(x, u, t) \quad (i = 1, \dots, n),$$

записываемой далее в векторном виде

$$\frac{dx}{dt} = f(x, u, t), \quad (1)$$

где  $f = (f_1 \dots f_n)^T$ . Если правая часть уравнения (1) не зависит от  $t$ , то управляемый объект называется стационарным.

Для однозначного определения фазовых переменных необходимо к системе (1) присоединить соотношение

$$x(t_0) = x_0, \quad (2)$$

характеризующее начальное положение объекта. Если некоторому управлению  $u(t)$  соответствует единственное решение  $x(t)$  дифференциального уравнения

$$\frac{dx}{dt} = f(x, u(t), t), \quad (3)$$

удовлетворяющее начальному условию (2), то  $x(t)$  называют фазовой траекторией, соответствующей управлению  $u(t)$ .

Частным случаем системы (1) является система

$$\frac{dx}{dt} = A(t)x + b(u, t). \quad (4)$$

Здесь  $A(t)$  – квадратная матрица размеров  $n \times n$ ,  $b(u, t) = (b_1(u, t), \dots, b_n(u, t))^T$  – вектор-функция переменных  $u, t$ . Объект, описываемой системой (4), будем называть линейным по фазовым переменным.

Дальнейшей специализацией системы (1) является система

$$\frac{dx}{dt} = A(t)x + B(t)u + c(t), \quad (6)$$

где  $A(t)$  имеет тот же смысл, что и в (4),  $B(t)$  – матрица размеров  $n \times m$ ,  $c(t)$  – вектор-функция. Соответствующий (5) управляемый объект назовём линейным по фазовым переменным и управлениям.

**2. Допустимые управления.** Управление  $u: [t_0, \infty) \rightarrow \mathbb{R}^m$  будем называть кусочно-непрерывным, если на каждом отрезке  $[t_0, t_1]$  оно непрерывно всюду за исключением конечного числа точек, в которых отображение  $u(t)$  имеет разрывы первого рода. Значение кусочно-непрерывного отображения  $u(t)$  в точке разрыва не играет существенной роли в дальнейшем. Условимся считать, что все рассматриваемые ниже кусочно-непрерывные функции непрерывны справа. Совокупность кусочно-непрерывных отображений  $[t_0, \infty)$  в  $\mathbb{R}^m$  обозначим символом  $\mathcal{D}_m$ .

В качестве класса допустимых управлений чаще всего используют совокупность таких отображений  $u \in \mathcal{D}_m$ , что

$$u(t) \in \Omega \quad (t \geq t_0). \quad (6)$$

Здесь  $\Omega$  – некоторое замкнутое подмножество  $\mathbb{R}^m$ ; случай  $\Omega = \mathbb{R}^m$  не исключается. Множество отображений  $u$  класса  $\mathcal{D}_m$ , для которых выполнено включение (6), обозначим символом  $\mathcal{U}(\Omega)$ .

Под решением задачи Коши (2), (3) ниже понимается непрерывная при  $t \geq t_0$  функция  $x(t)$ , удовлетворяющая уравнению (3) всюду за исключением точек



разрыва управления  $u(t)$ . Функция  $x(t)$  является решением интегрального уравнения

$$x(t) = x_0 + \int_{t_0}^t f[x(s), u(s), s] ds. \quad (7)$$

Если функции  $f_i(x, u, t), \partial f_i / \partial x_j(x, u, t)$  ( $i, j = 1, \dots, n$ ) непрерывны по совокупности переменных  $x \in \mathbb{R}^n, u \in \Omega, t \geq t_0$ , то задача Коши (2), (3) имеет единственное решение, определённое на некотором отрезке  $[t_0, t_0 + h]$ . Величина  $h$  может зависеть от выбора управления  $u$ . Если сверх того имеет место оценка

$$xf(x, u, t) \leq c(R, T_0)(1 + |x|^2), \quad (8)$$

в которой  $xy$  – скалярное произведение в  $\mathbb{R}^n$ ,  $|x|$  – евклидова норма в  $\mathbb{R}^n$ ,  $c(R, T_0)$  – постоянная не зависящая от  $x$  из  $\mathbb{R}^n$  и  $u$  из  $\Omega$ , то число  $h$  может быть произвольным.

Соотношение (8) имеет место для линейной по фазовым переменным системы (4), если только матрица  $A(t)$  и вектор-функция  $b(u, t)$  непрерывны по совокупности переменных. В общем случае (8) является достаточным условием нелокальной продолжимости решений задачи Коши (2), (3). Оно может выполняться и для некоторых нелинейных управляемых объектов.

**3. Критерий качества управления.** Пусть задан управляемый объект, описываемый системой (1), фиксированы класс допустимых управлений  $\mathcal{U}(\Omega)$ , начальное состояние объекта и множество  $V$  конечных состояний объекта. Будем говорить, что управление  $u \in \mathcal{U}(\Omega)$  переводит точку  $x_0$  в множество  $V$  за время  $t_1 - t_0$ , если правый конец  $x(t_1)$  траектории  $x(t)$ , соответствующей управлению  $u(t)$ , принадлежит  $V$ . Если существует несколько управлений, переводящих точку  $x_0$  в множество  $V$ , то возникает задача о выборе оптимального (в смысле какого-нибудь критерия) управления. Определим критерий качества управления равенством

$$J(u) = g(x(t_1), t_1), \quad (9)$$

где  $g: \mathbb{R}^n \times [t_0, \infty) \rightarrow \mathbb{R}$  – некоторая скалярная функция,  $x(t)$  – траектория, соответствующая управлению  $u(t)$ .

Задача о минимизации функционала (9) на множестве  $\mathcal{U}(\Omega)$  будет записываться следующим образом

$$J(u) = g(x(t_1), t_1) \rightarrow \min, \quad \frac{dx}{dt} = f(x, u, t), \quad (10)$$

$$x(t_0) = x_0, \quad x(t_1) \in V, \quad u \in \mathcal{U}(\Omega).$$

Процесс  $(x(t), u(t), t_1)$  называется допустимым, если  $x(t)$  – фазовая траектория, соответствующая управлению  $u(t)$  ( $t_0 \leq t \leq t_1$ ), а  $x(t_1) \in V$ . Допустимый процесс  $(\hat{x}(t), \hat{u}(t), \hat{t}_1)$  называется оптимальным, если  $J(u) = g(x(t_1), t_1) \geq J(\hat{u}) =$

$g(\hat{x}(\hat{t}_1), \hat{t}_1)$  для любого допустимого процесса. В этом случае  $\hat{u}(t)$  называют оптимальным управлением,  $\hat{x}(t)$  – оптимальной траекторией,  $\hat{t}_1$  – оптимальным временем окончания процесса.

Функционал (9) называют терминальным или функционалом качества в форме Майера. Задача с критерием качества в форме Больца

$$J_1(u) = g(x(t_1), t_1) + \int_{t_0}^{t_1} L(x(t), u(t), t) dt$$

легко сводится к задаче (10). Для этого к системе (1) присоединяют ещё одно уравнение

$$\frac{dx_{n+1}}{dt} = L(x, u, t), \quad x_{n+1}(t_0) = 0.$$

Так как

$$x_{n+1}(t_1) = \int_{t_0}^{t_1} L(x(t), u(t), t) dt,$$

то  $J_1(u) = g(x(t_1), t_1) + x_{n+1}(t_1)$ , т.е.  $J_1$  является терминальным функционалом.

Если  $V = \mathbb{R}^n$ , то (10) называют задачей со свободным правым концом. В этом случае условие  $x(t_1) \in V$  не накладывает ограничений на правый конец траектории. Задачи со свободным правым концом в ряде отношений являются наиболее простыми.

В приложениях представляет интерес вариант задачи (10), в котором время  $t_1$  окончания процесса фиксировано и не подлежит оптимизации. Введённые выше определения допустимого и оптимального процессов сохраняются, нужно лишь считать время окончания процесса  $t_1$  постоянной величиной. В этом случае будем предполагать, что управление  $u(t)$  непрерывно слева в точке  $t_1$ .

## 2.2 Оптимизация линейных систем

**1. Простейшая задача оптимального управления.** Изучение задач оптимального управления начнём с рассмотрения частного случая, весьма простого с точки зрения используемых средств, но вместе с тем достаточно характерного для данного круга вопросов. На взгляд автора, простейшей представляется задача

$$\begin{aligned} J(u) &= px(t_1) \rightarrow \min, \\ \frac{dx}{dt} &= A(t)x + b(u, t), \\ x(t_0) &= x_0, \quad u \in \mathcal{U}(\Omega). \end{aligned} \tag{1}$$

Здесь  $p \in (\mathbb{R}^n)^*$  – вектор-строка,  $px = p_1x_1 + \dots + p_nx_n$  – произведение строки  $p = (p_1 \dots p_n)$  на столбец  $x = (x_1 \dots x_n)^T$ ,  $t_1$  – фиксированная величина ( $t_1 > t_0$ ),  $A(t), b(u, t)$  имеют тот же смысл, что и в (1.4).

Будем предполагать, что  $b: \Omega \times [t_0, t_1] \rightarrow \mathbb{R}^n$  – непрерывное по совокупности переменных отображение, элементы  $a_{ij}(t)$  матрицы  $A(t)$  кусочно непрерывны на отрезке  $[t_0, t_1]$ .

Задача (1) представляет весьма упрощенный вариант задачи (1.10). Упрощение связано с тем, что рассматривается линейный по фазовым переменным управляемый объект, время окончания процесса фиксировано, правый конец предполагается свободным, критерий качества управления линейно зависит от правого конца траектории.

Далее будет неоднократно использоваться одно вспомогательное утверждение о линейной системе

$$\frac{dx}{dt} = A(t)x + w(t). \quad (2)$$

Здесь  $w(t)$  – кусочно-непрерывная на отрезке  $[t_0, t_1]$  вектор-функция. Сопряжённой к (2) называется однородная система

$$\frac{d\psi}{dt} = -\psi A(t), \quad (3)$$

где  $\psi$  – вектор-строка с  $n$  компонентами. Вектор  $\psi$  называется импульсом, а его компоненты  $\psi_1, \dots, \psi_n$  – сопряжёнными переменными. Под решением системы (3) понимается непрерывная вектор-строка  $\psi(t)$ , удовлетворяющая интегральному уравнению

$$\psi(t) = \psi(t_1) + \int_t^{t_1} \psi(\tau) A(\tau) d\tau;$$

если матричная функция  $A(\cdot)$  непрерывна в точке  $t$ , то функция  $\psi$  дифференцируема в точке  $t$  и справедливо равенство (3).

**Лемма 1.** Пусть  $x(t)$  – решение системы (2),  $\psi(t)$  – решение системы (3). Тогда

$$\psi(t_1)x(t_1) - \psi(t_0)x(t_0) = \int_{t_0}^{t_1} \psi(t)w(t) dt. \quad (4)$$

◀ Всюду (кроме точек разрыва матричной функции  $A(t)$  и вектор-функции  $w(t)$ ) имеем

$$\begin{aligned} \frac{d}{dt}(\psi(t)x(t)) &= \frac{d\psi(t)}{dt}x(t) + \psi(t)\frac{dx(t)}{dt} = \\ &= -\psi(t)A(t)x(t) + \psi(t)(A(t)x(t) + w(t)) = \psi(t)w(t). \end{aligned}$$

Интегрируя полученное равенство по отрезку  $[t_0, t_1]$ , приходим (в силу формулы Ньютона–Лейбница) к соотношению (4). ►

Равенство (4) будем называть основным свойством сопряжённой системы.

**2. Решение простейшей задачи.** Ниже будет дана характеристика решения задачи (10), приводящая к достаточно простому способу его отыскания.

**Теорема 1.** Пусть  $\hat{\psi}(t)$  есть решение обратной задачи Коши

$$\frac{d\psi}{dt} = -\psi A(t), \quad \psi(t_1) = -p. \quad (5)$$

Для того чтобы управление  $\hat{u}(t)$  было решением задачи (1), необходимо и достаточно, чтобы для каждого  $t$  из отрезка  $[t_0, t_1]$  имело место соотношение

$$\hat{\psi}(t)b(\hat{u}(t), t) = \max_{v \in \Omega} \hat{\psi}(t)b(v, t). \quad (6)$$

◀ Пусть  $u \in \mathcal{U}(\Omega)$ ,  $x(\cdot)$  – соответствующая управлению  $u$  фазовая траектория. Согласно лемме 1

$$\hat{\psi}(t_1)x(t_1) - \hat{\psi}(t_0)x(t_0) = \int_{t_0}^{t_1} \hat{\psi}(t)b(u(t), t) dt,$$

Отсюда вытекает равенство

$$J(u) = -\hat{\psi}(t_0)x_0 - \int_{t_0}^{t_1} \hat{\psi}(t)b(u(t), t) dt,$$

означающее, что задача (1) эквивалентна задаче

$$J_1(u) = \int_{t_0}^{t_1} \hat{\psi}(t)b(u(t), t) dt \rightarrow \max, \quad u \in \mathcal{U}(\Omega). \quad (7)$$

Если управление  $\hat{u}$  удовлетворяет соотношению (6), то, очевидно,  $J_1(\hat{u}) \geq J_1(u) \forall u \in \mathcal{U}(\Omega)$ , т.е.  $\hat{u}$  – решение задач (1), (7). В части достаточности теорема 1 доказана.

Пусть теперь  $\hat{u}$  – решение задачи (1). Тогда  $\hat{u}$  является и решением задачи (7). Предположим, что в некоторой точке  $\tau$  из  $\Delta$  соотношение (6) не имеет места, т.е.  $\hat{\psi}(t)b(\hat{u}(\tau), \tau) < \hat{\psi}(\tau)b(v, \tau)$  для некоторого  $v$  из  $\Omega$ . Так как функция  $\hat{\psi}(t)b(u, t)$  непрерывна, а функция  $\hat{u}$  кусочно-непрерывна, то найдётся такой отрезок  $[\tau_0, \tau_1]$  положительной длины, что  $[\tau_0, \tau_1] \subset \Delta$ ,

$$\hat{\psi}(t)b(\hat{u}(t), t) < \hat{\psi}(t)b(v, t) \quad \forall t \in [\tau_0, \tau_1] \quad (8)$$

и управление  $\hat{u}$  непрерывно в точках  $\tau_0, \tau_1$ . Положим

$$\tilde{u}(t) = \begin{cases} \hat{u}(t), & \text{если } t \notin [\tau_0, \tau_1] \\ v, & \text{если } t \in [\tau_0, \tau_1]. \end{cases}$$

Как нетрудно видеть,  $\tilde{u} \in U(\Omega)$ . В силу (8)  $J_0(\tilde{u}) > J_0(\hat{u})$ , что эквивалентно неравенству  $J(\tilde{u}) < J(\hat{u})$ . Но это противоречит оптимальности управления  $\hat{u}$ . ▶

Процесс отыскания решения задачи (1) удобно разбить на два этапа. Вначале находится решение обратной задачи Коши (5). Особенно просто первый этап реализуется в случае постоянной матрицы  $A$ . Второй этап связан с отысканием управления  $\hat{u}$ , удовлетворяющего принципу максимума (6). Сложность этой задачи зависит от структуры множества  $\Omega$  и функции  $b(u, t)$ . Если  $\Omega$  – полиэдр, а функция  $b(u, t)$  линейна по  $u$ , то нахождение  $\hat{u}$  приводит к некоторой задаче линейной оптимизации.

В качестве примера рассмотрим задачу о накоплении возмущений. Пусть поведение системы с одной степенью свободы описывается линейным уравнением

$$m \frac{d^2 q}{dt^2} + c \frac{dq}{dt} + kq = F(t), \quad (9)$$

где  $q$  – обобщённая координата системы,  $F$  – внешняя сила,  $m$  – масса,  $c$  – коэффициент трения,  $k$  – коэффициент жёсткости. Будем считать, что внешняя сила ограничена по модулю:

$$|F(t)| \leq F_0. \quad (10)$$

Пусть известно начальное состояние системы:

$$q(t_0) = q_0, \quad q'(t_0) = q_1. \quad (11)$$

Задача о накоплении возмущений состоит в отыскании такого воздействия  $F(t)$ , которое сообщает отклонению системы  $I = q(t_1)$  максимальное значение.

Положим  $x_1 = q$ ,  $x_2 = q'$ ,  $\gamma = \frac{c}{m}$ ,  $\delta = \frac{k}{m}$ ,  $u = \frac{F}{m}$ . Тогда задача о накоплении возмущений сводится к следующему частному случаю задачи (1):

$$\begin{aligned} J(u) &= -x_1(t_1) \rightarrow \min, \\ \frac{dx_1}{dt} &= x_2, \quad \frac{dx_2}{dt} = -\delta x_1 - \gamma x_2 + u, \\ x_1(t_0) &= q_0, \quad x_2(t_0) = q_1, \quad |u(t)| \leq \frac{F_0}{m}, \end{aligned} \quad (12)$$

В рассматриваемом случае

$$A = \begin{pmatrix} 0 & 1 \\ -\delta & -\gamma \end{pmatrix}, \quad p = (-1, 0), \quad \Omega = \left[ -\frac{F_0}{m}, \frac{F_0}{m} \right], \quad b(u, t) = \begin{pmatrix} 0 \\ u \end{pmatrix}.$$

Вектор-функция  $\hat{\psi}(t)$  определяется как решение обратной задачи Коши

$$\frac{d}{dt}(\psi_1, \psi_2) = -(\psi_1, \psi_2) \begin{pmatrix} 0 & 1 \\ -\delta & -\gamma \end{pmatrix}, \quad (\psi_1(t_1), \psi_2(t_1)) = (1, 0).$$

Принцип максимума (6) приводит к соотношениям

$$\hat{u}(t)\psi_2(t) = \max_{v \in \Omega} v\psi_2(t), \quad \hat{u}(t) = \frac{F_0}{m} \operatorname{sgn}(\psi_2(t)). \quad (13)$$

Как нетрудно видеть, функция  $\psi_2(t)$ , фигурирующая в (13), есть решение следующей обратной задачи Коши

$$\frac{d^2 z}{dt^2} - \gamma \frac{dz}{dt} + \delta z = 0, \quad z(t_1) = 0, \quad z'(t_1) = -1. \quad (14)$$

Поэтому для решения задачи о накоплении возмущений достаточно найти решение обратной задачи Коши (14), а затем воспользоваться формулой (13). Здесь возможны два разных случая. Если трение достаточно велико, то характеристическое уравнение  $\lambda^2 - \gamma\lambda + \delta = 0$  для дифференциального уравнения (14) имеет действительные корни. В этой ситуации на каждом отрезке  $[t_0, t_1]$  решение  $z(t)$  обращается в нуль не более чем один раз, поэтому оптимальное управление  $\hat{u}(t)$  либо постоянно, либо имеет лишь одну точку переключения. При малом трении корни характеристического уравнения перестают быть действительными, а соответствующее решение  $z(t)$  становится колеблющимся, число его корней зависит от величины отрезка  $[t_0, t_1]$ . Решение задачи (12) в этом случае есть кусочно-постоянная функция, принимающая поочередно значения  $\frac{F_0}{m}$  и  $-\frac{F_0}{m}$ .

**3. Задача об оптимальном быстродействии.** Рассматривается управляемый объект, эволюция которого описывается соотношениями

$$\begin{aligned} \frac{dx}{dt} &= A(t)x + B(t)u + c(t), \\ x(t_0) &= x_0, \quad u \in \mathcal{U}(\Omega). \end{aligned} \quad (15)$$

Здесь  $x = (x_1 \dots x_n)^T$  – фазовый вектор,  $u = (u_1 \dots u_m)^T$  – управление, матрицы  $A(t), B(t), c(t)$  кусочно-непрерывны при  $t \geq t_0$ ,  $x_0$  – фиксированный элемент  $\mathbb{R}^n$ ,  $\Omega$  – непустое выпуклое замкнутое подмножество  $\mathbb{R}^m$ . Таким образом, рассматривается линейный по фазовым переменным и управлениям объект, класс допустимых управлений которого есть выпуклое подмножество  $\mathcal{D}_m$ .

Каждому числу  $t_1 > t_0$  сопоставим множество  $\mathcal{K}(t_1)$ , составленное из положений в момент  $t_1$  фазовых траекторий, соответствующих всевозможным допустимым управлениям. Иными словами,

$$\mathcal{K}(t_1) := \{x \in \mathbb{R}^n, \quad x = x(t_1; x_0, u), \quad u \in \mathcal{U}(\Omega)\},$$

где через  $x(t; x_0, u)$  обозначена фазовая траектория, соответствующая управлению  $u$  и начальному значению  $x_0$ . Определенное таким образом множество  $\mathcal{K}(t_1)$  называют множеством достижимости управляемого объекта (15).

**Лемма 2.** При любом  $t_1 > t_0$  множество достижимости  $\mathcal{K}(t_1)$  выпукло.

◀ Пусть  $y, z$  две точки из  $\mathcal{K}(t_1)$ ,  $u, v$  – допустимые управления, переводящие точку  $x_0$  в точки  $y, z$  соответственно за время  $t_1 - t_0$ . При любом  $\lambda$  из  $(0, 1)$  управление  $w = (1 - \lambda)u + \lambda v$  является допустимым, поскольку  $\Omega$  – выпуклое подмножество  $\mathbb{R}^m$ . Из линейности управляемого объекта вытекает, что  $w$  переводит точку  $x_0$  в точку  $(1 - \lambda)y + \lambda z$  за время  $t_1 - t_0$ . Поэтому  $(1 - \lambda)y + \lambda z \in \mathcal{K}(t_1)$  при любом  $\lambda$  из  $(0, 1)$ , т.е.  $\mathcal{K}(t_1)$  выпуклое множество. ▶

Пусть  $V$  – подмножество  $\mathbb{R}^n$ , не содержащее точку  $x_0$ . Пусть существует такой управляемый процесс  $(x(t), u(t), t)$ , что  $x(t_1) \in V$  при некотором  $t_1 > t_0$ . Задача оптимального быстрогодействия заключается в нахождении допустимого управления  $u(t)$ , переводящего точку  $x_0$  в множество  $V$  за минимальное время. Она записывается следующим образом

$$\begin{aligned} J(u) = t_1 \rightarrow \min, \\ \frac{dx}{dt} = A(t)x + B(t)u + c(t), \\ x(t_0) = x_0, \quad x(t_1) \in V, \quad u \in \mathcal{U}(\Omega). \end{aligned} \quad (16)$$

При выводе необходимых условий минимума для задачи (16) существенную роль играет

**Лемма 3.** Пусть  $C$  – выпуклое подмножество  $\mathbb{R}^q$ , не содержащее нулевую точку 0. Тогда найдутся такие числа  $a_1, \dots, a_q$ , не все равные нулю одновременно, что для любой точки  $\xi = (\xi_1, \dots, \xi_q)^T$  из  $C$  выполняется неравенство

$$a_1\xi_1 + \dots + a_q\xi_q \geq 0.$$

◀ Лемма вытекает из теорем отделимости [1], [3]. ▶

Если  $X, Y$  – два подмножества  $\mathbb{R}^n$ , то определяемое равенством

$$Z = \{z \in \mathbb{R}^n, z = y - x, y \in Y, x \in X\}$$

множество  $Z$  называют алгебраической разностью множеств  $X, Y$  и обозначают символом  $Y - X$ . Несложно показать, что алгебраическая разность двух выпуклых подмножеств  $\mathbb{R}^n$  есть выпуклое множество в  $\mathbb{R}^n$ .

**Теорема 2.** Пусть  $V$  – выпуклое подмножество  $\mathbb{R}^n$ ,  $(\hat{x}(t), \hat{u}(t), \hat{t})$  – оптимальный по быстродвижению процесс. Тогда существует такое ненулевое решение  $\hat{\psi}(t)$  сопряженной системы (4), что 1° при всех  $t$  из отрезка  $[t_0, t_1]$  справедлив принцип максимума

$$\hat{\psi}(t)B(t)u(t) = \max_{v \in \Omega} \hat{\psi}(t)B(t)v; \quad (17)$$

2° имеют место соотношения

$$\hat{\psi}(t_1)(v - \hat{x}(t_1)) \geq 0 \quad \forall v \in V, \quad \hat{\psi}(t_1) \frac{d\hat{x}}{dt}(\hat{t}_1) \geq 0. \quad (18)$$

◀ Пусть  $\tau_s$  – последовательность, сходящаяся к  $\hat{t}_1$ , причём  $\tau_s < \hat{t}_1$ . Положим  $C_s = V - \mathcal{K}(\tau_s)$ . При любом  $s$  множество выпукло и не содержит нуля. Действительно, включение  $0 \in C_s$  означает, что множество  $V$  имеет непустое пересечение с множеством  $\mathcal{K}(\tau_s)$ . Отсюда следовало бы, что  $\hat{t}_1 \leq \tau_s$ . Это противоречит определению последовательности  $\tau_s$ .

Поскольку  $0 \notin C_s$ , то согласно лемме 3 найдутся такие числа  $a_1^s, \dots, a_n^s$ , что  $(a_1^s)^2 + \dots + (a_n^s)^2 = 1$  и для любой точки  $z = (z_1, \dots, z_n)^T$  из  $C_s$  выполняется неравенство

$$\sum_{i=1}^n a_i^s z_i \geq 0. \quad (19)$$

Не нарушая общности, будем считать, что последовательность  $a_i^s$  сходится к  $a_i$  при  $s \rightarrow \infty$ ,  $i = 1, \dots, n$ . Очевидно равенство  $a_1^2 + \dots + a_n^2 = 1$ .

Определим вектор-строки  $a^s, a$  равенствами

$$a^s = (a_1^s, \dots, a_n^s), \quad a = (a_1, \dots, a_n).$$

Из неравенства (19) следует соотношение

$$a_s y \geq a_s z \quad (y \in V, z \in \mathcal{K}(\tau_s)).$$

В частности, если  $x(t)$  – траектория, соответствующая допустимому управлению  $u(t)$ , то

$$a_s y \geq a_s x(\tau_s) \quad (y \in V). \quad (20)$$

Устремляя  $s$  к  $\infty$ , приходим к неравенству

$$a y \geq a x(\hat{t}_1) \quad (y \in V). \quad (21)$$

В неравенстве (21)  $x(\hat{t}_1)$  может быть произвольным элементом  $\mathcal{K}(\hat{t}_1)$ .

Поскольку  $\hat{x}(\hat{t}_1) \in V \cap \mathcal{K}(\hat{t}_1)$ , то из (21) вытекают соотношения

$$a y \geq a \hat{x}(\hat{t}_1) \quad (y \in V), \quad (22)$$

$$a \hat{x}(\hat{t}_1) \geq a x \quad (x \in \mathcal{K}(\hat{t}_1)). \quad (23)$$

Из неравенства (23) следует, что управление  $\hat{u}(t)$  является решением простейшей задачи (1) с  $t_1 = \hat{t}_1, b(u, t) = B(t)u + c(t), p = -a$ . В силу теоремы 1 управление  $\hat{u}(t)$  удовлетворяет принципу максимума (6), если в качестве  $\hat{\psi}(t)$  взять решение обратной задачи

$$\frac{d\psi}{dt} = -\psi A(t), \quad \psi(\hat{t}_1) = a.$$

Учитывая равенство  $b(u, t) = B(t)u + c(t)$ , приходим к принципу максимума (17). Так как  $\hat{\psi}(t_1) = a$ , то из (22) следует первое из неравенств (18).

Согласно (20)  $a_s \hat{x}(\hat{t}_1) \geq a_s \hat{x}(\tau_s)$ , поэтому

$$\lim_{s \rightarrow \infty} a(\hat{x}(\hat{t}_1) - \hat{x}(\tau_s)) \geq 0.$$

Отсюда получаем неравенство

$$\lim_{s \rightarrow \infty} a \left( \frac{\hat{x}(\hat{t}_1) - \hat{x}(\tau_s)}{\hat{t}_1 - \tau_s} \right) \geq 0,$$



эквивалентное второму из неравенств (18). ►

Неравенство  $\hat{\psi}(t_1)(v - \hat{x}(t_1)) \geq 0 \forall v \in V$  называют условием трансверсальности в задаче быстродействия. Из него вытекает, например, что если целевое множество  $V$  есть полупространство  $pv \geq h, (p \neq 0)$ , то в теореме 2 можно взять  $\hat{\psi}(\hat{t}_1) = p$ . По этой причине задачи на быстродействие наиболее просты для целевого множества, совпадающего с полупространством или ограничивающей его гиперплоскостью.

**4. Достаточные условия оптимальности по быстродействию.** При некоторых дополнительных ограничениях на управляемый объект и целевое множество  $V$  принцип максимума оказывается достаточным условием оптимальности по быстродействию. Говорят, что управляемый объект (15) обладает свойством единственности, если для каждого нетривиального решения  $\hat{\psi}(t)$  сопряжённой системы (3) существует единственное управление  $\hat{u}(t)$ , удовлетворяющее принципу максимума (17). Обозримые условия, обеспечивающие свойство единственности, можно найти в [2], [4], [12].

Множество  $V \subset \mathbb{R}^n$  называют устойчивым для объекта (15), если, каковы бы ни были  $x_1 \in V$  и отрезок  $[\theta_0, \theta_1]$  ( $\theta_0 > t_0$ ), найдётся такое допустимое управление  $u(t)$ , что соответствующая этому управлению траектория  $x(t)$ , удовлетворяющая условию  $x(\theta_0) = x_1$ , при  $t > \theta_0$  целиком расположена в множестве  $V$ . Устойчивость целевого множества  $V$  означает, что, из какой бы точки этого множества мы ни начинали движение, мы сможем (надлежащим образом подбирая допустимое управление) как угодно долго оставаться в множестве  $V$ .

**Теорема 3.** Пусть управляемый объект (15) обладает свойством единственности, а целевое множество выпукло и устойчиво. Пусть  $\hat{u}(t)$  ( $t_0 \leq t \leq \hat{t}_1$ ) – допустимое управление, переводящее точку  $x_0$  в точку  $x_1 \in V$  за время  $\hat{t}_1 - t_0$ ,  $\hat{x}(t)$  – соответствующая траектория, причём  $\hat{x}(t) \notin V$  при  $t < \hat{t}_1$ . Пусть существует нетривиальное решение  $\hat{\psi}(t)$  сопряжённой системы, для которого справедлив принцип максимума (16) и условие трансверсальности.

Тогда  $(\hat{x}(t), \hat{u}(t), \hat{t}_1)$  – оптимальный по быстродействию процесс.

◀ Если процесс  $(\hat{x}(t), \hat{u}(t), \hat{t}_1)$  не оптимален по быстродействию, то существуют допустимое управление  $\tilde{u}(t)$  и соответствующая ему траектория  $\tilde{x}(t)$ , такие, что  $\tilde{x}(\theta) \in V$  при некотором  $\theta < \hat{t}_1$ . Выберем такое допустимое управление, что соответствующая ему фазовая траектория, выходящая в момент  $\theta$  из положения  $\tilde{x}(\theta)$ , будет находиться в множестве  $V$  (такое управление существует в силу устойчивости множества  $V$ ). Выбранное управление мы снова обозначим через  $\tilde{u}(t)$ , а соответствующую траекторию – через  $\tilde{x}(t)$ . Теперь  $\tilde{u}(t), \tilde{x}(t)$  определены на отрезке  $[t_0, \hat{t}_1]$ , причём  $\tilde{x}(t) \in V$  при  $\theta \leq t \leq \hat{t}_1$ . Из включения  $\hat{x}(\hat{t}_1) \in V$  вытекает в силу условия трансверсальности неравенство

$$\hat{\psi}(\hat{t}_1)(\tilde{x}(\hat{t}_1) - \hat{x}(\hat{t}_1)) \geq 0. \quad (24)$$

Поскольку управление  $\hat{u}(t)$  удовлетворяет принципу максимума (16), то оно минимизирует на  $\mathcal{U}(\Omega)$  критерий качества  $J(u) = -\hat{\psi}(\hat{t}_1)x(\hat{t}_1)$ . В силу (24) это

возможно лишь при условии  $J(\hat{u}) = J(\tilde{u})$ . Следовательно, оба управления минимизируют один и тот же специальный функционал  $J$ ; поэтому

$$\hat{\psi}(t)B(t)\hat{u}(t) = \hat{\psi}(t)B(t)\tilde{u}(t) = \max_{v \in \Omega} \hat{\psi}(t)B(t)v.$$

Так как управляемая система обладает свойством единственности, то  $\tilde{u}(t) = \hat{u}(t)$  ( $t_0 \leq t \leq t_1$ ). Отсюда вытекает равенство  $\tilde{x}(t) = \hat{x}(t)$  ( $t_0 \leq t \leq t_1$ ). Но тогда  $\hat{x}(\theta) \in V$ , что противоречит условиям теоремы. ►

Рассмотрим следующий частный случай задачи (16):

$$J(u) = t_1 \rightarrow \min,$$

$$\frac{dx_1}{dt} = x_2, \quad \frac{dx_2}{dt} = u, \quad (25)$$

$$x(0) = (\xi_1, \xi_2)^T, \quad -\alpha \leq x_1(t_1) \leq \alpha, x_2(t_1) = 0, \quad |u| \leq 1.$$

В данном примере  $n = 2$ ,  $m = 1$ ,  $\Omega = [-1, 1]$ ,  $t_0 = 0$ ,  $\alpha \geq 0$ ,

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = (0 \ 1)^T, \quad c = (0 \ 0)^T, \quad V = [-\alpha, \alpha] \times \{0\}.$$

Множество  $V$  устойчиво для управляемого объекта (25). Действительно, траектория  $x(t)$ , соответствующая управлению  $u = 0$  и удовлетворяющая условию  $x(0) \in V$ , при  $t > 0$  целиком расположена в множестве  $V$ , поскольку  $x(t) = x(\theta)$  при  $t > 0$ .

Нетривиальные решения сопряжённой системы

$$\frac{d}{dt}(\psi_1 \ \psi_2) = -(\psi_1 \ \psi_2) \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

имеют вид  $\psi_1(t) = c_1$ ,  $\psi_2(t) = d_1 - c_1 t$ , причём хотя бы одна из констант  $d_1$  и  $c_1$  отлична от нуля. Каждому нетривиальному решению  $\psi(t) = (c_1, d_1 - c_1 t)$  сопряжённой системы соответствует лишь одно управление  $u(t) = \text{sgn}(d_1 - c_1 t)$ , удовлетворяющее принципу максимума. Таким образом, управляемый объект (25) обладает свойством единственности.

Поскольку ненулевая линейная функция меняет знак не более чем один раз, то управление, оптимальное по быстродействию, является кусочно-постоянной функцией времени и может переключаться не более одного раза. Следовательно, оптимальными могут быть только четыре последовательности:

$$\{+1\}, \quad \{-1\}, \quad \{+1, -1\}, \quad \{-1, +1\}.$$

Вид оптимального управления зависит от начального положения объекта. Введем в рассмотрение множества

$$G_\Delta = \{(x_1 \ x_2)^T \in \mathbb{R}^2 : -\frac{x_2|x_2|}{2} - \alpha \leq x_1 \leq -\frac{x_2|x_2|}{2} + \alpha, \quad \Delta x_2 < 0\},$$

$$R_{\Delta} = \{(x_1, x_2)^T \in \mathbb{R}^2 : \Delta x_1 < -\Delta \frac{x_2 |x_2|}{2} - \Delta \alpha\},$$

$$\gamma_{\Delta} = \{(x_1, x_2)^T \in \mathbb{R}^2 : x_1 = -\frac{x_2 |x_2|}{2} - \Delta \alpha\},$$

где  $\Delta = \pm 1$ . Очевидно, что  $G_1 \cup G_{-1} \cup R_1 \cup R_{-1} \cup V = \mathbb{R}^2$ .

Если  $\xi = (\xi_1, \xi_2)^T \in G_1$ , то управление  $\hat{u}(t) = 1$  есть решение задачи (25). Действительно, пусть  $\hat{x}(t)$  – соответствующая этому управлению траектория. Как нетрудно видеть,  $\hat{x}(t) \notin V$  при  $t < |\xi_2|$  и  $\hat{x}(|\xi_2|) \in V$ . Положим  $\hat{\psi}(t) = (0, 1)$ . Тогда  $\hat{\psi}(t)$  – нетривиальное решение сопряжённой системы, для которого справедлив принцип максимума и условие трансверсальности. В силу теоремы 3  $\hat{u}(t) \equiv 1$  есть оптимальное по быстродействию управление для рассматриваемого случая. Совершенно аналогичные рассуждения показывают, что если  $\xi \in G_{-1}$ , то решением задачи (25) является управление  $\hat{u}(t) \equiv -1$ .

Пусть  $\xi \in R_1$ . Докажем, что в данном случае оптимальное по быстродействию управление имеет вид

$$\hat{u}(t) = \begin{cases} +1, & \text{если } t \in [0, \tau); \\ -1, & \text{если } t \in [\tau, \hat{t}_1]. \end{cases}$$

где  $\tau, \hat{t}_1$  – некоторые константы. Для отыскания констант  $\tau, \hat{t}_1$  обозначим через  $\hat{x}(t)$  соответствующую управлению  $\hat{u}(t)$  траекторию. Определим числа  $\tau, \hat{t}_1$  из условий

$$\hat{x}(\tau) \in \gamma_1, \quad \hat{x}(\hat{t}_1) = (-\alpha, 0)^T.$$

Несложные вычисления показывают, что этими условиями числа  $\tau, \hat{t}_1$  определяются однозначно, причём  $\hat{x}(t) \notin V$  при  $t < \hat{t}_1$ . Управление  $\hat{u}(t)$  удовлетворяет принципу максимума с  $\hat{\psi}(t) = (1, \tau - t)$ . Очевидно,  $\hat{\psi}(t)$  есть нетривиальное решение сопряжённой системы, для которого справедливо условие трансверсальности. Из теоремы 3 вытекает, что  $\hat{u}(t)$  – оптимальное по быстродействию управление.

В случае  $\xi \in R_{-1}$  управление, оптимальное по быстродействию, имеет вид

$$\hat{u}(t) = \begin{cases} -1, & \text{если } t \in [0, \tau); \\ +1, & \text{если } t \in [\tau, \hat{t}_1]. \end{cases}$$

где  $\tau, \hat{t}_1$  – константы, определяемые из условия

$$\hat{x}(\tau) \in \gamma_{-1}, \quad \hat{x}(\hat{t}_1) = (\alpha, 0)^T.$$

Соответствующая этому случаю функция  $\hat{\psi}(t)$  определяется равенством  $\hat{\psi}(t) = (-1, t - \tau)$ .

Приведённые выше результаты сохраняются и для  $\alpha = \infty$ . В этом случае  $R_{\Delta}, \gamma_{\Delta}$  – пустые множества,  $G_1, G_{-1}$  – нижняя и верхняя полуплоскости соответственно, оптимальное быстродействие постоянно, причём  $\hat{u}(t) = -\Delta$  для  $\xi \in G_{\Delta}$ .

**5. Задача терминального управления.** Пусть  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  – дифференцируемая функция  $n$  переменных. Рассматривается задача терминального управления

$$\begin{aligned} J(u) &= g(x(t_1)) \rightarrow \min, \\ \frac{dx}{dt} &= A(t)x + B(t)u + c(t), \\ x(t_0) &= x_0, \quad u \in \mathcal{U}(\Omega). \end{aligned} \quad (26)$$

Число  $t_1 > t_0$  и фиксировано. Как и в п. 3,  $x = (x_1, \dots, x_n)^T$  – фазовый вектор,  $u = (u_1, \dots, u_m)^T$  – управление, матрицы  $A(t), B(t), c(t)$  кусочно-непрерывны при  $t \in [t_0, t_1]$ ,  $x_0$  – фиксированный элемент  $\mathbb{R}^n$ ,  $\Omega$  – непустое выпуклое замкнутое подмножество  $\mathbb{R}^m$ ,  $\mathcal{K}(t_1)$  – множество достижимости рассматриваемой управляемой системы.

В силу леммы 2 множество  $\mathcal{K}(t_1)$  выпукло. Для изучения задачи (26) нам потребуются ещё два вспомогательных утверждения.

**Лемма 4.** *Для того чтобы пара  $(\hat{x}(t), \hat{u}(t))$  была оптимальным процессом задачи (26), необходимо и достаточно, чтобы элемент  $\hat{x}(t_1)$  являлся решением задачи*

$$g(x) \rightarrow \min, \quad x \in \mathcal{K}(t_1). \quad (27)$$

◀ Если  $(\hat{x}(t), \hat{u}(t))$  – оптимальный процесс, то  $J(u) = g(x(t_1)) \geq J(\hat{u}) = g(\hat{x}(t_1))$ , т.е.  $g(x) \geq g(\hat{x}(t_1))$  для всех  $x$  из  $\mathcal{K}(t_1)$ . Таким образом,  $\hat{x}(t_1)$  – решение задачи (27).

Обратно, пусть  $\hat{x}(t_1)$  – решение задачи (27). Это значит, что  $g(\hat{x}(t_1)) \leq g(x)$  для всех  $x$  из  $\mathcal{K}(t_1)$ . Следовательно,  $J(u) = g(x(t_1)) \geq g(\hat{x}(t_1)) = J(\hat{u})$  для любого допустимого управления  $u$ . ▶

**Лемма 5.** *Пусть  $K$  – выпуклое подмножество  $\mathbb{R}^n$ ,  $f: K \rightarrow \mathbb{R}$  – действительнозначная функция на  $K$ , дифференцируемая в точке  $z \in K$ . Для того чтобы элемент  $z$  минимизировал функцию  $f$  на множестве  $K$ , необходимо, а в случае выпуклости функции  $f$  и достаточно, чтобы точка  $z$  была решением экстремальной задачи*

$$f'(z)v \rightarrow \min, \quad v \in K. \quad (28)$$

Доказательство леммы 5 и более общих результатов приведено в приложении А.

**Теорема 4.** *Пусть  $(\hat{x}(t), \hat{u}(t))$  – оптимальный процесс задачи (26),  $\hat{\psi}(t)$  – решение обратной задачи Коши*

$$\frac{d\psi}{dt} = -\psi A(t), \quad \psi(t_1) = -g'(\hat{x}(t_1)). \quad (29)$$

Тогда для каждого  $t$  из отрезка  $[t_0, t_1]$  справедлив принцип максимума (17).

◀ В силу леммы 4  $\hat{x}(t_1)$  есть решение задачи (27). Применяя лемму 5 к функции  $f(x) = g(x)$  и  $K = \mathcal{K}(t_1)$ , получаем, что  $\hat{x}(t_1)$  минимизирует на  $\mathcal{K}(t_1)$

функцию  $v \rightarrow g'(\hat{x}(t_1))v$ . Используя вновь лемму 4, находим, что управление  $\hat{u}(t)$  есть решение следующей задачи

$$\begin{aligned} J_0(u) &= g'(\hat{x}(t_1))x(t_1) \rightarrow \min, \\ \frac{dx}{dt} &= A(t)x + B(t)u + c(t), \\ x(t_0) &= x_0, \quad u \in \mathcal{U}(\Omega). \end{aligned} \quad (30)$$

Задача (30) представляет частный случай задачи (1), возникающий при  $b(u, t) = B(t)u + c(t)$ ,  $p = g'(x(t_1))$ . Соотношения (30) являются вариантами равенств (5), а принцип максимума (17) это версия равенства (6). Теперь доказываемое утверждение вытекает из теоремы 1. ►

При дополнительных предположениях относительно функции  $g$  принцип максимума (17) становится достаточным условием оптимальности процесса.

**Теорема 5.** Пусть функция  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  выпукла. Для оптимальности процесса  $(\hat{x}(t), \hat{u}(t))$  необходимо и достаточно, чтобы управление  $\hat{u}(t)$  удовлетворяло принципу максимума (17), в котором  $\hat{\psi}(t)$  – решение обратной задачи Коши (29).

◄ Пусть функция  $\hat{u}(t)$  удовлетворяет принципу максимума (17). В силу теоремы 1  $\hat{u}$  является решением задачи (30). Отсюда согласно лемме 4 вытекает, что  $\hat{x}(t_1)$  минимизирует функцию  $v \rightarrow g'(\hat{x}(t_1))v$  на множестве  $\mathcal{K}(t_1)$ . Так как  $g$  выпуклая функция, то снова из леммы 5 следует, что  $\hat{x}(t_1)$  минимизирует функцию  $g$  на множестве  $\mathcal{K}(t_1)$ . Снова применяя лемму 4, приходим к достаточности принципа максимума. Его необходимость установлена выше. ►

**6. Принцип максимума для задачи со скользящим правым концом.** Сформулируем вариант принципа максимума для задачи

$$\begin{aligned} J(u) &= g_0(x(t_1)) \rightarrow \min, \\ \frac{dx}{dt} &= A(t)x + B(t)u + c(t), \\ x(t_0) &= x_0, \quad u \in \mathcal{U}(\Omega), \\ g_1(x(t_1)) &\leq 0, \dots, g_k(x(t_1)) \leq 0. \end{aligned} \quad (31)$$

Она отличается от рассмотренной в предшествующем пункте задачи наличием ограничений на правые концы траекторий. Функции  $g_0, g_1, \dots, g_k$  определены и непрерывно дифференцируемы на  $\mathbb{R}^n$ . Время  $t_1 > t_0$  предполагается фиксированным.

Рассуждения, аналогичные использованным при доказательстве леммы 4, приводят к следующему утверждению.

**Лемма 6.** Для того чтобы процесс  $(\hat{x}(t), \hat{u}(t))$  был оптимальным для задачи (31), необходимо и достаточно, чтобы элемент  $\hat{x}(t_1)$  являлся решением задачи

$$g_0(x) \rightarrow \min, \quad g_i(x) \leq 0 \quad (i = 1, \dots, k), \quad x \in \mathcal{K}(t_1). \quad (32)$$

Применяя к задаче (32) необходимые условия локального минимума (см. приложение А), получаем следующее утверждение.

**Теорема 6.** Пусть  $(\hat{x}(t), \hat{u}(t))$  – оптимальный процесс задачи (31). Тогда существует такой ненулевой набор  $\hat{\Lambda} = (\hat{\lambda}_0, \hat{\lambda}_1, \dots, \hat{\lambda}_k)$ , что  $\hat{\lambda}_i \geq 0$ , и если  $\hat{\psi}(t)$  – решение обратной задачи Коши

$$\frac{d\psi}{dt} = -\psi A(t), \quad \psi(t_1) = -\sum_{i=0}^k \hat{\lambda}_i g'_i(\hat{x}(t_1)),$$

то справедлив принцип максимума (17) и условия дополнительной нежёсткости  $\hat{\lambda}_i g_i(\hat{x}(t_1)) = 0$  ( $i = 1, \dots, k$ ).

Читателю предлагается доказать теорему 6 самостоятельно. Новым в ней по сравнению с теоремой 4 является наличие  $k + 1$  априори неизвестных множителей  $\hat{\lambda}_i$ . Их возникновение обусловлено дополнительными ограничениями на правые концы траекторий. Набор  $\hat{\Lambda}$  можно считать нормированным. Условие нормировки даёт ещё одно уравнение для определения неизвестных параметров  $\hat{\lambda}_i$ . При дополнительных предположениях (выпуклость функций  $g_i$ , неравенство  $\hat{\lambda}_0 > 0$ ) принцип максимума становится достаточным условием оптимальности процесса  $(\hat{x}(t), \hat{u}(t))$ .

## 2.3 Леммы

**1. Варианты правила множителей Лагранжа.** Для распространения принципа максимума на нелинейные управляемые системы нам потребуется ряд вспомогательных утверждений. В этом пункте приводятся формулировки необходимых условий локального минимума для специальных задач нелинейного программирования. Их доказательства можно найти в приложении А и во многих руководствах по конечномерной оптимизации (см. например, [19], [25]).

Пусть  $\square_N := \{\alpha = (\alpha_j), 0 \leq \alpha_j \leq 1, j = 1, \dots, n\}$  – стандартный куб с ребром 1 в пространстве  $\mathbb{R}^N$ ,  $\Phi(\alpha)$  – функция, определенная и непрерывно дифференцируемая на некоторой окрестности куба  $C_\eta = \eta \square_N$  ( $\eta > 0$ ).

**Лемма 1.** Пусть точка  $\vec{0} = (0 \dots 0)^T$  доставляет минимум функции  $\Phi$  на кубе  $C_\eta$ . Тогда

$$\frac{\partial \Phi}{\partial \alpha_j}(\vec{0}) \geq 0 \quad (j = 1, \dots, N). \quad (1)$$

Аналогичный лемме 1 результат верен для задачи на условный экстремум

$$\Phi_0(\alpha) \rightarrow \min, \quad \alpha \in C_\eta, \quad (2)$$

$$\Phi_i(\alpha) = 0 \quad (i \in I_0), \quad \Phi_i(\alpha) \leq 0 \quad (i \in I_+). \quad (3)$$

Здесь  $I_0, I_+$  – непересекающиеся конечные множества натуральных чисел (не исключается случай пустых множеств  $I_0, I_+$ ). Предполагается, что точка  $\vec{0}$  удовлетворяет ограничениям (3), а функции  $\Phi_0, \Phi_i (i \in I = I_0 \cup I_+)$  определены и непрерывно дифференцируемы на некоторой окрестности куба  $C_\eta$ .

**Лемма 2.** Пусть точка  $\vec{0}$  есть решение задачи (2), (3). Тогда найдутся числа  $\lambda_0, \lambda_i (i \in I)$ , не все равные 0 одновременно и такие, что

$$1^\circ \quad \lambda_0 \geq 0, \quad \lambda_i \geq 0, \quad \lambda_i \Phi_i(\vec{0}) = 0 \quad (i \in I_+);$$

2° для функции

$$\Phi(\alpha) = \lambda_0 \Phi_0(\alpha) + \sum_{i \in I} \lambda_i \Phi_i(\alpha) \quad (4)$$

имеют место соотношения (1).

Числа  $\lambda_0, \lambda_i (i \in I)$  называют множителями Лагранжа, а определяемую равенством (4) функцию  $\Phi$  – функцией Лагранжа для задачи (2), (3). Лемма 2 гарантирует существование такого ненулевого набора  $\Lambda = (\lambda_0, \lambda_i)$ , что точка  $\vec{0}$  удовлетворяет необходимым условиям минимума в задаче

$$\Phi(\alpha) \rightarrow \min, \quad \alpha \in Q_\eta. \quad (5)$$

Переход от задачи (2), (3) к задаче (5) выражает правило множителей Лагранжа, основная цель которого заключается в избавлении от условий связи путём введения дополнительных переменных.

Набор  $\Lambda$ , обладающий требуемыми свойствами, определяется неоднозначно. Очевидно, например, что вместе с  $\Lambda$  описанными выше свойствами обладает и набор  $\mu\Lambda$  при любом положительном числе  $\mu$ . За счёт выбора множителя  $\mu$  всегда можно добиться того, что длина набора  $\Lambda$  равнялась 1:

$$|\Lambda| = \left( \lambda_0^2 + \sum_{i \in I} \lambda_i^2 \right)^{1/2} = 1;$$

соответствующий набор  $\Lambda$  будем называть нормированным.

Если  $\lambda_0 > 0$ , то, умножая набор  $\Lambda$  на положительное число  $\mu = \lambda_0^{-1}$ , можно добиться равенства  $\lambda_0 = 1$ . В этом случае условия связи (3) будем называть регулярными. Достаточные условия регулярности можно найти в [13].

Обозначим через  $Q_\eta(\tau)$  произведение куба  $Q_\eta$  и отрезка  $[\tau - \eta, \tau + \eta]$ , т.е.

$$Q_\eta(\tau) = \{(\alpha, t) \in \mathbb{R}^N \times \mathbb{R}, \quad \alpha \in Q_\eta, \quad t \in [\tau - \eta, \tau + \eta]\}.$$

**Лемма 3.** Пусть точка  $(\vec{0}, \tau)$  является решением задачи

$$\Phi_0(\alpha, t) \rightarrow \min, \quad (\alpha, t) \in Q_\eta(\tau), \quad (6)$$

$$\Phi_i(\alpha, t) = 0 \quad (i \in I_0), \quad \Phi_i(\alpha, t) \leq 0 \quad (i \in I_+). \quad (7)$$

Пусть функции  $\Phi_0, \Phi_i (i \in I = I_0 \cup I_+)$  определены и непрерывно дифференцируемы в окрестности точки  $(\vec{0}, \tau)$ . Тогда существует такой нормированный набор  $\lambda_0, \lambda_i (i \in I)$ , что выполняются соотношения

$$\lambda_0 \geq 0, \quad \lambda_i \geq 0, \quad \lambda_i \Phi_i(\vec{0}, \tau) = 0 \quad (i \in I_+), \quad (8)$$

$$\sum_{i \in \{0\} \cup I} \lambda_i \frac{\partial \Phi_i}{\partial \alpha_j}(\vec{0}, \tau) \geq 0 \quad (j = 1, \dots, N), \quad (9)$$

$$\sum_{i \in \{0\} \cup I} \lambda_i \frac{\partial \Phi_i}{\partial t}(\vec{0}, \tau) = 0, \quad (10)$$

Как нетрудно понять, из леммы 3 вытекают утверждения лемм 1, 2.

**2. Оператор сдвига по траекториям дифференциальных уравнений.** Пусть  $F(x, t) = (F_1(x, t) \dots F_n(x, t))^T : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$  – непрерывная вектор-функция, компоненты  $F_1, \dots, F_n$  дифференцируемы по пространственным переменным и частные производные  $\partial F_i / \partial x_j(x, t)$  непрерывны по совокупности переменных. При этих предположениях справедливы классические результаты о зависимости решений задачи Коши

$$\frac{dx}{dt} = F(x, t), \quad (11)$$

$$x(\delta) = \xi \quad (12)$$

от начальных данных  $(\tau, \xi)$ . В частности, справедлива (см., например, [1])

**Лемма 4.** Пусть  $x : [\theta_0, \theta_1] \rightarrow \mathbb{R}^n$  – решение уравнения (1),  $\Gamma = \{(x(t), t) : \theta_0 \leq t \leq \theta_1\}$  – график этого решения.

Тогда существует такое  $\delta > 0$  и такое открытое в  $\mathbb{R}^{n+1}$  множество  $G \supset \Gamma$ , что для любых  $(\xi, \tau) \in G$  решение  $x(t, \tau, \xi)$  задачи Коши (11), (12) определено на  $[\theta_0 - \delta, \theta_1 + \delta]$  и является непрерывно дифференцируемой по совокупности аргументов функцией в  $[\theta_0 - \delta, \theta_1 + \delta) \times G$ .

Точка  $\xi \in \mathbb{R}^n$ , двигаясь по траекториям системы (11), за время от  $\tau$  до  $t$ , перейдёт в новую точку  $x$ . Оператор  $\mathcal{U}(t, \tau)$  перехода от  $\xi$  к  $x$  называется оператором сдвига по траекториям системы. Этот оператор определяется равенством

$$\mathcal{U}(t, \tau)\xi = x(t, \tau, \xi).$$

Область его определения зависит от  $\tau$  и  $t$ . Если выполнены предположения леммы 4 и  $|t - \theta_1| < \delta, |\tau - \theta_0| < \delta$ , то оператор  $\mathcal{U}(t, \tau)$  определён и непрерывно дифференцируем в некоторой окрестности точки  $x(\theta_0)$ .

**3. Определение игольчатых вариаций.** Ниже рассматривается нелинейный управляемый объект

$$\frac{dx}{dt} = f(x, u, t), \quad x(t_0) = x_0, \quad u \in \mathcal{U}(\Omega). \quad (13)$$

Здесь и далее  $x = (x_1, \dots, x_n)^T, u = (u_1, \dots, u_m)^T$  – фазовый и управляющий векторы соответственно,  $\Omega$  – замкнутое подмножество  $\mathbb{R}^m$ , компоненты



$f_1, \dots, f_n$  вектор-функции  $f$  дифференцируемы по фазовым переменным, причём функции  $f_i, \partial f_i / \partial x_j$  ( $i, j = 1, \dots, n$ ) непрерывны по совокупности переменных на  $\mathbb{R}^n \times \Omega \times [t_0, \infty)$ .

Для изучения объекта (13) будет использоваться семейство управляемых процессов, называемое пакетом игольчатых вариаций. Пусть  $u \in \mathcal{U}(\Omega)$ . Игольчатая вариация управления  $u$  зависит от следующих параметров:

конечный момент времени  $t_1$ ;

набор  $\vec{\tau} = (\tau_1 \dots \tau_N)$ , где  $t_0 < \tau_1 < \dots < \tau_N < t_1$ , причём среди чисел  $\tau_i$  содержатся все точки разрыва управления  $u(t)$  на  $(t_0, t_1)$ ;

набор  $\vec{\alpha} = (\alpha_1 \dots \alpha_N)$ , где  $\alpha_i$ , ( $i = 1, \dots, N$ ) – неотрицательные достаточно малые числа;

набор  $\vec{v} = (v_1 \dots v_N)$ , где  $v_i \in \Omega$  ( $i = 1, \dots, N$ ). Игольчатая вариация управления  $u$ , отвечающая наборам  $\vec{\tau}, \vec{\alpha}, \vec{v}$ , определяется равенством

$$u(t, \vec{\alpha}) = \begin{cases} u(t), & t \notin \bigcup_{i=1}^N [\tau_i, \tau_{i+1}), \quad t < t_1 \\ v_i, & t \in [\tau_i, \tau_{i+1}) \\ u(t_1 - 0), & t \geq t_1. \end{cases}$$

Очевидно,  $u(t, \vec{\alpha}) \in \mathcal{U}(\Omega)$ . В пределах этого параграфа наборы  $\vec{\tau}, \vec{v}$  фиксированы, поэтому зависимость от них не отмечается.

Числа  $\alpha_i$  предполагаются настолько малыми, что промежутки  $[\tau_i, \tau_i + \alpha_i)$  ( $i = 1, \dots, N$ ) попарно не пересекаются и принадлежат  $(t_0, t_1)$ . На этих промежутках разность  $u(t, \vec{\alpha}) - u(t)$  может быть значительной, вне их объединения при  $t < t_1$  она равна нулю.

Управлениям  $u(t)$  и  $u(t, \vec{\alpha})$  соответствуют фазовые траектории  $x(t)$  и  $x(t, \vec{\alpha})$  управляемого объекта (13). Траектория  $x(t, \vec{\alpha})$  называется игольчатой вариацией траектории  $x(t)$ . Будем считать, что  $x(t)$  определена на отрезке  $[t_0, t_1]$ . Совокупность пар  $x(t, \vec{\alpha}), u(t, \vec{\alpha})$  образует пакет игольчатых вариаций. Очевидно, что  $(x(t), u(t)) = (x(t, \vec{0}), u(t, \vec{0}))$ .

**4. Формулировка леммы о пакете иголок.** Управляемому процессу  $(x(t), u(t))$  сопоставим матрицу

$$A(t) = f'_x(t) = \left( \frac{\partial f_i}{\partial x_j}(x(t), u(t)) \right)_{i,j=1}^n.$$

Элементы  $a_{ij}(t) = \frac{\partial f_i}{\partial x_j}(x(t), u(t), t)$  этой матрицы – кусочно-непрерывные при  $t \geq t_0$  функции. В дальнейшем важную роль играет система

$$\frac{dz}{dt} = A(t)z. \quad (14)$$

Через  $z_k(t)$  ниже обозначается решение системы (14), удовлетворяющее начальному условию

$$z(\tau_k) = f(x(\tau_k), v_k, \tau_k) - f(x(\tau_k), u(\tau_k), \tau_k). \quad (15)$$

**Лемма 5. (Лемма о пакете иголок).**

1) Существует такое  $\eta > 0$ , что если  $\vec{\alpha} \in C_\eta$ , то игольчатая вариация  $x(t, \vec{\alpha})$  определена на отрезке  $[t_0, t_1 + \eta]$ .

2) Если  $\vec{\alpha} \rightarrow \vec{0}$ , то  $x(t, \vec{\alpha}) \rightarrow x(t)$  равномерно на отрезке  $[t_0, t_1]$ .

3) Отображение  $(\vec{\alpha}, t) \rightarrow x(t, \vec{\alpha})$  может быть продолжено до отображения, определённого и непрерывно дифференцируемого в некоторой окрестности точки  $(\vec{0}, t_1)$ , при этом справедливы равенства

$$\frac{\partial x}{\partial t}(t_0, \vec{0}) = f(x(t_1), u(t_1 - 0), t_1), \quad (16)$$

$$\frac{\partial x}{\partial \alpha_k}(t_1, \vec{0}) = z_k(t_1), \quad (17)$$

где  $z_k(t)$  – решение задачи (13), (14).

Доказательство леммы о пакете иголок приводится в следующем пункте. Его основу составляет специальное представление игольчатой вариации траектории. Положим для единообразия  $\tau_0 = t_0, \tau_{N+1} = t_1$ . Обозначим через  $w_i(t)$  ( $i = 0, 1, \dots, N$ ) непрерывное отображение из  $\mathbb{R}$  в  $\Omega$ , совпадающее с  $u(t)$  на интервале  $(\tau_i, \tau_{i+1})$ . Например, можно положить

$$w_i(t) = \begin{cases} u(\tau), & \text{если } t \leq \tau_i \\ u(t), & \text{если } t \in (\tau_i, \tau_{i+1}) \\ u(\tau_{i+1} - 0), & \text{если } t \geq \tau_{i+1}. \end{cases}$$

Продолжим функцию  $f$  на  $\mathbb{R}^n \times \Omega \times \mathbb{R}$ , полагая  $f(x, u, t) = f(x, u, t_0)$  при  $t < t_0$ .

Функции  $f(x, w_i(t), t), f(x, v_i, t)$  ( $i = 1, \dots, N$ ) удовлетворяют всем условиям гладкости леммы 4. Поэтому к дифференциальным уравнениям

$$\frac{dx}{dt} = f(x, w_i(t), t), \quad (18)$$

$$\frac{dx}{dt} = f(x, v_i, t) \quad (19)$$

применимы классические результаты о дифференцируемой зависимости решений от начальных данных. Введём операторы сдвига  $P_i(t, \tau), Q_i(t, \tau)$  по траекториям систем (18), (19) соответственно ( $i = 1, \dots, N$ ).

Если игольчатая вариация  $x(t, \vec{\alpha})$  траектории  $x(t)$  определена на отрезке  $[\tau_0, \tau_{k+1}]$ , то

$$x(t, \vec{\alpha}) = Q_k(t, \tau_k)x(\tau_k, \vec{\alpha}), \quad (\tau_k \leq t \leq \tau_k + \alpha_k), \quad (20)$$

$$x(t, \vec{\alpha}) = P_k(t, \tau_k + \alpha_k)x(\tau_k + \alpha_k, \vec{\alpha}), \quad (\tau_k + \alpha_k \leq t \leq \tau_{k+1}). \quad (21)$$

В частности, если траектория  $x(t, \vec{\alpha})$  определена на отрезке  $[t_0, t_1]$ , то формулы (20), (21) верны для  $k = 1, \dots, N$ .

Для доказательства равенства (20) достаточно учесть, что на отрезке  $[\tau_k, \tau_k + \alpha_k]$  функция  $x(t, \vec{\alpha})$  есть решение уравнения (19). Равенство (21) вытекает из того, что на отрезке  $[\tau_k + \alpha_k, \tau_{k+1}]$  функция  $x(t, \vec{\alpha})$  удовлетворяет уравнению (18).

**5. Доказательство леммы о пакете иголок.** Разобьём доказательство на несколько этапов.

а) Построение специальной последовательности отображений.

Пусть  $z_{2k} = z_{2k-1} = x(\tau_k)$  ( $k = 1, \dots, N$ ). Определим отображения  $T_i: \mathbb{R}^n \times \mathbb{R}^N \rightarrow \mathbb{R}^n$  ( $i = 1, \dots, 2N - 1$ ), полагая

$$T_{2k-1}(\xi, \vec{\alpha}) = Q_k(\tau_k + \alpha_k, \tau_k)\xi \quad (k = 1, \dots, N),$$

$$T_{2k}(\xi, \vec{\alpha}) = P_k(\tau_{k+1}, \tau_k + \alpha_k)\xi \quad (k = 1, \dots, N - 1).$$

Как нетрудно видеть,  $z_{j+1} = T_j(z_j, \vec{0})$  ( $j = 1, \dots, 2N - 1$ ).

Отображение  $T_i: \mathbb{R}^n \times \mathbb{R}^N \rightarrow \mathbb{R}^n$  определено и непрерывно дифференцируемо на некоторой окрестности точки  $(z_i, \vec{0})$ . Для отображения  $T_{2k-1}$  это следует из того, что в силу леммы 4 отображение  $(t, \tau, z) \rightarrow Q_k(t, \tau)z$  непрерывно дифференцируемо в некоторой окрестности точки  $(\tau_k, \tau_k, x(\tau_k))$ . Непрерывная дифференцируемость отображения  $T_{2k}$  вытекает из того, что согласно лемме 4 отображение  $(t, \tau, z) \rightarrow P_k(t, \tau)z$  является непрерывно дифференцируемым в окрестности точки  $(\tau_{k+1}, \tau_k, x(\tau_k))$ .

б) Доказательство первого утверждения. Введём последовательность отображений  $\mathcal{Y}_j: \mathbb{R}^N \rightarrow \mathbb{R}^n$ , полагая

$$\mathcal{Y}_i(\vec{\alpha}) = z_1, \quad \mathcal{Y}_{j+1}(\vec{\alpha}) = T_j(\mathcal{Y}_j(\vec{\alpha}), \vec{\alpha}) \quad (j = 1, \dots, 2N - 1).$$

Очевидно,  $\mathcal{Y}(0) = z_j$  ( $j = 1, \dots, 2N$ ). Поскольку суперпозиция непрерывно дифференцируемых отображений есть непрерывно дифференцируемое отображение, то найдётся окрестность  $W_1$  точки  $\vec{0}$ , в которой все отображения  $\mathcal{Y}_i$  ( $i = (1, \dots, 2N)$ ) являются непрерывно дифференцируемыми.

Рассмотрим отображения

$$\Phi(t, \vec{\alpha}) = P_N(t, \tau_N + \alpha_N)\mathcal{Y}_{2N}(\vec{\alpha}).$$

Очевидно, что  $\Phi(t_0, \vec{0}) = x(t_1)$ . Из леммы 4 вытекает существование такого  $\delta > 0$  и такой окрестности  $W$  точки  $\vec{0}$ , что  $W \subset W_1$  и отображение  $\Phi(t, \vec{\alpha})$  определено и непрерывно дифференцируемо на  $(t_1 - \delta, t_1 + \delta) \times W$ .

Пусть  $0 < \eta < \delta$  и число  $\eta$  настолько мало, что куб  $C_\eta = \eta \square_N$  принадлежит  $W$ . Если  $\vec{\alpha} \in C_\eta$ , то

$$\mathcal{Y}_1(\vec{\alpha}) = z_1 = x(\tau_1),$$

$$\mathcal{Y}_2(\vec{\alpha}) = Q_1(\tau_1 + \alpha_1, \tau_1)x(\tau_1).$$

Последнее равенство означает, что траектория  $x(t, \vec{\alpha})$  определена на отрезке  $[t_0, \tau_1 + \alpha_1]$  и  $\mathcal{Y}_2(\vec{\alpha}) = x(\tau_1 + \alpha_1, \vec{\alpha})$ . Из равенства

$$\mathcal{Y}_3(\vec{\alpha}) = P_1(\tau_2, \tau_1 + \alpha_1)x(\tau_1 + \alpha_1, \vec{\alpha})$$

следует, что траектория  $x(t, \vec{\alpha})$  определена на отрезке  $[t_0, \tau_2]$  и  $\mathcal{Y}_3(\vec{\alpha}) = x(\tau_2, \vec{\alpha})$ . Аналогичные рассуждения приводят к равенствам

$$\mathcal{Y}_{2k}(\vec{\alpha}) = x(\tau_k + \alpha_k, \vec{\alpha}), \quad \mathcal{Y}_{2k+1} = x(\tau_{k+1}, \vec{\alpha}), \quad (22)$$

$$\Phi(t, \vec{\alpha}) = x(t, \vec{\alpha}). \quad (23)$$

Как показывает равенство (23), при любом  $\vec{\alpha}$  из  $C_\eta$  игольчатая вариация  $x(t, \vec{\alpha})$  определена на отрезке  $[t_0, t_1 + \eta]$ . Тем самым доказано первое утверждение леммы 5.

в) Доказательство второго утверждения. Объединяя (22) с равенствами (20), (21), приходим к соотношениям

$$x(t, \vec{\alpha}) = \begin{cases} P_k(t, \tau_k + \alpha_k) \mathcal{Y}_{2k}(\vec{\alpha}), & t \in [\tau_k, \tau_{k+1}] \\ Q_k(t, \tau_k) \mathcal{Y}_{2k-1}(\vec{\alpha}), & t \in [\tau_k, \tau_k + \alpha_k], \end{cases} \quad (24)$$

в которых  $k = 1, \dots, N, \vec{\alpha} \in C_\eta$ . Так как  $\mathcal{Y}_{2k-1}(\vec{\alpha}) \rightarrow x(\tau_k)$ ,  $\mathcal{Y}_{2k}(\vec{\alpha}) \rightarrow x(\tau_k)$  при  $\vec{\alpha} \rightarrow \vec{0}$ , то из соотношений (24) и леммы 4 вытекает, что

$$x(t, \vec{\alpha}) \rightarrow x(t)$$

равномерно на отрезке  $[\tau_k, \tau_{k+1}]$  ( $k = 1, \dots, N$ ). Подобное рассуждение применимо к каждому из отрезков  $[\tau_k, \tau_{k+1}]$  ( $k = 1, \dots, N$ ). Отсюда следует второе утверждение леммы 5.

г) Обоснование равенства (16). Как отмечалось выше, отображение  $\Phi(t, \vec{\alpha})$  непрерывно дифференцируемо в окрестности точки  $(t_1, \vec{0})$ . В силу равенства (23) его можно рассматривать как продолжение отображения  $(t, \vec{\alpha}) \rightarrow x(t, \vec{\alpha})$ .

Пусть

$$\tilde{u}(t) = \begin{cases} u(t) & , t < t_1 \\ u(t_1 - 0) & , t \geq t_1. \end{cases}$$

Управление  $\tilde{u}(t)$  непрерывно в точке  $t_1$ , а траектория  $z(t)$ , соответствующая управлению  $\tilde{u}(t)$ , при  $t < t_1$  совпадает с  $x(t)$ . В частности,

$$\frac{\partial x}{\partial t}(t_1, \vec{0}) = \frac{dx}{dt}(t_1) = f(x(t_1), u(t_1 - 0), t_1).$$

Равенство (16) доказано.

д) Доказательство равенства (17). Пусть  $0 < \lambda < \eta$ ,  $u_\lambda(t) = u(t, \lambda e_k)$ , где  $e_k$  – вектор-строка, у которой  $k$ -ая компонента равна 1, а остальные компоненты равны 0;  $x_\lambda(t) = x(t, u_\lambda)$ . Поскольку  $u_\lambda(t) = u(t)$  при  $t < \tau_k$ , то  $x_\lambda(t) = x(t)$  при  $t < \tau_k$ . На отрезке  $[\tau_k, t_1]$  функция  $x_\lambda(t)$  удовлетворяет интегральному уравнению

$$x_\lambda(t) = x(\tau_k) + \int_{\tau_k}^t f(x_\lambda(s), u_\lambda(s), s) ds.$$

В частности, справедливы равенства

$$x_\lambda(\tau_k + \lambda) = x(\tau_k) + \int_{\tau_k}^{\tau_k + \lambda} f(x_\lambda(s), u_\lambda(s), s) ds, \quad (25)$$

$$x_\lambda(t) = x(\lambda + \tau_k) + \int_{\lambda + \tau_k}^t f(x_\lambda(s), u_\lambda(s), s) ds \quad (t \geq \tau_k + \lambda). \quad (26)$$

В силу (24)

$$x_\lambda(t) = P_k(t, \tau_k + \lambda) \mathcal{Y}_{2k}(\lambda e_k) \quad (\tau_k + \lambda \leq t \leq \tau_{k+1}). \quad (27)$$

Из (21) вытекает равенство

$$x_\lambda(t) = P_1(t, \tau_k + \lambda) \dots P_{k+1}(\tau_{k+2}, \tau_{k+1}) x_\lambda(\tau_{k+1}), \quad (28)$$

где  $(\tau_s \leq t \leq \tau_{s+1}; s = k+1, \dots, N)$ .

Введём в рассмотрение области

$$G_0 = \{(t, \lambda) : 0 < \lambda < \eta, \tau_k < \lambda < t < \tau_{k+1}\},$$

$$G_i = \{(t, \lambda) : 0 < \lambda < \eta, \tau_{k+i} < \lambda < t < \tau_{k+i+1}\} \quad (i = 1, \dots, N - k).$$

Из равенств (27), (28) и леммы 4 легко выводится, что функции

$$(t, \lambda) \rightarrow x_\lambda(t), \quad \frac{\partial x_\lambda(t)}{\partial t}, \quad \frac{\partial x_\lambda(t)}{\partial \lambda}$$

непрерывны в областях  $G_i$  и могут быть продолжены с сохранением непрерывности на замыкания этих областей.

Дифференцируя равенство (25) по параметру  $\lambda$ , получаем

$$\frac{\partial x_\lambda}{\partial \lambda}(\tau_k + \lambda) = f(x_\lambda(\tau_k + \lambda), v_k, \tau_k + \lambda). \quad (29)$$

Интеграл, находящийся в правой части (26), представим в виде суммы

$$\begin{aligned} & \int_{\tau_k + \lambda}^t f[x_\lambda(s), u(s), s] ds + \int_{\tau_k + \lambda}^{\tau_{k+1}} f[x_\lambda(s), u(s), s] ds + \\ & + \sum_{j=k+1}^{N-1} \int_{\tau_j}^{\tau_{j+1}} f[x_\lambda(s), u(s), s] ds + \int_{\tau_N}^t f[x_\lambda(s), u(s), s] ds. \end{aligned} \quad (30)$$

В силу отмеченных выше дифференциальных свойств отображения  $(t, \lambda) \rightarrow x_\lambda(t)$  к каждому из интегралов в правой части (30) применимо классическое правило Лейбница дифференцирования интегралов по параметру. Используя это правило и аддитивность интеграла, приходим к равенству

$$\begin{aligned} \frac{d}{d\lambda} \int_{\tau_k + \lambda}^t f[x_\lambda(s), u(s), s] ds &= -f[x_\lambda(\tau_k + \lambda), u(\tau_k + \lambda), \tau_k + \lambda] + \\ &+ \int_{\tau_k + \lambda}^t f'_x[x_\lambda(s), u(s), s] \frac{\partial x_\lambda}{\partial s}(s) ds. \end{aligned} \quad (31)$$

Положим

$$z(t) = \left. \frac{\partial x_\lambda(t)}{\partial \lambda} \right|_{\lambda=0} \quad (t \geq \tau_k).$$

Из соотношений (29), (31) вытекает уравнение

$$z(t) = f[x(\tau_k), v_k, \tau_k] - f[x(\tau_k), u(\tau_k), \tau_k] + \int_{\tau_k}^t f'_x[x(s), u(s), s] z(s) ds.$$

Это означает, что  $z(t)$  есть решение задачи Коши (14), (15). Теперь формула (17) следует из равенства

$$\frac{\partial x}{\partial \alpha_k}(t_1, 0) = \left. \frac{\partial x_\lambda(t_1)}{\partial \lambda} \right|_{\lambda=0} = z(t_1).$$

Формула (17), а вместе с ней и лемма о пакете иголок доказаны.

## 2.4 Оптимизация нелинейных систем

**1. Локально оптимальные процессы.** В этом разделе изучается задача

$$\begin{aligned} J(u) &= g(x(t_1)) \rightarrow \min, \\ \frac{dx}{dt} &= f(x, u, t), \\ x(t_0) &= x_0, \quad x(t_1) \in V, \quad u \in \mathcal{U}(\Omega) \end{aligned} \quad (1)$$

с фиксированным временем окончания процесса  $t_1$ . Рассматриваемый управляемый объект (1) удовлетворяет предположениям предшествующего раздела. Как и выше, процесс  $(x(t), u(t))$  назовём допустимым, если  $x(t)$  – траектория, соответствующая управлению  $u(t)$ , а  $x(t_1) \in V$ . Допустимый процесс  $(\hat{x}(t), \hat{u}(t))$

называется локально оптимальным для задачи (1), если найдётся такое  $\varepsilon > 0$ , что для любого допустимого процесса  $(x(t), u(t))$ , удовлетворяющего условию

$$\max_{t \in [t_0, t_1]} |x(t) - \hat{x}(t)| < \varepsilon,$$

выполняется соотношение  $J(u) \geq J(\hat{u})$ . Очевидно, оптимальный процесс является и локально оптимальным. Отсюда вытекает, в частности, что необходимое условие оптимальности процесса  $(\hat{x}(t), \hat{u}(t))$  является и необходимым условием локальной оптимальности процесса.

В общем случае локально оптимальный процесс может не быть оптимальным. В качестве примера рассмотрим следующий частный случай задачи (1):

$$J_0(u) = x_2(1) \rightarrow \min,$$

$$\frac{dx_1}{dt} = u, \quad \frac{dx_2}{dt} = (x_1^2 - x_1^4)u^2,$$

$$x_1(0) = x_1(1) = x_2(0) = 0, \quad -\infty < u(t) < \infty.$$

Процесс  $\hat{x}(t) = 0, \hat{u}(t) = 0$  является локально оптимальным. Действительно, если  $|x(t)| < 1/2$ , то

$$J_0(u) = x_2(1) = \int_0^1 (x_1^2(t) - x_1^4(t)) \left( \frac{dx_1}{dt} \right)^2 dt \geq 0 = J_0(0).$$

Поскольку

$$\inf_{\lambda} J_0(\lambda u) = -\infty$$

для любого ненулевого управления  $u$ , то в рассматриваемой задаче оптимального процесса нет.

## 2. Принцип максимума для задачи со свободным правым концом.

Ниже выводятся необходимые условия оптимальности в задаче

$$J(u) = g(x(t_1)) \rightarrow \min,$$

$$\frac{dx}{dt} = f(x, u, t), \tag{2}$$

$$x(t_0) = x_0, \quad u \in \mathcal{U}(\Omega).$$

Характерным для изучаемого случая является отсутствие ограничений на правые концы траекторий. Считаем, что в точке  $t_1$  рассматриваемые управления непрерывны слева. Функция  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  предполагается непрерывно дифференцируемой;  $g'(x) = \left( \frac{\partial g}{\partial x_1} \dots \frac{\partial g}{\partial x_n} \right)$  – вектор-строка, компоненты которой совпадают с частными производными функции  $g$ .

Для того чтобы придать необходимым условиям минимума достаточно обобщимую форму, удобно ввести функцию  $H$  переменных  $\psi_1, \dots, \psi_n, x_1, \dots, x_n$ ,

$u_1, \dots, u_m, t$ , называемую функцией Понтрягина объекта (1). Она определяется равенством

$$H(\psi, x, u, t) = \sum_{i=1}^n \psi_i f_i(x, u, t),$$

где  $(f_1 \cdots f_n)^T = f$ . Функция Понтрягина линейна по переменным  $\psi_1, \dots, \psi_n$ , непрерывно дифференцируема по переменным  $x_1, \dots, x_n$  и непрерывна по совокупности переменных.

**Теорема 1.** Пусть  $(\hat{x}(t), \hat{u}(t))$  – локально оптимальный процесс задачи (2),  $\hat{\psi}(t) = (\hat{\psi}_1(t) \dots \hat{\psi}_n(t))$  – решение обратной задачи Коши

$$\frac{d\psi_i}{dt} = -\frac{\partial H}{\partial x_i}(\psi, \hat{x}(t), t), \quad (i = 1, \dots, n) \quad (3)$$

$$\psi(t_1) = -g'[\hat{x}(t_1)]. \quad (4)$$

Тогда для любого  $t$  из отрезка  $[t_0, t_1]$  справедливо равенство

$$H(\hat{\psi}(t), \hat{x}(t), \hat{u}(t), t) = \max_{v \in \Omega} H(\hat{\psi}(t), \hat{x}(t), v, t). \quad (5)$$

◀ Пусть  $u(t, \vec{\alpha})$  – игольчатая вариация управления  $\hat{u}(t)$ , отвечающая наборам  $\vec{\alpha} = (\alpha_1 \cdots \alpha_N)$ ,  $\vec{\tau} = (\tau_1 \dots \tau_N)$ ,  $\vec{v} = (v_1 \cdots v_N)$ ,  $x(t, \vec{\alpha})$  – соответствующая игольчатая вариация траектории  $\hat{x}(t)$ . Поскольку  $x(t, \vec{\alpha}) \rightarrow \vec{x}(t)$  равномерно на  $[t_0, t_1]$  при  $\vec{\alpha} \rightarrow \vec{0}$ , а  $(\vec{x}(t), \vec{u}(t))$  – локально оптимальный процесс задачи (2), то найдётся такое  $\eta > 0$ , что

$$g[x(t, \vec{\alpha})] \geq g[\vec{x}(t_1)] \quad (\vec{\alpha} \in C_\eta). \quad (6)$$

Определим на кубе  $C_\eta$  функцию  $\Phi$  равенством  $\Phi(\vec{\alpha}) = g[x(t_1, \vec{\alpha})]$ . Неравенство (6) означает, что точка  $\vec{0}$  является решением задачи

$$\Phi(\vec{\alpha}) \rightarrow \min, \quad \vec{\alpha} \in C_\eta.$$

В силу леммы 3.1 отсюда вытекают неравенства

$$\frac{\partial \Phi}{\partial \alpha_k}(\vec{0}) \geq 0 \quad (k = 1, \dots, N). \quad (7)$$

Представим левую часть (7) в удобной для анализа форме. Очевидно,

$$\frac{\partial \Phi}{\partial \alpha_k}(\vec{0}) = g'[\hat{x}(t_1)] \frac{\partial x}{\partial \alpha_k}(t_1, \vec{0}).$$

Используя лемму об игольчатых вариациях, приходим к равенству

$$\frac{\partial \Phi}{\partial \alpha_k}(\vec{0}) = g'[\hat{x}(t_1)] z_k(t_1), \quad (8)$$



где  $z_k(t)$  – решение задачи Коши (3.14), (3.15).

Система (3) является сопряженной к системе (3.14). В силу основного свойства сопряжённых систем справедливо равенство  $\hat{\psi}(t_1)z_k(t_1) = \hat{\psi}(\tau_k)z_k(\tau_k)$ . Согласно (4)  $\hat{\psi}(t_1) = -g'[\hat{x}(t_1)]$ , поэтому из (7), (8) вытекает неравенство

$$0 \geq \hat{\psi}(\tau_k)z_k(\tau_k) = \hat{\psi}(\tau_k)[f(\hat{x}(\tau_k), v_k, \tau_k) - f(\hat{x}(\tau_k), \hat{u}(\tau_k), \tau_k)].$$

Таким образом,

$$\hat{\psi}(\tau_k)f(\hat{x}(\tau_k), \hat{u}(\tau_k), \tau_k) \geq \hat{\psi}(\tau_k)f(\hat{x}(\tau_k), v_k, \tau_k). \quad (9)$$

В предшествующих рассуждениях в качестве  $\tau_k$  можно брать произвольное число  $t$  из интервала  $(t_0, t_1)$ , а в качестве  $v_k$  – произвольный элемент из  $\Omega$ . Поэтому (9) приводит к (5) при всех  $t$  из  $(t_0, t_1)$ . При любом  $v$  из  $\Omega$  справедливы неравенства

$$\begin{aligned} H[\hat{\psi}(t_1), \hat{x}(t_1), v, t_1] &= \lim_{t \rightarrow t_1-0} H[\hat{\psi}(t), \hat{x}(t), v, t] \leq \\ &\leq \lim_{t \rightarrow t_1-0} H[\hat{\psi}(t), \hat{x}(t), \hat{u}(t), t] = H[\hat{\psi}(t_1), \hat{x}(t_1), \hat{u}(t_1), t_1]. \end{aligned}$$

Следовательно,

$$\sup_{v \in \Omega} H[\hat{\psi}(t_1), \hat{x}(t_1), v, t_1] \leq H[\hat{\psi}(t_1), \hat{x}(t_1), \hat{u}(t_1), t_1].$$

Поскольку обратное неравенство очевидно, то равенство (5) верно и при  $t = t_1$ . Случай  $t = t_0$  рассматривается аналогично. ►

Равенство (5) есть принцип максимума для задачи (2). Отметим, что уравнения, описывающие изменения фазовых и сопряжённых переменных, могут быть записаны в форме

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial \psi_i}, \quad \frac{d\psi_i}{dt} = -\frac{\partial H}{\partial x_i}, \quad (1, \dots, n) \quad (10)$$

напоминающей известные в механике уравнения Гамильтона.

**3. Обсуждение принципа максимума.** В формулировке принципа максимума имеется  $2n + m$  неизвестных функций:  $x_1, \dots, x_n, \psi_1, \dots, \psi_n, u_1, \dots, u_m$ . Определим количество соотношений для определения этих функций.

Рассмотрим прежде всего равенство (5), считая  $\Omega$  ограниченной замкнутой областью в  $\mathbb{R}^m$  с достаточно гладкой границей, а функцию  $f$  – дифференцируемой по  $u$ . Нетрудно понять, что оно даёт  $m$  соотношений между неизвестными функциями. Действительно, если  $\hat{u}(t)$  является внутренней точкой множества  $\Omega$ , то для выполнения (5) необходимо обращение в нуль частных производных

$$\left. \frac{\partial H[\hat{\psi}(t), \hat{x}(t), v, t]}{\partial v_i} \right|_{v=\hat{u}(t)},$$

что даёт  $m$  соотношений между неизвестными функциями. Если элемент  $\hat{u}(t)$  принадлежит границе множества  $\Omega$ , то количество соотношений остаётся прежним: действительно, имеем  $m - 1$  условий стационарности по касательным к  $\partial\Omega$  направлениям плюс одно условие принадлежности к границе.

Кроме соотношения (5), в принципе максимума фигурируют  $2n$  дифференциальных уравнений вида (10). Общее решение системы из  $2n$  дифференциальных уравнений зависит от  $2n$  произвольных постоянных, для определения которых можно использовать начальное условие  $x(t_0) = x_0$  и равенство (4). Поэтому естественно ожидать, что существуют лишь отдельные управляемые процессы, удовлетворяющие условиям теоремы 1. Только эти управляемые процессы могут оказаться оптимальными.

**4. Принцип максимума для задачи со скользящим правым концом.** Установим принцип максимума для задачи

$$\begin{aligned} J(u) &= g_0[x(t_1)] \rightarrow \min, \\ \frac{dx}{dt} &= f(x, u, t), \quad x(t_0) = x_0, \quad u \in \mathcal{U}(\Omega), \\ g_i[x(t_1)] &= 0 \quad (i \in I_0), \quad g_i[x(t_1)] \leq 0 \quad (i \in I_+). \end{aligned} \quad (11)$$

В рассматриваемом частном случае задачи (1.10) целевое множество  $V$  определено равенством

$$V = \{x \in \mathbb{R}^n, g_i(x) = 0 \quad (i \in I_0), \quad g_i(x) \leq 0 \quad (i \in I_+)\}.$$

Здесь и далее  $I_0, I_+$  – непересекающиеся конечные множества натуральных чисел (не исключается, что одно из этих множеств пусто);  $I = I_0 \cup I_+$ ; функции  $g_i$  ( $i \in I_0 \cup I_+$ ) определены и непрерывно дифференцируемы на пространстве  $\mathbb{R}^n$ . Ненулевой набор  $\Lambda = (\lambda_i)_{i \in \{0\} \cup I}$  назовём допустимым для точки  $x$  из  $V$ , если выполняются соотношения

$$\lambda_0 \geq 0, \quad \lambda_i \geq 0, \quad \lambda_i g_i(x) = 0 \quad (i \in I_+).$$

Допустимый набор  $\Lambda$  именуем нормированным, если его длина

$$|\Lambda| = \left( \sum_{i \in \{0\} \cup I} \lambda_i^2 \right)^{1/2} = 1.$$

**Теорема 2.** Пусть  $(\hat{x}(t), \hat{u}(t))$  – локально оптимальный процесс задачи (11). Тогда существует такой допустимый набор  $\hat{\Lambda} = (\hat{\lambda}_i)_{i \in \{0\} \cup I}$ , что если  $\hat{\psi}(t)$  – решение системы (3), удовлетворяющее условию

$$\psi(t_1) = - \sum_{i \in I_0 \cup I} \hat{\lambda}_i g'_i(\hat{x}(t_1)), \quad (12)$$

то справедливо равенство (5).

◀ Разобьём доказательство теоремы 2 на четыре этапа.

а) Сведение к конечномерной задаче. Фиксируем наборы  $\vec{\tau} = (\tau_1 \dots \tau_N)$ ,  $\vec{v} = (v_1 \dots v_N)$  и обозначим через  $u(t, \vec{\alpha})$ ,  $x(t, \vec{\alpha})$  игольчатые вариации управления  $\hat{u}(t)$  и траектории  $\hat{x}(t)$  соответственно. Рассуждая так же, как при доказательстве теоремы 1, найдём такое  $\eta > 0$ , что

$$g_0(x(t_1, \vec{\alpha})) \geq g_0(\hat{x}(t_1)),$$

если только  $\vec{\alpha} \in C_\eta$ ,  $x(t, \vec{\alpha}) \in V$ .

Определим на кубе  $C_\eta$  функции  $\Phi_i(\vec{\alpha})$  ( $i \in \{0\} \cup I$ ) равенствами

$$\Phi_i(\vec{\alpha}) = g_i[x(t_1, \vec{\alpha})].$$

Из определения числа  $\eta$  вытекает, что  $\vec{0}$  есть решение задачи

$$\Phi_0(\vec{\alpha}) \rightarrow \min, \quad \vec{\alpha} \in C_\eta \cap V.$$

К этой задаче применима лемма 3.2. Поэтому существует такой нормированный набор  $L = (l_i)_{i \in \{0\} \cup I}$ , что функция

$$\Phi = \sum_{i \in \{0\} \cup I} l_i \Phi_i$$

удовлетворяет неравенствам

$$\frac{\partial \Phi}{\partial \alpha_k}(\vec{0}) \geq 0 \quad (k = 1, \dots, N). \quad (13)$$

б) Сеточный принцип максимума. Неравенства (13) аналогичны неравенствам (7), поэтому, заменяя в доказательстве теоремы 1 функцию  $g$  на функцию

$$\sum_{i \in \{0\} \cup I} l_i g_i,$$

придём к неравенствам

$$\psi_L(\tau_k) f(\hat{x}(\tau_k), \hat{u}(\tau_k), \tau_k) \geq \psi_L(\tau_k) f(\hat{x}(\tau_k), v_k, \tau_k), \quad (14)$$

где  $k = 1, \dots, N$ ,  $\psi_L(t)$  – решение системы (3), удовлетворяющее условию

$$\psi(t_1) = - \sum_{i \in \{0\} \cup I} l_i g'_i(\hat{x}(t_1)). \quad (15)$$

Неравенство (14) аналогично неравенству (9), и может быть названо сеточным принципом максимума.

Принципиальное отличие (4) от (9) связано с тем, что набор  $L$  и соответствующая ему функция  $\psi_L$  могут быть разными для разных наборов  $\vec{\tau}, \vec{v}$ . Обозначим через  $\Lambda(\vec{\tau}, \vec{v})$  совокупность таких нормированных наборов  $L$ , для которых справедливо неравенство (14) с  $\psi_L$ , определяемой как решение задачи (3), (15).

в) Случай кратных  $\tau_k$ . Выше всюду предполагалось, что все числа  $\tau_k$  различны. Однако сеточный принцип максимума легко распространяется и на тот случай, когда некоторые из чисел  $\tau_k$  являются кратными, т.е. повторяются в наборе  $\vec{\tau} = (\tau_1 \dots \tau_N)$  более чем один раз. Действительно, для любого набора  $\vec{\tau}$  можно построить последовательность наборов  $\vec{\tau}^s$  с попарно различными компонентами и такую, что  $\vec{\tau}^s \rightarrow \vec{\tau}$ . В силу уже доказанного существует набор  $L^s$  из  $\Lambda(\vec{\tau}^s, v)$ . В частности, имеют место неравенства

$$\psi_{L^s}(\tau_{ks})f(\hat{x}(\tau_{ks}), \hat{u}(\tau_{ks}), \tau_{ks}) \geq \psi_{L^s}(\tau_{ks})f(\hat{x}(\tau_{ks}), v_k, \tau_{ks}) \quad (k = 1, \dots, N). \quad (16)$$

Без ограничения общности можно считать, что последовательность наборов  $L^s$  сходится к некоторому нормированному набору  $L$ . Переходя в (16) к пределу при  $s \rightarrow \infty$ , приходим к (14) и в рассматриваемом случае. Распространим введенное выше обозначение  $\Lambda(\vec{\tau}, \vec{v})$  и на случай наборов  $\vec{\tau}$  с повторяющимися компонентами.

г) Существование универсальных множителей. Решение обратной задачи Коши (3), (14) непрерывно (и даже линейно) зависит от набора  $L$ . Поэтому левая и правая части неравенств (14) суть непрерывные функции от  $L$ . Следовательно, множество  $\Lambda(\vec{\tau}, \vec{v})$  определяется некоторой совокупностью нестрогих неравенств, связывающих непрерывные функции. Отсюда вытекает замкнутость множеств  $\Lambda(\vec{\tau}, \vec{v})$  для любых наборов  $\vec{\tau}, \vec{v}$ .

Пусть  $(\vec{\tau}^1, \vec{v}^1), \dots, (\vec{\tau}^s, \vec{v}^s)$  – некоторые наборы, имеющие, возможно, разное число компонент,  $(\vec{\tau}, \vec{v})$  – объединение этих наборов. Как нетрудно видеть,

$$\Lambda(\vec{\tau}, \vec{v}) \subset \bigcap_{j=1}^s \Lambda(\vec{\tau}^j, \vec{v}^j),$$

поэтому любое конечное число множеств вида  $\Lambda(\vec{\tau}, \vec{v})$  имеет непустое пересечение, т.е. система множеств  $\Lambda(\vec{\tau}, \vec{v})$  центрирована. (Напомним, что система множеств  $\mathcal{M}_\gamma$ , зависящих от параметра  $\gamma$ , называется центрированной, если пересечение любого конечного числа множеств из системы  $\mathcal{M}_\gamma$  непусто.

Из топологии известно, что пересечение центрированной системы замкнутых подмножеств компакта непусто [1], [3]. В частности, существует такой нормированный набор  $\Lambda$ , что

$$\Lambda \in \bigcap_{\vec{\tau}, \vec{v}} \Lambda(\vec{\tau}, v).$$

Этот набор является искомым. Действительно, неравенство

$$\psi_{\Lambda}(\tau_k)f(\hat{x}(\tau_k), \hat{u}(\tau_k), \tau_k) \geq \psi_{\Lambda}(\tau_k)f(\hat{x}(\tau_k), v_k, \tau_k)$$

будет выполняться для любых наборов  $\vec{\tau} = (\tau_1 \dots \tau_n)$ ,  $\vec{v} = (v_1 \dots v_N)$ . ►

Основное отличие теоремы 2 от теоремы 1 заключается в появлении неизвестных множителей в условии (12), называемом, как и для линейных управляемых объектов, условием трансверсальности. Возникновение этих множителей вызвано ограничениями на правые концы траекторий. Для задачи (11) можно провести анализ полноты условий, вполне аналогичный описанному в предшествующем пункте.

## 2.5 Управляемые процессы с неизвестным временем окончания

**1. Формулировка основных теорем.** Если время окончания процесса не фиксировано, то возникает вопрос о его оптимальном выборе. Допустимый процесс  $(\hat{x}(t), \hat{u}(t), \hat{t}_1)$  называется локально оптимальным для задачи (4.1), если найдётся такое  $\varepsilon > 0$ , что для любого допустимого процесса  $(x(t), u(t), t_1)$ , удовлетворяющего условиям

$$|t_1 - \hat{t}_1| < \varepsilon, \quad |x(t) - \hat{x}(t)| < \varepsilon \quad \forall t \in [t_0, t_1] \cap [t_0, \hat{t}_1],$$

выполняется неравенство  $g[x(t_1)] \geq g[\hat{x}(\hat{t}_1)]$ .

Случай  $t_1 = \hat{t}_1$  не исключается. Поэтому если  $(\hat{x}(t), \hat{u}(t), \hat{t}_1)$  – локально оптимальный процесс для задачи (4.1) с нефиксированным временем окончания, то пара  $(\hat{x}(t), \hat{u}(t))$  есть локально оптимальный процесс для варианта задачи (4.1) с фиксированным и равным  $\hat{t}_1$  временем окончания процесса. В силу этого очевидного замечания необходимые условия оптимальности процесса, выведенные в предшествующем разделе, сохраняются и для задачи с переменным  $t_1$ . Однако к прежним необходимым условиям добавляется ещё одно, что совершенно естественно ввиду наличия дополнительного неизвестного  $t_1$ .

Приведём аналогии теорем 4.1, 4.2 для задач с переменным временем окончания процесса.

**Теорема 1.** Пусть  $(\hat{x}(t), \hat{u}(t), \hat{t}_1)$  – локально оптимальный процесс задачи (4.2),  $\hat{\psi}(t) = (\hat{\psi}_1(t) \dots \hat{\psi}_n(t))$  – решение обратной задачи Коши

$$\frac{d\psi_i}{dt} = -\frac{\partial H}{\partial x_i}[\psi(t), \hat{x}(t), \hat{u}(t), t] \quad (i = 1, \dots, n), \quad (1)$$

$$\psi(t_1) = -g'[\hat{x}(\hat{t}_1)]. \quad (2)$$

Тогда для каждого  $t$  из  $[t_0, \hat{t}_1]$  справедливо равенство

$$H[\hat{\psi}(t), \hat{x}(t), \hat{u}(t), t] = \max_{v \in \Omega} H[\hat{\psi}(t), \hat{x}(t), v, t], \quad (3)$$

причём

$$H[\hat{\psi}(t), \hat{x}(t), \hat{u}(t), t] \Big|_{t=\hat{t}_1-0} = 0. \quad (4)$$

**Теорема 2.** Пусть  $(\hat{x}(t), \hat{u}(t), \hat{t}_1)$  – локально оптимальный процесс задачи (4.11). Тогда существует такой допустимый набор  $\hat{\Lambda} = (\hat{\lambda}_i)$  ( $i \in \{0\} \cup I$ ), что если  $\hat{\psi}(t)$  – решение системы (1), удовлетворяющее граничному условию

$$\psi(t_1) = - \sum_i \hat{\lambda}_i g'_i[\hat{x}(\hat{t}_1)], \quad (5)$$

то имеют место соотношения (3), (4).

Доказательства теорем 1, 2 приводятся в следующем пункте. Изменяя значение  $\hat{u}(t)$  в точке  $\hat{t}_1$ , можно добиться равенства  $\hat{u}(\hat{t}_1) = \hat{u}(\hat{t}_1 - 0)$ . Такое переопределение не влияет на траекторию, соответствующую управлению  $\hat{u}(t)$ , поэтому ниже считаем, что управление  $\hat{u}(t)$  непрерывно слева в точке  $\hat{t}_1$ .

**2. Доказательства теорем 1, 2.** Равенство (4) есть новое дополнительное ограничение. Особенно легко оно устанавливается для задачи со свободным правым концом. Действительно, в этом случае  $g(\hat{x}(t)) \geq g(\hat{x}(\hat{t}_1))$  для  $t$  достаточно близких к  $\hat{t}_1$ . Поэтому  $\hat{t}_1$  – критическая точка функции  $g[x(t)]$ . Это влечёт за собой равенства

$$\begin{aligned} 0 &= \frac{d}{dt} g[\hat{x}(t)]|_{t=\hat{t}_1} = g'[\hat{x}(\hat{t}_1)] \frac{d\hat{x}(\hat{t}_1)}{dt} = -\hat{\psi}(\hat{t}_1) f(\hat{x}(\hat{t}_1), \hat{u}(\hat{t}_1), \hat{t}_1) = \\ &= -H[\hat{\psi}(t), \hat{x}(t), \hat{u}(t), t] \Big|_{t=\hat{t}_1}. \end{aligned}$$

Полученное равенство в рассматриваемом случае эквивалентно (4). Поскольку для задачи с переменным временем окончания процесса сохраняются необходимые условия оптимальности, установленные в предшествующем разделе, то равенство (3) следует из (4.5). Теорема 1 доказана.

Доказательство теоремы 2 проводится по той же схеме, что и доказательство теоремы 4.2. Укажем лишь на необходимые изменения в этапах а) – г).

На первом этапе вводятся функции

$$\Phi_i(\vec{\alpha}, t) = g_i[x(t, \vec{\alpha})] \quad (i \in \{0\} \cup I).$$

Здесь  $x(t, \vec{\alpha})$  – игольчатая вариация траектории  $\hat{x}(t)$ , соответствующая наборам  $\vec{\alpha}, \vec{\tau}, \vec{v}$ . Показывается, что если  $(\hat{x}(t), \hat{u}(t), \hat{t}_1)$  – локально оптимальный процесс задачи (4.11), то при некотором  $\eta > 0$  точка  $(\vec{0}, \hat{t}_1)$  есть решение задачи (3.6), (3.7).

На втором этапе используется лемма 3.3, из которой выводится сеточный принцип максимума. Именно, находится такой нормированный набор  $L = (l_i)$ , что если  $\psi_L(t)$  – решение обратной задачи Коши (1), (5), то справедливы соотношения

$$\psi_L(\tau_k) f(\hat{x}(\tau_k), \hat{u}(\tau_k), \tau_k) \geq \psi_L(\tau_k) f(\hat{x}(\tau_k), v_k, \tau_k), \quad (6)$$

$$\psi_L(t) f(\hat{x}(t), \hat{u}(t), t)|_{t=\hat{t}_1} = 0. \quad (7)$$

Соотношения (6), (7) выводятся из равенств (3.9), (3.10) (см. лемму 3.3).

Ненулевой набор  $L$ , для которого выведен сеточный принцип максимума, может зависеть от наборов  $\vec{\tau}, \vec{v}$ . Повторяя рассуждения, проведённые на этапах в), г) доказательства теоремы 4.2, можно убедиться в существовании универсального набора  $\hat{\Lambda}$ , обеспечивающего справедливость сеточного принципа максимума для произвольных наборов  $\vec{\tau}, \vec{v}$ . Если  $\psi_{\hat{\Lambda}}(t)$  – соответствующее универсальному нормированному набору  $\hat{\Lambda}$  решение задачи (1), (5), то соотношения (6), (7) верны для любых наборов  $\vec{\tau}, \vec{v}$ . Очевидно, что из соотношений (6), (7) вытекает теорема 2.

**3. Задачи с интегральным критерием качества.** В этом пункте описывается способ сведения к задаче (1.10) следующей экстремальной задачи

$$\begin{aligned} J(u) &= g_0(x(t_1)) + \int_{t_0}^{t_1} F_0(x(t), u(t), t) dt \rightarrow \min, \\ \frac{dx}{dt} &= f(x, u, t), \quad x(t_0) = x_0, \quad u \in \mathcal{U}(\Omega), \\ g_i(x(t_1)) + \int_{t_0}^{t_1} F_i(x(t), u(t), t) dt &= 0 \quad (i = 1, \dots, k). \end{aligned} \quad (8)$$

Здесь  $F_0, \dots, F_k$  – функции на  $\mathbb{R}^n \times \Omega \times [t_0, \infty)$ , причём как сами функции, так и их частные производные по пространственным переменным непрерывны по совокупности переменных.

Введём новые фазовые переменные  $y_0, y_1, \dots, y_k$ , эволюция которых описывается соотношениями

$$\frac{dy_i}{dt} = F_i(x, u, t), \quad y_i(t_0) = 0, \quad (i = 0, 1, \dots, k).$$

Задача (8) эквивалентна следующему варианту задачи (1.10):

$$\begin{aligned} J(u) &= g_0(x(t_1)) + y_0(t_1) \rightarrow \min, \\ \frac{dx}{dt} &= f(x, u, t), \quad \frac{dy}{dt} = F(x, u, t), \\ x(t_0) &= x_0, \quad y(t_0) = \vec{0}, \\ g_i(x(t_1)) + y_i(t_1) &= 0, \quad (i = 1, \dots, k), \\ u &\in \mathcal{U}(\Omega), \end{aligned} \quad (9)$$

где  $y = (y_1 \dots y_k)^T$ ,  $F = (F_1 \dots F_k)^T$ . Для доказательства совпадения множества оптимальных управлений в задачах (8), (9) достаточно заметить, что

$$y_i(t_1) = \int_{t_0}^{t_1} F_i(x(t), u(t), t) dt \quad (i = 0, 1, \dots, k).$$

Функция Понтрягина в задаче (9) имеет вид

$$\tilde{H} = \sum_{i=1}^n \psi_i f_i(x, u, t) + \sum_{i=0}^k \varphi_i F_i(x, u, t).$$

Из теоремы 2 вытекает существование такого допустимого набора  $\hat{\Lambda} = (\hat{\lambda}_i)$ , что если  $\hat{x}(t), \hat{y}(t), \hat{u}(t), t$  есть локально оптимальный процесс задачи (9),

$$(\hat{\psi}(t), \hat{\phi}(t)) = (\hat{\psi}_1(t) \dots \hat{\psi}_n(t), \hat{\varphi}_0(t) \dots \varphi_k(t))$$

– решение обратной задачи Коши

$$\frac{d\psi_i}{dt} = -\frac{\partial \tilde{H}}{\partial x_i} \quad (i = 1, \dots, n),$$

$$\frac{d\varphi_i}{dt} = -\frac{\partial \tilde{H}}{\partial t} = 0 \quad (i = 0, \dots, k),$$

$$\psi(t_1) = -\sum_{i=0}^k \hat{\lambda}_i g'_i(\hat{x}(t_1)), \quad \hat{\varphi}(t_1) = -\hat{\Lambda},$$

то

$$\tilde{H}(\hat{\psi}(t), \hat{\varphi}(t), \hat{x}(t), \hat{u}(t), t) = \max_{v \in \Omega} \tilde{H}(\hat{\psi}(t), \hat{\varphi}(t), \hat{x}(t), v, t), \quad (10)$$

$$\tilde{H}[\hat{\psi}(t), \hat{\varphi}(t), \hat{x}(t), \hat{y}(t), \hat{u}(t), t) \Big|_{t=\hat{t}_1} = 0. \quad (11)$$

Поскольку  $\frac{d\varphi_i}{dt} = 0$ , то  $\hat{\varphi}_i(t) = -\hat{\lambda}_i$ . Таким образом, вспомогательные функции  $\hat{\varphi}_0(t), \dots, \hat{\varphi}_k(t)$  постоянны и противоположны множителям Лагранжа. Соотношение (10) выражает принцип максимума для задачи (9). Равенство (11) есть дополнительное условие для определения времени окончания процесса. В задачах с фиксированным временем окончания процесса оно отсутствует.

**4. Задача о быстродействии.** Будем называть (8) задачей о быстродействии, если  $g_0 = 0, F_0 = 1, F_1 = \dots = F_k = 0$ . В этом частном случае речь идёт о минимизации времени перехода объекта из начального положения  $x_0$  на целевое множество

$$V = \{x \in \mathbb{R}^n : g_1(x) = \dots = g_k(x) = 0\}. \quad (12)$$

**Теорема 3.** Пусть  $(\hat{x}(t), \hat{u}(t), \hat{t}_1)$  – оптимальный процесс в задаче о быстродействии. Тогда найдутся такие постоянные  $\hat{\lambda}_1, \dots, \hat{\lambda}_k$ , не все равные нулю, что если  $\hat{\psi}(t)$  – решение системы (1), удовлетворяющее условию

$$\hat{\psi}(\hat{t}_1) = -\sum_{i=1}^k \hat{\lambda}_i g'_i(\hat{x}(\hat{t}_1)), \quad (13)$$



то выполняется соотношение (3) и

$$\tilde{H}(\hat{\psi}(\hat{t}_1), \hat{x}(\hat{t}_1), \hat{u}(\hat{t}_1), \hat{t}_1) \geq 0. \quad (14)$$

◀ Так как  $F_0 = 1, F_1 = \dots = F_k = 0$ , то функция Понтрягина для задачи быстродействия имеет вид

$$\tilde{H} = \sum_{i=1}^n \psi_i f_i(x, u, t) + \varphi_0.$$

Если  $\hat{\Lambda} = (\hat{\lambda}_0, \dots, \hat{\lambda}_k)$  – допустимый набор, для которого справедливы соотношения (10), (11), то хотя бы одна из постоянных  $\hat{\lambda}_1, \dots, \hat{\lambda}_k$  отлична от нуля. Действительно, в противном случае  $\hat{\psi}(t) \equiv 0$  и из (11) следовало бы, что  $\hat{\varphi}_0(\hat{t}_1) = 0$ . Как показано в предшествующем пункте,  $\hat{\lambda}_0 = -\hat{\varphi}_0(\hat{t}_1) = 0$ . Следовательно, набор  $\hat{\Lambda}$  оказывается нулевым. Полученное противоречие приводит к неравенству  $\hat{\lambda}_1^2 + \dots + \hat{\lambda}_k^2 > 0$ .

Поскольку  $\hat{\Lambda}$  – допустимый набор, то  $\hat{\lambda}_0 \geq 0$ . Поэтому  $\hat{\varphi}_0(\hat{t}_1) = -\hat{\lambda}_0 \leq 0$ . Из равенства (11) вытекает (14). Соотношение (10) влечёт (3). ▶

Остановимся на геометрической интерпретации условия трансверсальности (13). Определяемое равенством (12) множество  $V$  называют  $(n-k)$ -мерным многообразием в  $\mathbb{R}^n$ , если выполнено следующее условие: в каждой точке  $x \in V$  вектор-строки  $g'_1(x), \dots, g'_k(x)$  линейно независимы. Данное условие равносильно требованию, чтобы ранг функциональной матрицы

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_k}{\partial x_1} & \dots & \frac{\partial g_k}{\partial x_n} \end{pmatrix}$$

был максимальным (т.е. равнялся  $k$ ). Вектор  $h$  из  $\mathbb{R}^n$  называют касательным к многообразию  $V$  в точке  $x \in V$ , если существует некоторая линия  $\Gamma$  в пространстве  $\mathbb{R}^n$ , целиком лежащая на многообразии  $V$  и проходящая через точку  $x$ , причём вектор  $h$  касается линии  $\Gamma$  в точке  $x$ . Совокупность подобных векторов  $h$  образует линейное пространство, называемое касательным к многообразию  $V$  в точке  $x$ , и обозначается символом  $T_x V$ . Условие трансверсальности означает, что импульс-вектор  $\hat{\psi}(\hat{t}_1)$  ортогонален всем элементам пространства  $T_x V$  при  $x = \hat{x}(\hat{t}_1)$ . Иногда (краткости ради) говорят, что импульс-вектор ортогонален многообразию  $M$ .

**5. Обобщения принципа максимума.** Принцип максимума Понтрягина многократно обобщался в различных направлениях. Не претендуя на обзор всевозможных модификаций, усилений и приложений, остановимся здесь лишь на некоторых вариантах принципа максимума. Выше всюду предполагалось, что рассматриваемые функции  $g_i, f_j$  определены на всём пространстве  $\mathbb{R}^n$  и непрерывно дифференцируемы по фазовым переменным. Однако на самом деле основные рассуждения носили локальный характер, поэтому они без

труда переносятся на случай, когда фазовые переменные принадлежат открытому множеству, содержащему исследуемые фазовые траектории. Значительно сложнее ситуация в случае, когда множество допустимых значений фазовых переменных замкнуто в  $\mathbb{R}^n$ . В этом случае и формулировка принципа максимума, и его доказательство становятся более трудными (см., например, [10], [19], [24] и приведённую там литературу).

Принцип максимума доказан выше для кусочно-непрерывных управлений  $u(t)$  и кусочно-гладких фазовых траекторий  $x(t)$ . Незначительные изменения в доказательствах позволяют распространить его на случай измеримых ограниченных управлений  $u(t)$  и удовлетворяющих условию Липшица траекторий  $x(t)$ . Соответствующие результаты изложены, например, в [4], [12].

Привлечение класса измеримых отображений позволяет дополнить необходимые условия экстремума теоремами о разрешимости соответствующих задач оптимального управления. Например, удаётся установить замкнутость множеств достижимости как линейных, так и нелинейных управляемых систем. Применение интеграла Лебега вместо классического интеграла Римана даёт возможность использовать достаточно тонкие теоремы о предельном переходе под знаком интеграла, а также хорошо развитые методы функционального анализа. Вместе с тем часто оказывается, что оптимальные управления являются кусочно-непрерывными, а иногда даже кусочно-постоянными функциями.

## 2.6 Принцип максимума и вариационное исчисление

В этом разделе обсуждается связь принципа максимума с классическим вариационным исчислением. Чтобы избежать громоздких выкладок, подробные рассуждения проведём лишь для простейшей вариационной задачи. Принцип максимума позволяет получить для этой задачи три из четырёх известных необходимых условий минимума, а также условия в угловых точках и условия трансверсальности.

**1. Уравнения Эйлера–Лагранжа.** Пусть  $-\infty < t_0 < t_1 < \infty$ ,  $\Delta = [t_0, t_1]$ . Функцию  $x: \Delta \rightarrow \mathbb{R}$  назовём кусочно-гладкой на отрезке  $\Delta$ , если существует такой конечный набор чисел  $\vartheta_0, \vartheta_1, \dots, \vartheta_m$ , что

$$t_0 = \vartheta_0 < \vartheta_1 < \dots < \vartheta_{m-1} < \vartheta_m = t_1,$$

и сужение функции  $x$  на каждый из отрезков  $[\vartheta_{i-1}, \vartheta_i]$  ( $i = 1, \dots, m$ ) принадлежит классу  $C^1([\vartheta_{i-1}, \vartheta_i])$ . Совокупность кусочно-гладких на отрезке  $\Delta$  функций обозначим символом  $KC^1(\Delta)$ . Каждая из функций  $x$  класса  $KC^1(\Delta)$  непрерывна на отрезке  $\Delta$ . В каждой точке  $\vartheta$  интервала  $(t_0, t_1)$  существуют односторонние производные  $x'(\vartheta-0)$ ,  $x'(\vartheta+0)$ . Если при этом  $x'(\vartheta-0) \neq x'(\vartheta+0)$ , то  $\vartheta$  именуют

угловой точкой функции  $x$ ; подобных точек лишь конечное число, при этом

$$x'(\vartheta - 0) = \lim_{t \rightarrow \vartheta - 0} x'(t), \quad x'(\vartheta + 0) = \lim_{t \rightarrow \vartheta + 0} x'(t).$$

Исследуемая вариационная задача будет записываться следующим образом

$$f(x) = \int_{t_0}^{t_1} L(t, x(t), x'(t)) dt \rightarrow \min, \quad (1)$$

$$x \in KC^1(\Delta), \quad x(t_0) = a_0, \quad x(t_1) = a_1. \quad (2)$$

Здесь  $L(t, x, u)$  – функция определённая и непрерывная на  $\Delta \times \mathbb{R}^2$  вместе с частными производными  $L'_x, L'_u$  по переменным  $x, u$ ,  $a_0, a_1$  – фиксированные числа. Предположения о функции  $L(t, x, u)$ , называемой интегрантом функционала (1), будут далее меняться. Значения производной  $x'(t)$  в конечном числе точек излома функции  $x$  не влияют на величину функционала  $f$ . Определённости ради, полагаем функцию  $u(t) = x'(t)$  непрерывной справа.

Функцию  $\hat{x}(t)$ , удовлетворяющую условиям (2), назовём сильным решением задачи (1), (2), если существует такое  $\delta > 0$ , что

$$f(x) \geq f(\hat{x}) \quad (3)$$

для всех функций  $x$ , удовлетворяющих условиям (2) и неравенству

$$|x(t) - \hat{x}(t)| < \delta.$$

Понятие сильного решения вариационной задачи (1), (2) отличается от понятия слабого решения простейшей вариационной задачи. Как нетрудно видеть, сильное решение одновременно является и слабым решением той же вариационной задачи. Обратное, вообще говоря, неверно.

Сопоставим задаче (1), (2) следующую задачу терминального управления

$$J(u) = x_0(t_1) \rightarrow \min,$$

$$\frac{dx_0}{dt} = L(t, x_1, u), \quad \frac{dx_1}{dt} = u, \quad (4)$$

$$x_0(t_0) = 0, \quad x_1(t_0) = a_0, \quad x_1(t_1) - a_1 = 0, \quad u \in \mathcal{U}(\mathbb{R}).$$

Задача (4) есть вариант рассмотренной в п. 5.3 задачи оптимального управления со скользящим правым концом, возникающий при  $x = (x_0 \ x_1)^T$ ,  $g_0(x) = x_0$ ,

$$g_1(x) = x_1 - a_1, \quad f_0(t, x, u) = L(t, x_1, u), \quad f_1(t, x, u) = u, \quad \Omega = \mathbb{R}.$$

**Лемма 1.** Если  $\hat{x}(t)$  – сильное решение задачи (1), (2) и  $\hat{u}(t) = \hat{x}'(t)$ , то пара  $(\hat{x}(t), \hat{u}(t))$  – локально оптимальный процесс задачи (4).

◀ Пусть функция  $x(t)$  ( $t \in \Delta$ ) удовлетворяет условиям (2). Положим  $x_1(t) = x(t)$ ,  $u(t) = x'(t)$  и определим функцию  $x_0(t)$  равенством

$$x_0(t) = \int_{t_0}^t L(s, x(s), x'(s)) ds = \int_{t_0}^t L(s, x_1(s), u(s)) ds.$$

Как нетрудно видеть,  $(x_0(t), x_1(t), u(t))$  есть допустимый процесс для задачи (4), при этом  $J(u) = x_0(t_1)$ . Поэтому соотношение (3) влечёт за собой локальную оптимальность процесса  $(\hat{x}(t), \hat{u}(t))$  в задаче (4). ▶

Функция Понтрягина  $H$  для задачи (4) такова

$$H = \psi_0 L(t, x_1, u) + \psi_1 u.$$

Уравнения для сопряжённых переменных  $\psi_0, \psi_1$  имеют вид

$$\frac{d\psi_0}{dt} = -\frac{\partial H}{\partial x_0} = 0; \quad (5)$$

$$\frac{d\psi_1}{dt} = -\frac{\partial H}{\partial x_1} = -\psi_0 \frac{\partial L}{\partial x_1}(t, x_1, u). \quad (6)$$

Условия трансверсальности на правом конце траектории с учётом определяющих функции  $g_0, g_1$  соотношений выглядят следующим образом

$$\psi_0(t_1) = -\lambda_0, \quad \psi_1(t_1) = -\lambda_1, \quad (7)$$

причём  $\lambda_0 \geq 0$  и хотя бы один из множителей Лагранжа  $\lambda_0, \lambda_1$  отличен от нуля. В соответствии с принципом максимума существуют такие числа  $\lambda_0, \lambda_1$ , что выполняется соотношение

$$\psi_0(t)L(t, \hat{x}_1(t), \hat{u}(t)) + \psi_1(t)\hat{u}(t) = \max_{v \in \mathbb{R}} \{ \psi_0(t)L(t, \hat{x}_1(t), v) + \psi_1(t)v \}. \quad (8)$$

Так как  $\hat{u}(t)$  доставляет максимум функции

$$v \rightarrow \psi_0(t)L(t, \hat{x}_1(t), v) + \psi_1(t)v$$

на действительной прямой  $\mathbb{R}$ , то  $\hat{u}(t)$  – стационарная точка этой функции. Это приводит к равенствам

$$0 = \psi_0 L'_u + \psi_1, \quad \psi_1 = -\psi_0 L'_u. \quad (9)$$

Установим неравенство  $\lambda_0 > 0$ . В предположении противного  $\lambda_0 = 0$  и из (5)-(7) вытекает тогда, что  $\psi_0(t) \equiv 0$ . Но в силу (9) в этом случае и  $\psi_1(t) \equiv 0$ . Это

противоречит (7) – набор  $\lambda_0, \lambda_1$  должен быть ненулевым. Итак,  $\lambda_0 > 0$ , поэтому можно считать, что  $\lambda_0 = 1, \psi_0(t) \equiv -1$ . Согласно (9)  $\psi_1 = L_u$ . Равенство

$$\frac{d\psi_1}{dt} = -\psi_0 \frac{\partial L}{\partial x_1} = \frac{\partial L}{\partial x}$$

влечёт за собой соотношение

$$\frac{d}{dt} L'_u = L'_x, \quad (10)$$

совпадающее с уравнением Эйлера–Лагранжа.

Следует отметить, что соотношения (5), (10) справедливы лишь в точках непрерывности функции  $\hat{u}(t)$ , т.е. всюду за исключением конечного числа точек. Дифференциальное уравнение (10) эквивалентно интегральному равенству

$$L'_u(t, \hat{x}(t), \hat{x}'(t)) = \int_{t_0}^t L'_x(s, \hat{x}(s), \hat{x}'(s)) ds + const, \quad (11)$$

верному при всех  $t$  из отрезка  $\Delta$  – непрерывность функции  $\psi_1(t)$ .

**2. Условия Вейерштрасса и Лежандра.** Проведённый в предшествующем пункте анализ позволяет считать, что принцип максимума (8) выполнен с  $\psi_0(t) = -1, \psi(t) = L'_u(t, \hat{x}(t), \hat{x}'(t))$ . Это приводит к неравенству

$$\begin{aligned} L(t, \hat{x}(t), \hat{x}'(t)) - L'_u(t, \hat{x}(t), \hat{x}'(t)) \hat{x}'(t) &\leq \\ &\leq L(t, \hat{x}(t), v) - L'_u(t, \hat{x}(t), \hat{x}'(t)) v, \end{aligned} \quad (12)$$

в котором  $v$  – произвольный элемент из  $\mathbb{R}$ . Неравенство (12) известно как условие Вейерштрасса сильного локального минимума в простейшей вариационной задаче. Оно означает, что функция

$$v \rightarrow L(t, \hat{x}(t), v) - L'_u(t, \hat{x}(t), \hat{x}'(t)) v$$

достигает своего минимума на  $\mathbb{R}$  при  $v = \hat{x}'(t)$ . Если функция  $v \rightarrow L(t, \hat{x}(t), v)$  дважды дифференцируема в точке  $\hat{x}'(t)$ , то вторая производная этой функции в данной точке неотрицательна. Поэтому справедливо неравенство

$$L''_{v^2}(t, \hat{x}(t), v)|_{v=\hat{x}'(t)} \geq 0,$$

называемое условием Лежандра локального минимума в простейшей вариационной задаче.

Условия Вейерштрасса и Лежандра заведомо выполнены, если интегрант  $L(t, x, u)$  является выпуклой функцией по переменной  $u$ . Это вытекает из известных критериев выпуклости дифференцируемых функций. Требование выпуклости  $L$  по  $u$  именуют условием квазирегулярности. Н. Н. Боголюбовым ещё в 1928 году была доказана важная теорема, из которой следует, что для любой

простейшей вариационной задачи существует эквивалентная ей задача с квазирегулярным интегрантом. Точную формулировку теоремы Боголюбова и её обсуждение можно найти, например, в [10, с. 398–401].

Подводя итог пунктам 1, 2, можно сказать, что принцип максимума Понтрягина влечёт за собой три необходимых условия минимума для простейшей вариационной задачи: уравнение Эйлера–Лагранжа, условия Вейерштрасса и условие Лежандра. Четвёртое необходимое условие (условие Якоби) не следует непосредственно из принципа максимума.

**3. Условие в угловых точках.** Решение  $\hat{x}(t)$  задачи (1), (2) есть функция класса  $KC^1(\Delta)$ . В частности, могут существовать точки излома  $\hat{x}(t)$ , т.е. такие точки  $\vartheta$  из  $(t_0, t_1)$ , что  $\hat{x}(\vartheta - 0) \neq \hat{x}(\vartheta + 0)$ . Оказывается, левая и правая производные функции  $\hat{x}$  в точке  $\vartheta$  связаны определёнными соотношениями, называемыми условиями Вейерштрасса–Эрдмана.

Для их вывода вспомним, что импульс-функция  $\psi_1(t) = L'_u(t, \hat{x}(t), \hat{x}'(t))$  непрерывна всюду на отрезке  $I = [t_0, t_1]$ . Это приводит к первому условию Вейерштрасса–Эрдмана

$$L'_u(\vartheta, \hat{x}(\vartheta), \hat{x}'(\vartheta - 0)) = L'_u(\vartheta, \hat{x}(\vartheta), \hat{x}'(\vartheta + 0)). \quad (13)$$

Для вывода ещё одного условия воспользуемся принципом максимума (8). Поскольку  $\hat{x}'(t)$  – ограниченная функция, то  $|\hat{x}'(t)| < r \forall t \in \Delta$ . Поэтому

$$\max_{v \in \mathbb{R}} \{ \psi_0(t) L(t, \hat{x}_1(t), v) + \psi_1(t) v \} = \max_{|v| \leq r} \{ \psi_0(t) L(t, \hat{x}_1(t), v) + \psi_1(t) v \},$$

из которого вытекает непрерывность по  $t$  правой части (8). Но тогда и левая часть (8) также непрерывна на отрезке  $\Delta$ . Учитывая найденные ранее соотношения для функций  $\psi_0, \psi_1$ , приходим к равенству

$$(L(\vartheta, \hat{x}(\vartheta), v) - v L_u(\vartheta, \hat{x}(\vartheta), v))|_{v=\hat{x}'(\vartheta-0)} = (L(\vartheta, \hat{x}(\vartheta), v) - v L_u(\vartheta, \hat{x}(\vartheta), v))|_{v=\hat{x}'(\vartheta+0)},$$

именуемому вторым условием Вейерштрасса–Эрдмана.

Решения простейшей вариационной задачи, имеющие угловые точки, появляются в задачах со сколь угодно гладкими интегрантами  $L$ . В качестве примера рассмотрим задачу

$$f(x) = \int_0^2 x'^2(t) (1 - x'(t))^2 dt \rightarrow \min, \quad (14)$$

$$x \in KC^1(I), \quad x(0) = 0, \quad x(2) = 1 \quad (I = [0, 2]). \quad (15)$$

Так как интегрант  $L = u^2(1 - u)^2 \geq 0$ , то  $f(x) \geq 0$ . Поэтому если для какой-нибудь функции  $x$ , удовлетворяющей условиям (15), имеет место равенство

$f(x) = 0$ , то  $x$  — решение задачи (14), (15). Например, в качестве подобной функции  $x$  можно взять

$$x(t) = \begin{cases} t, & 0 \leq t \leq 1 \\ 1, & 1 < t \leq 2. \end{cases}$$

С другой стороны, легко видеть, что  $f(x) > 0$  для любой функции  $x$  класса  $C^1(I)$ . Действительно, подынтегральная функция обращается в нуль только при  $x = t + C_1$  или  $x = C_2$ , но линии, составленные из отрезков прямых, проходящих через точки  $(0, 0)$  и  $(2, 1)$  могут быть лишь ломаными.

Выпишем условия Вейерштрасса–Эрдмана для функционала (14). В рассматриваемом случае

$$L = u^2(1 - u)^2, \quad L'_u = 2u(u - 1)(2u - 1), \quad L - uL'_u = u^2(u - 1)(1 - 3u).$$

Если  $\vartheta$  — точка излома экстремали  $x(t)$ , то условия Вейерштрасса–Эрдмана принимают вид

$$u(u - 1)(2u - 1)|_{u=x'(\vartheta-0)} = u(u - 1)(2u - 1)|_{u=x'(\vartheta+0)},$$

$$u^2(u - 1)(1 - 3u)|_{u=x'(\vartheta-0)} = u^2(u - 1)(1 - 3u)|_{u=x'(\vartheta+0)}.$$

Пусть  $x(\vartheta - 0) \neq x(\vartheta + 0)$ . Тогда данная система имеет два решения:

$$x'(\vartheta - 0) = 0, \quad x'(\vartheta + 0) = 1$$

или

$$x'(\vartheta - 0) = 1, \quad x'(\vartheta + 0) = 0.$$

Поэтому ломанные экстремали могут состоять только из отрезков прямых, принадлежащих семействам  $x = C_1$  и  $x = t + C_2$ .

**4. Условия трансверсальности.** Выше всюду предполагалось, что граничные точки  $(t_0, a_0), (t_1, a_1)$  фиксированы. Предположим теперь, что одна из этих точек, например,  $(t_1, a_1)$  может перемещаться по кривой  $\Gamma$ , задаваемой уравнением  $g(x, t) = 0$ . Выражаясь более точно, будем считать, что множество

$$\Gamma = \{(t, x) \in \mathbb{R}^2, g(x, t) = 0\}$$

есть непустой компакт, функция  $g(x, t)$  определена и непрерывно дифференцируема в некоторой окрестности компакта  $\Gamma$ , причём

$$|g'_t(x, t)| + |g'_x(x, t)| > 0 \quad (t, x) \in \Gamma.$$

Допустимыми функциями в этом случае являются функции  $x(t)$  ( $t_0 \leq t \leq t_1$ ), удовлетворяющие предположениям

$$x \in KC^1([t_0, t_1]), \quad x(t_0) = a_0, \quad g(x(t_1), t_1) = 0. \quad (16)$$

Число  $t_1 > t_0$  в рассматриваемой ситуации не фиксировано, требуется лишь, чтобы точка  $(t_1, x(t_1))$  принадлежала кривой  $\Gamma$ .

Допустимую функцию  $z(t)$  ( $t_0 \leq t \leq T$ ) назовём локальным решением задачи (1), (16), если найдётся такое  $\delta > 0$ , что для любой допустимой функции  $x(t)$  ( $t_0 \leq t \leq t_1$ ,  $T - \delta < t_1 < T + \delta$ ), удовлетворяющей соотношению

$$|x(t) - z(t)| < \delta \forall t \in [t_0, t_1] \cap [t_0, T],$$

выполняется неравенство

$$\int_{t_0}^{t_1} L(t, x(t), x'(t)) dt \geq \int_{t_0}^T L(t, z(t), z'(t)) dt. \quad (17)$$

В неравенстве (17) не исключается случай  $t_1 = T$ , поэтому локальное решение задачи (1), (16) реализует сильный локальный минимум и для задачи (1), (2) с  $t_1 = T$ . Поэтому функция  $z(t)$  удовлетворяет всем ранее выведенным необходимым условиям сильного локального минимума. В частности, функция  $z(t)$  должна быть решением уравнения Эйлера–Лагранжа для определяемого равенством (1) функционала  $f$ .

Общее решение этого уравнения зависит от двух произвольных постоянных, для определения которых требуется иметь два условия. В случае задачи с закреплёнными границами такими условиями были равенства  $x(t_0) = a_0$ ,  $x(t_1) = a_1$ . В рассматриваемом же случае условие  $x(t_1) = a_1$  отсутствует. Для определения произвольных постоянных требуется получить ещё одно условие.

Для вывода соответствующего условия трансверсальности сопоставим задаче (1), (16) следующую задачу оптимального управления

$$\begin{aligned} J(u) = x_0(t_1) &\rightarrow \min, \\ \frac{dx_0}{dt} &= L(x_2, x_1, u), \quad \frac{dx_1}{dt} = u, \quad \frac{dx_2}{dt} = 1, \\ x_0(t_0) &= 0, \quad x_1(t_0) = a_0, \quad x_2(t_0) = t_0, \\ g(x(t_1), t_1) &= 0, \\ u &\in \mathcal{U}(\mathbb{R}). \end{aligned} \quad (18)$$

Задача (18) есть частный случай рассмотренной выше задачи оптимального управления со скользящим правым концом и нефиксированным временем окончания, возникающий при

$$\begin{aligned} x &= (x_0 \ x_1 \ x_2)^T, \quad g_0(x) = x_0, \quad g_1(x) = g(x_1, x_2), \\ f_1(t, x, u) &= L(x_2, x_1, u), \quad f_2(t, x, u) = u, \quad f_3(t, x, u) = 1. \end{aligned}$$



**Лемма 2.** Пусть  $z(t)$  ( $t_0 \leq t \leq T$ ) – локальное решение задачи (1), (16). Тогда  $(z(t), z'(t), T)$  – локально оптимальный процесс для задачи (18).

Доказательство леммы 2 не приводится ввиду его полной аналогии с доказательством леммы 1.

Функция Понтрягина  $H$  для задачи (18) такова

$$H = \psi_0 L(x_2, x_1, u) + \psi_1 u + \psi_2.$$

Будем считать, что функция  $L(t, x, u)$  определена и непрерывно дифференцируема по совокупности переменных при  $t \geq t_0$  и произвольных действительных  $x, u$ . Уравнения для сопряжённых переменных  $\psi_0, \psi_1, \psi_2$  имеют вид

$$\begin{aligned} \frac{d\psi_0}{dt} &= -\frac{\partial H}{\partial x_0} = 0, \\ \frac{d\psi_1}{dt} &= -\frac{\partial H}{\partial x_1} = -\psi_0 \frac{\partial L}{\partial x_1}, \\ \frac{d\psi_2}{dt} &= -\frac{\partial H}{\partial x_2} = -\psi_0 \frac{\partial L}{\partial x_2}. \end{aligned} \quad (19)$$

Условия на правом конце  $T$  выглядят следующим образом

$$\psi_0(T) = -\lambda_0, \quad \frac{d\psi_1}{dt} = -\lambda_1 \frac{\partial g}{\partial x_1}, \quad \frac{d\psi_2}{dt} = -\lambda_1 \frac{\partial g}{\partial x_2}. \quad (20)$$

В соответствии с доказанными выше необходимыми условиями минимума (теорема 5.2) найдутся такие числа  $\lambda_0 \geq 0$ ,  $\lambda_1 \in \mathbb{R}$ , что  $|\lambda_0| + |\lambda_1| > 0$ , и выполняются соотношения

$$\psi_0(t)L(t, z(t), z'(t)) + \psi_1(t)z'(t) = \max_{v \in \mathbb{R}} (\psi_0(t)L(t, z(t), v) + \psi_1(t)v), \quad (21)$$

$$(\psi_0(t)L(t, z(t), z'(t)) + \psi_1(t)z'(t) + \psi_2(t))|_{t=T} = 0. \quad (22)$$

Повторяя рассуждения, проведённые в п. 1., можно доказать неравенство  $\lambda_0 > 0$ , что позволяет положить  $\lambda_0 = 1$ . Это в свою очередь влечёт за собой соотношения  $\psi_0(t) \equiv -1$ ,  $\psi_1(t) = L'_u$ . Из (22) вытекает, что

$$\psi_2(T) = (L - uL_u)(T).$$

Объединяя полученные соотношения, приходим к равенствам

$$L_u = -\lambda_1 \frac{\partial g}{\partial x_1}, \quad L - uL'_u = -\lambda_1 \frac{\partial g}{\partial x_2}. \quad (23)$$

В исходных переменных равенства (23), называемые условиями трансверсальности, могут быть записаны следующим образом

$$L'_u(T, z(T), z'(T)) = -\lambda_1 \frac{\partial g}{\partial x}(z(T), T) \quad (24)$$

$$L(T, z(T), z'(T)) - z'(T)L'_u(T, z(T), z'(T)) = -\lambda_1 \frac{\partial g}{\partial t}(z(T), T). \quad (25)$$

Неизвестный параметр  $\lambda_1$  можно исключить из условий трансверсальности. Например, из (23) вытекает равенство

$$L'_u g'_t = (L - uL'_u)g'_x, \quad (26)$$

также называемое условием трансверсальности. Отметим два частных случая. Если конечная кривая  $\Gamma$  задается явным уравнением  $x = \varphi(t)$ , то можно положить  $g(x, t) = x - \varphi(t)$ . В этой ситуации  $g'_x = 1$ ,  $g'_t = -\varphi'(t)$  и условие (26) принимает вид

$$L = (u - \varphi'(T))L'_u.$$

Другой интересный случай возникает, если  $L(t, x, u) = c(t, x)\sqrt{1 + u^2}$ . Интегранты такого рода характерны для задач о преломлении света. После очевидных вычислений условие трансверсальности (26) принимает форму

$$z'(T)g'_t(z(T), T) = g'_x(z(T), T). \quad (27)$$

Условие (27) имеет прозрачный геометрический смысл. Оно означает, что кривая  $z = z(t)$  в точке  $(T, z(T))$  ортогональна кривой  $\Gamma$ .

Аналогичным образом может быть исследован случай, когда левая граничная точка  $(t_0, a_0)$  перемещается вдоль некоторой кривой  $\Gamma_0$ . Например, если  $\Gamma_0$  задаётся уравнением  $x = \psi(t)$ , то соответствующее условие трансверсальности на левом конце оптимальной траектории имеет вид

$$(L + (\psi - z')L_u)|_{t=t_0} = 0.$$

Без принципиальных изменений изложенная выше схема переносится на случай, когда  $x(t)$  – вектор-функция со значениями в конечномерном пространстве  $\mathbb{R}^n$ . Естественно считать при этом, что интегрант  $L$  есть функция на  $\Delta \times \mathbb{R}^{2n}$ ,  $a_0, a_1$  – фиксированные векторы из  $\mathbb{R}^n$ . Вместо одного уравнения Эйлера–Лагранжа в данном случае возникает система уравнений Эйлера–Лагранжа. Соответствующим образом модифицируются условия сильного минимума Вейерштрасса, условия в угловых точках и условия трансверсальности.

Лучший способ освоения данного материала (как и многих истин) – это попытаться получить какие-то более общие утверждения. Этот способ не назовёшь простым, но его эффективность многократно проверена. С целью облегчения участи потенциального читателя укажу, что богатый фактический материал можно найти в учебниках по вариационному исчислению (см., например, [7]–[11], [15], [21], [28]–[30]). Для тех, кто предпочитает справочники, будет полезен следующий пункт.

**5. Минисправочник.** Приведём формулировки основных результатов.

### 1. Уравнение Эйлера для простейшей вариационной задачи

Простейшая вариационная задача (п.в.з.) записывается следующим образом

$$f(x) = \int_{\Delta} L(t, x(t), x'(t)) dt \rightarrow \min, \quad x(t_0) = x_0, \quad x(t_1) = x_1.$$

Локальное решение  $\hat{x}$  п.в.з. удовлетворяет уравнению Эйлера

$$\frac{d}{dt} L_u = L_x.$$

В общем случае уравнение Эйлера – это дифференциальное уравнение второго порядка. Его решения называют экстремальми функционала  $f$ ; они образуют семейство функций, зависящих от двух параметров, значения которых находят из граничных условий  $x(t_0) = x_0, x(t_1) = x_1$ , где  $[t_0, t_1] = \Delta$ .

### 2. Первые интегралы уравнения Эйлера

Если функция  $L = L(t, x, u)$  не зависит явным образом от  $x$ , то справедливо равенство  $L_u(t, x'(t)) = C_1$ , называемое интегралом импульса. Если же функция  $L = L(t, x, u)$  не зависит явным образом от  $t$ , то имеет место соотношение  $x' L_u(x, x') - L = C_1$ , именуемое интегралом энергии.

### 3. Достаточные условия абсолютного минимума в п.в.з.

Если  $f$  – выпуклый функционал,  $\hat{x}$  – экстремаль функционала  $f$ , удовлетворяющая граничным условиям, то  $\hat{x}$  реализует абсолютный минимум функционала  $f$ . Например, это так, если при любом  $t$  из отрезка  $\Delta = [t_0, t_1]$  функция  $L(t, \cdot, \cdot) : \mathbb{R}^2 \rightarrow \mathbb{R}$  выпукла по совокупности переменных.

### 4. Изопериметрическая задача

Изопериметрической называют вариационную задачу вида

$$\begin{aligned} f_0(x) &= \int_{\Delta} L_0(t, x(t), x'(t)) dt \rightarrow \min, \\ f_i(x) &= \int_{\Delta} L_i(t, x(t), x'(t)) dt = C_i \quad (i = 1, \dots, m); \\ x(t_0) &= x_0, \quad x(t_1) = x_1. \end{aligned}$$

Если  $\hat{x}$  – локальное решение изопериметрической задачи, то существует такой ненулевой набор чисел  $\hat{\lambda}_0, \hat{\lambda}_1, \dots, \hat{\lambda}_m$ , что  $\hat{x}$  есть экстремаль функционала  $f = \hat{\lambda}_0 f_0 + \hat{\lambda}_1 f_1 + \dots + \hat{\lambda}_m f_m$  – правило множителей Лагранжа для изопериметрической задачи.

## 5. Уравнение Эйлера–Пуассона

Рассматривается следующее обобщение п.в.з.

$$f(x) = \int_{\Delta} L(t, x(t), x'(t), \dots, x^{(m)}(t)) dt \rightarrow \min,$$

$$x^j(t_0) = x_{0j}, \quad x^j(t_1) = x_{1j} \quad (j = 0, 1, \dots, m-1).$$

Локальное решение  $\hat{x}$  этой задачи удовлетворяет дифференциальному уравнению

$$\sum_{k=0}^m (-1)^k \frac{d^k}{dt^k} L'_{u_k}(t, x(t), x'(t), \dots, x^{(m)}(t)) = 0,$$

называемому уравнением Эйлера – Пуассона. В общем случае порядок этого дифференциального уравнения равен  $2m$ .

## 6. Задача Больца

В задаче Больца

$$f(x) = \int_{\Delta} L(t, x(t), x'(t)) dt + g[x(t_0), x(t_1)] \rightarrow \min$$

отсутствуют априорные ограничения на значения функций  $x$  в точках  $t_0$  и  $t_1$ , вместе с тем минимизируемый функционал наряду с интегральным критерием качества содержит терминальную составляющую  $-g[x(t_0), x(t_1)]$ . Локальное решение этой задачи удовлетворяет уравнению Эйлера и двум естественным краевым условиям

$$\left( \frac{\partial L}{\partial u}(t, x(t), x'(t)) \right) \Big|_{t=t_1} + g_{z_1}[x(t_0), x(t_1)] = 0,$$

$$\left( \frac{\partial L}{\partial u}(t, x(t), x'(t)) \right) \Big|_{t=t_0} = g_{z_0}[x(t_0), x(t_1)].$$

В рассмотренном случае числа  $t_0, t_1$  фиксированы. Имеют смысл варианты задачи Больца, в которых начальная и конечная точки  $(t_0, x(t_0)), (t_1, x(t_1))$  принадлежат некоторым подмножествам  $M_0, M_1$  плоскости  $(t, x)$ ; чаще всего  $M_0, M_1$  – гладкие кривые, и речь идёт об оптимальном переходе с начальной кривой на конечную. В этой ситуации появляются так называемые условия трансверсальности, обобщающие естественные краевые условия; обсуждение условий трансверсальности можно найти в предшествующем пункте этого раздела. Полезно иметь перед глазами простейшую задачу такого рода: найти кратчайший путь перехода с одной кривой на другую.

**7. Принцип Ферма.** Пусть  $v = v(x, y, z)$  – скорость распространения луча света в неоднородной среде,  $\Gamma$  – траектория луча света. Время прохождения лучом света траектории  $\Gamma$  называют оптической длиной пути. Принцип Ферма утверждает: *из всех траекторий, соединяющих фиксированные точки, свет выбирает ту, оптическая длина которой минимальна.* Вариационная задача Лопиталя

$$J(y) = \int_{x_0}^{x_1} \frac{\sqrt{1 + y'^2(x)}}{v[y(x)]} dx \rightarrow \min, \quad y(x_0) = y_0, \quad y(x_1) = y_1$$

возникла в оптике (начало вариационного исчисления  $\approx 1696$  г).

**8. Принцип максимума для задачи со свободным правым концом.** Сформулируем необходимые условия оптимальности в задаче

$$\begin{aligned} J(u) &= g(x(t_1)) \rightarrow \min, \\ \frac{dx}{dt} &= f(x, u, t), \\ x(t_0) &= x_0, \quad u \in \mathcal{U}(\Omega). \end{aligned} \quad (1)$$

Характерным для изучаемого случая является отсутствие ограничений на правые концы траекторий. Функция Понтрягина объекта (1) определяется равенством

$$H(\psi, x, u, t) = \sum_{i=1}^n \psi_i f_i(x, u, t),$$

где  $(f_1 \cdots f_n)^T = f$ .

**Теорема 1.** Пусть  $(\hat{x}(t), \hat{u}(t))$  – локально оптимальный процесс задачи (1),  $\hat{\psi}(t) = (\hat{\psi}_1(t) \dots \hat{\psi}_n(t))$  – решение обратной задачи Коши

$$\frac{d\psi_i}{dt} = -\frac{\partial H}{\partial x_i}(\psi, \hat{x}(t), t) \quad (i = 1, \dots, n), \quad (2)$$

$$\psi(t_1) = -g'[\hat{x}(t_1)]. \quad (3)$$

Тогда для любого  $t$  из отрезка  $[t_0, t_1]$  справедливо равенство

$$H(\hat{\psi}(t), \hat{x}(t), \hat{u}(t), t) = \max_{v \in \Omega} H(\hat{\psi}(t), \hat{x}(t), v, t). \quad (4)$$

**9. Принцип максимума для задачи со скользящим правым концом.** Приведём принцип максимума для задачи

$$J(u) = g_0[x(t_1)] \rightarrow \min,$$

$$\frac{dx}{dt} = f(x, u, t), \quad x(t_0) = x_0, \quad u \in \mathcal{U}(\Omega), \quad (5)$$

$$g_i[x(t_1)] = 0 \quad (i \in I_0), \quad g_i[x(t_1)] \leq 0 \quad (i \in I_+).$$

Здесь  $I_0, I_+$  – непересекающиеся конечные множества натуральных чисел (не исключается, что одно из этих множеств пусто);  $I = I_0 \cup I_+$ . Ненулевой набор  $\Lambda = (\lambda_i)_{i \in \{0\} \cup I}$  называют допустимым для точки  $x$ , если выполняются соотношения

$$\lambda_0 \geq 0, \quad \lambda_i \geq 0, \quad \lambda_i g_i(x) = 0 \quad (i \in I_+).$$

**Теорема 2.** Пусть  $(\hat{x}(t), \hat{u}(t))$  – локально оптимальный процесс задачи (11). Тогда существует такой допустимый для точки  $\hat{x}(t_1)$  набор  $\hat{\Lambda} = (\hat{\lambda}_i)_{i \in \{0\} \cup I}$ , что если  $\hat{\psi}(t)$  – решение системы (2), удовлетворяющее условию

$$\psi(t_1) = - \sum_{i \in \{0\} \cup I} \hat{\lambda}_i g'_i(\hat{x}(t_1)), \quad (6)$$

то справедливо равенство (4).

### 10. Принцип максимума для задачи на быстродействие.

Рассматривается задача

$$\begin{aligned} J(u) = t_1 \rightarrow \min, \\ \frac{dx}{dt} = f(x, u, t), \quad x(t_0) = x_0, \quad u \in \mathcal{U}(\Omega), \\ g_i(x(t_1)) = 0 \quad (i = 1, \dots, k). \end{aligned}$$

В этом частном случае речь идёт о минимизации времени перехода объекта из начального положения  $x_0$  на целевое множество

$$V = \{x \in \mathbb{R}^n : g_1(x) = \dots = g_k(x) = 0\}.$$

**Теорема 3.** Пусть  $(\hat{x}(t), \hat{u}(t), \hat{t}_1)$  – оптимальный процесс в задаче о быстродействии. Тогда найдутся такие постоянные  $\hat{\lambda}_1, \dots, \hat{\lambda}_k$ , не все равные нулю, что если  $\hat{\psi}(t)$  – решение системы (2), удовлетворяющее условию

$$\hat{\psi}(\hat{t}_1) = - \sum_{i=1}^k \hat{\lambda}_i g'_i(\hat{x}(\hat{t}_1)), \quad (7)$$

то выполняется соотношение (4) и

$$\tilde{H}(\hat{\psi}(\hat{t}_1), \hat{x}(\hat{t}_1), \hat{u}(\hat{t}_1), \hat{t}_1) \geq 0.$$

Соотношения (3), (6), (7) называют условием трансверсальности в соответствующих задачах. В задаче на быстродействие это условие имеет простой геометрический смысл. Оно означает, что вектор-импульс  $\hat{\psi}$  в точке  $\hat{t}_1$  ортогонален целевому многообразию  $V$ .

# Приложение А

## Правило множителей Лагранжа

**1. Условия минимума на выпуклом множестве.** Ниже рассматривается задача минимизации функции на выпуклом подмножестве пространства  $\mathbb{R}^n$ . Результаты приложения неоднократно использовались ранее для вывода условий минимума в задачах вариационного исчисления и оптимального управления. В этом пункте формулируются простейшие утверждения.

Линейным функционалом на пространстве  $\mathbb{R}^n$  называется функция  $l: \mathbb{R}^n \rightarrow \mathbb{R}$ , обладающая свойствами

$$l(x + y) = l(x) + l(y), \quad l(\lambda x) = \lambda l(x) \quad \forall x, y \in \mathbb{R}^n, \quad \lambda \in \mathbb{R}.$$

Любая линейная функция на  $\mathbb{R}^n$  допускает представление  $l(x) = px = p_1 x_1 + \dots + p_n x_n$ , в котором  $x_i$  – компоненты вектор-столбца  $x$ , а  $p_i$  – компоненты вектор-строки  $p$ . Это позволяет идентифицировать пространство линейных функционалов на  $\mathbb{R}^n$  с линейным пространством  $\mathbb{R}_n$  вектор-строк  $p = (p_1 \dots p_n)$ , изоморфным пространству  $\mathbb{R}^n$ . В  $\mathbb{R}_n$  стандартным образом вводится структура евклидова пространства.

Пусть  $Q$  – выпуклое множество в  $\mathbb{R}^n$ ,  $z \in Q$ . Линейный функционал  $l$  называют опорным к множеству  $Q$  в точке  $z$ , если  $l(z) \geq l(v) \forall v \in Q$ . Совокупность опорных функционалов обозначим символом  $N_Q(z)$  и назовём конусом опорных функционалов к множеству  $Q$  в точке  $z$ . Очевидно включение  $0 \in N_Q(z)$ . Если  $z \in \overset{\circ}{Q}$ , то  $N_Q(z) = \{0\}$ . Действительно, производная в точке локального минимума равна 0.

Для точки  $z$ , принадлежащей границе множества  $Q$ , конус  $N_Q(z)$  может быть достаточно сложно устроенным множеством. Отметим некоторые свойства конуса опорных функционалов.

1°. Если  $l_1, l_2 \in N_Q(z)$ , то  $l_1 + l_2 \in N_Q(z)$ .

2°. Если  $l \in N_Q(z)$ ,  $\lambda \geq 0$ , то  $\lambda l \in N_Q(z)$ .

3°. Опорный конус есть замкнутое множество.

4°. Если  $B(z, R)$  – замкнутый шар радиуса  $R > 0$  с центром в точке  $z$ ,  $Q_1 = Q \cap B(z, R)$ , то  $N_Q(z) = N_{Q_1}(z)$  – свойство локальности конуса опорных функционалов.

Способы нахождения конуса опорных функционалов для достаточно типичных множеств рассматриваются в третьем пункте. В этом пункте обсудим роль этого понятия в экстремальных задачах.

Пусть  $g$  – числовая функция на открытом выпуклом множестве  $\mathcal{V} \subset \mathbb{R}^n$ ,  $Q$  – выпуклое подмножество  $\mathcal{V}$ . Рассматривается задача

$$g(x) \rightarrow \min, \quad x \in Q. \quad (1)$$

Стандартным образом вводятся понятия решения и локального решения задачи (1). Имеют место следующие утверждения.

**Теорема 1.** Пусть  $z$  – локальное решение задачи (1) и функция  $g$  дифференцируема в этой точке. Тогда  $-g'(z) \in N_Q(z)$  или в эквивалентной (и чаще используемой) форме

$$0 \in g'(z) + N_Q(z). \quad (2)$$

◀ Фиксируем  $v$  из  $Q$  и введём вспомогательную функцию  $\varphi(t) = g(z + t(v - z))$  ( $0 \leq t \leq 1$ ). Поскольку множество  $Q$  выпукло, то  $z + t(v - z) \in Q$ . При достаточно малых положительных  $t$  справедливы соотношения

$$\varphi(t) = g(z + t(v - z)) \geq g(z) = \varphi(0), \quad \frac{\varphi(t) - \varphi(0)}{t} \geq 0.$$

Отсюда вытекают неравенства

$$0 \leq \varphi'(0) = g'(z)(v - z) \quad \forall v \in Q.$$

Последнее неравенство влечёт за собой включение (2). ▶

**Теорема 2.** Пусть  $g$  – выпуклая на множестве  $Q$  и дифференцируемая в точке  $z \in Q$  функция. Если справедливо соотношение (2), то  $z$  – решение задачи (1).

◀ Из выпуклости функции  $g$  и её дифференцируемости в точке  $z$  вытекает неравенство

$$g(v) \geq g(z) + g'(z)(v - z) \quad \forall v \in Q.$$

Включение (2) эквивалентно соотношению

$$g'(z)(v - z) \geq 0 \quad \forall v \in Q.$$

Объединяя два последних неравенства, приходим к доказываемому утверждению. ▶

В случае  $z \in \overset{\circ}{Q}$  условие (2) сводится к условию  $f'(z) = 0$ . Расшифровка условия (2) для граничной точки  $z$  не столь проста. Здесь требуется информация об устройстве множества  $Q$  вблизи точки  $z$ .

**2. Основная теорема.** Результаты предшествующего пункта могут быть распространены на более сложную задачу с дополнительными ограничениями типа равенств и неравенств

$$g_0(x) \rightarrow \min, \quad x \in Q, \quad (3)$$



$$g_i(x) = 0 \quad (i \in I_0), \quad (4)$$

$$g_i(x) \leq 0 \quad (i \in I_+). \quad (5)$$

Здесь  $I_0, I_+$  – непересекающиеся множества натуральных чисел (не исключается случай, когда одно из этих множеств пусто), функции  $g_i$  – определены на некоторой выпуклой окрестности  $\mathcal{V}$  множества  $Q$ . Для формулировки аналогов теорем 1, 2 введём некоторые понятия и обозначения. Ниже  $I = I_0 \cup I_+, I_1 = \{0\} \cup I$ . Ненулевой набор  $\Lambda = (\lambda_i)_{i \in I_1}$  назовём допустимым, если

$$\lambda_i \geq 0 \quad (i \in \{0\} \cup I_+).$$

Допустимый набор  $\Lambda$  будем именовать нормированным, если его длина

$$|\Lambda| \stackrel{\text{def}}{=} \left( \sum_{i \in I_1} \lambda_i^2 \right)^{1/2} = 1.$$

**Теорема 3.** Пусть  $z$  – решение задачи (3)-(5), функции  $g_i$  ( $i \in I_1$ ) непрерывно дифференцируемы в окрестности точки  $z$ . Тогда существует такой допустимый набор  $\Lambda = (\lambda_i)_{i \in I_1}$ , что

1° функционал

$$l = - \sum_{i \in I_1} \lambda_i g'_i(z)$$

является опорным к множеству  $Q$  в точке  $z$ ;

2° выполнены условия дополнительной нежёсткости

$$\lambda_i g_i(z) = 0 \quad (i \in I_+). \quad (6)$$

◀ Разобьём доказательство теоремы на два этапа. 1-й этап:  $Q$  – компактное выпуклое множество. Не нарушая общности, будем считать, что функции  $g_i$  дифференцируемы в каждой точке множества  $Q$  и производные  $g'_i$  непрерывны на  $Q$ . Выполнения этого требования можно добиться, переходя в случае необходимости от множества  $Q$  к множеству  $Q_1 = Q \cap B(z, R)$  с достаточно малым положительным числом  $R$  – свойство 4 локальности конуса опорных функционалов.

Введём в рассмотрение функцию

$$h(t) = \begin{cases} \frac{t^2}{2}, & (t \geq 0) \\ 0, & (t < 0). \end{cases}$$

Функция  $h(t)$  дифференцируема и

$$h'(t) = \begin{cases} t, & (t \geq 0) \\ 0, & (t < 0). \end{cases}$$

Производная  $h'(t)$  непрерывна на всей оси. Положим

$$\Phi(x) = \frac{1}{2} \sum_{i \in I_0} g_i^2(x) + \sum_{i \in I_+} h(g_i(x)).$$

Функция  $\Phi$  определена на множестве  $\mathcal{V}$ . Как нетрудно видеть, равенство  $\Phi(x) = 0$  эквивалентно соотношениям (4), (5). Функция  $\Phi$  дифференцируема в каждой точке  $x \in Q$ , и производная  $\Phi'(x)$  непрерывна.

Каждому натуральному числу  $\nu$  поставим в соответствие функцию

$$\Phi_\nu(x) \stackrel{def}{=} g_0(x) + \nu\Phi(x) + \frac{|x - z|^2}{2}$$

и экстремальную задачу

$$\Phi_\nu(x) \rightarrow \min, \quad x \in Q. \quad (7)$$

Из теоремы Вейерштрасса вытекает существование решения  $x^\nu$  задачи (7). Имеет место неравенство  $\Phi_\nu(x^\nu) \leq \Phi_\nu(z)$  или более подробно

$$g_0(x_\nu) + \nu\Phi(x^\nu) + \frac{|x^\nu - z|^2}{2} \leq g_0(z). \quad (8)$$

В частности, из (8) вытекает оценка  $\nu\Phi(x^\nu) \leq g_0(z) - g_0(x^\nu)$  и равенство

$$\lim_{\nu \rightarrow \infty} \Phi(x^\nu) = 0.$$

Последовательность  $x^\nu$  принадлежит  $Q$ . Если  $\bar{z}$  – предельная точка этой последовательности, то  $\bar{z} \in Q$  и  $\Phi(\bar{z}) = 0$ . Из (8) следует неравенство

$$g_0(\bar{z}) + \frac{|\bar{z} - z|^2}{2} \leq g_0(z).$$

С другой стороны, поскольку  $z$  – решение задачи (3)-(5), то  $g_0(z) \leq g_0(\bar{z})$ , поэтому  $\bar{z} = z$ . Всякая предельная точка последовательности  $x^\nu$  совпадает с  $z$ . Это означает, что  $x^\nu \rightarrow z$ .

Так как  $x^\nu$  – решение задачи (7), то  $-\Phi'_\nu(x^\nu) \in N_Q(x^\nu)$ . Более подробно данное включение может быть записано в виде

$$-\sum_{i \in I_1} \mu_i^\nu g'_i(x^\nu) - (x^\nu - z)^T \in N_Q(x^\nu), \quad (9)$$

где

$$\mu_0^\nu = 1, \quad \mu_i^\nu = g_i(x^\nu) \quad (i \in I_0), \quad \mu_i^\nu = h'(g_i(x^\nu)) \quad (i \in I_+). \quad (10)$$

Положим

$$K_\nu = \left( \sum_{i \in I_1} (\mu_i^\nu)^2 \right)^{1/2}.$$

Очевидно, что  $K_\nu \geq 1$ . Разделив (9) на  $K_\nu$ , приходим к включению

$$-\sum_{i \in I_1} \lambda_i^\nu g'_i(x^\nu) - \frac{1}{K_\nu} (x^\nu - z)^T \in N_Q(x^\nu), \quad (11)$$

в котором

$$\lambda_i^\nu = \frac{\mu_i^\nu}{K_\nu}, \quad \sum_{i \in I_1} (\lambda_i^\nu)^2 = 1. \quad (12)$$

Каждая из последовательностей  $\lambda_i^\nu$ , ( $i \in I_1, \nu = 1, 2, \dots$ ) ограничена, поэтому, не нарушая общности, можно считать, что  $\lambda_i^\nu \rightarrow \lambda_i \forall i$  при  $\nu \rightarrow \infty$ . Из (12) вытекает равенство

$$\sum_{i \in I_1} \lambda_i^2 = 1,$$

в силу которого хотя бы одно из чисел  $\lambda_i$  отлично от нуля.

Переходя в (11) к пределу при  $\nu \rightarrow \infty$ , получаем включение

$$-\sum_{i \in I_1} \lambda_i g'_i(z) \in N_Q(z). \quad (13)$$

Если  $g_i(z) = 0$ , то, очевидно,  $\lambda_i g_i(z) = 0$  и условие дополнительной нежёсткости выполнено. Если же  $g_i(z) < 0$ , то  $g_i(x^\nu) < 0$  при достаточно больших  $\nu$ , следовательно,

$$\lambda_i^\nu = \frac{\nu}{K_\nu} h'(g_i(x^\nu)) = 0,$$

поэтому условие дополняющей нежёсткости выполнено и в этом случае.

В силу (12)

$$\lambda_0^\nu = \frac{1}{K_\nu} \geq 0, \quad \lambda_i^\nu \geq 0 \quad (i \in I_+),$$

поэтому  $\lambda_0 \geq 0, \lambda_i \geq 0 \forall i \in I_+$ . Это означает, что набор  $\Lambda = (\lambda_i)_{i \in I_1}$  является искомым. Теорема в случае компактного множества  $Q$  доказана (и даже в более сильной формулировке – набор  $\Lambda$  оказался нормированным).

2-й этап:  $Q$  – выпуклое (возможно, и незамкнутое) подмножество  $\mathbb{R}^n$ . Пусть  $v_1, \dots, v_m$  – произвольный конечный набор элементов  $v_k$  из множества  $Q$ ,  $P$  – выпуклая оболочка множества  $\{z, v_1, \dots, v_m\}$ . Очевидно,  $P$  – компактное выпуклое подмножество  $Q$  (многогранник),  $z \in P$  и  $z$  минимизирует функцию  $g_0$  на множестве элементов, принадлежащих  $P$  и удовлетворяющих ограничениям (3), (4). Применим к многограннику  $P$  доказанную часть теоремы. В силу уже доказанного существует такой нормированный набор  $\Lambda = (\lambda_i)$ , что

$$-\sum_{i \in I_1} \lambda_i g'_i(z) \in N_P(z) \quad (14)$$

и выполнены условия дополнительной нежёсткости (6). Из (14) вытекают неравенства

$$\sum_{i \in I_1} \lambda_i g'_i(z) (v_k - z) \geq 0 \quad (k = 1, \dots, m). \quad (15)$$

Обозначим через  $\Lambda(P)$  совокупность нормированных наборов  $\Lambda$ , удовлетворяющих условиям (14), (15). Очевидно, что  $\Lambda(P)$  – непустое замкнутое подмножество единичной сферы конечномерного пространства. Система множеств  $\Lambda(P)$ , соответствующих многогранникам  $P$  описанного выше вида, центрирована. Поэтому пересечение множеств вида  $\Lambda(P)$  непусто, т.е. существует универсальный нормированный набор  $\Lambda$ , для которого соотношения (14), (15) выполняются при любом выборе конечного множества  $\{v_k\}$ . Этот универсальный набор  $\Lambda$  является искомым. ►

Из теоремы 1 вытекает

**Следствие.** Если выполнены предположения теоремы 1, то существует такой допустимый набор  $\Lambda = (\lambda_i)$ , что справедливы условия дополнительной нежёсткости (6) и элемент  $z$  является решением экстремальной задачи

$$\sum_{i \in I_1} \lambda_i g'_i(z) v \rightarrow \min, \quad v \in Q. \quad (16)$$

Числа  $\lambda_i$  называют множителями Лагранжа, а сам переход от задачи (3)–(5) к задаче (16) – правилом множителей Лагранжа. С иными версиями этого правила и с иными доказательствами теоремы 3 и более общих утверждений можно познакомиться, например, в [1], [5], [10], [19], [25].

**3. Обсуждение правила множителей Лагранжа.** Варианты теоремы 3 сохраняются, если одно из множеств  $I_0, I_+$  пусто. Например, верна

**Теорема 4.** Пусть  $z$  – решение задачи (3), (5) и функции  $g_i$  непрерывно дифференцируемы в окрестности точки  $z$ . Тогда найдутся такие неотрицательные числа  $\lambda_i$  ( $i \in \{0\} \cup I_+$ , не все равные нулю, что выполнены соотношение (13) с  $I_1 = \{0\} \cup I_+$  и условия дополняющей нежёсткости (6).

Полезно следующее дополнение к теореме 4.

**Теорема 5.** Пусть функции  $g_i$  ( $i \in \{0\} \cup I_+$  выпуклы,  $\lambda_i \geq 0$  ( $i \in I_+$ ),  $\lambda_0 > 0$ . Пусть

$$z \in Q, \quad g_i(z) \leq 0, \quad \lambda_i g_i(z) = 0 \quad (i \in I_+)$$

и  $z$  есть решение задачи (16) с  $I_1 = \{0\} \cup I_+$ . Тогда  $z$  является решением задачи (3), (5).

◀ Функция

$$g(x) = \sum_{i \in \{0\} \cup I_+} \lambda_i g_i(x)$$

выпукла как неотрицательная линейная комбинация выпуклых функций. Поскольку  $z$  есть решение задачи (16), то  $-g'(z) \in N_Q(z)$ . В силу теоремы 2  $g(x) \geq g(z) \forall x \in Q$ . Используя это неравенство и условие дополнительной нежёсткости, получаем последовательно соотношения

$$\sum_{i \in \{0\} \cup I_+} \lambda_i g_i(x) \geq \sum_{i \in \{0\} \cup I_+} \lambda_i g_i(z) = \lambda_0 g_0(z) \quad (x \in Q).$$

Из этого соотношения вытекает, что если  $x \in Q$  и  $g_i(x) \leq 0$  ( $i \in I_+$ ), то  $\lambda_0 g_0(x) \geq \lambda_0 g_0(z)$ . Поскольку  $\lambda_0 > 0$ , то  $z$  – решение задачи (3), (5). ►

Коснёмся способов описания конуса  $N_Q(z)$  опорных функционалов к множеству  $Q \subset \mathbb{R}^n$ . Ограничимся случаем, когда множество  $Q$  определено равенством

$$Q = \{x \in \mathbb{R}^n, \Phi_1(x) \leq 0, \dots, \Phi_p(x) \leq 0\},$$

где  $\Phi_k: \mathbb{R}^n \rightarrow \mathbb{R}$  – выпуклые дифференцируемые функции. Пусть  $z \in Q$  и множество  $I(z) = \{i, \Phi_i(z) = 0\}$  непусто. При достаточно необременительных предположениях любой линейный функционал из  $N_Q(z)$  допускает представление

$$l = \sum_{i \in I(z)} \lambda_i \Phi'_i(z), \quad (17)$$

в котором  $\lambda_i$  – неотрицательные числа. Например, формула (17) справедлива, если функции  $\Phi_i$  аффинны или выполнено условие сильной совместности: существует элемент  $\bar{v}$ , для которого  $\Phi_i(\bar{v}) < 0 \forall i$ . В частности, представление (17) можно применить к случаю, когда  $Q$  есть прямое произведение отрезков действительной оси (именно этот частный случай был важен во многих конструкциях второй главы).

В правиле множителей Лагранжа возможна ситуация, когда  $\lambda_0 = 0$ . Это означает, что в необходимом условии минимума отсутствует сама минимизируемая функция  $g_0$ . Любое предположение, гарантирующее неравенство  $\lambda_0 > 0$ , называют условием регулярности данных задачи (3)-(5). Не обсуждая и здесь вопрос с максимальной общностью, ограничимся случаем, когда  $I_0 = \emptyset$  (ограничения типа равенств (4) отсутствуют), а функции  $g_i: \mathbb{R}^n \rightarrow \mathbb{R}$  ( $i \in I_+$ ) выпуклы и дифференцируемы. Сильная совместность ограничений (5) гарантирует неравенство  $\lambda_0 > 0$ . Именно, пусть существует элемент  $\bar{v}$  из  $Q$ , для которого  $g_i(\bar{v}) < 0$ . Предположение  $\lambda_0 = 0$  влечёт за собой неравенство

$$\sum_{i \in I_+} \lambda_i g'_i(z)(\bar{v} - z) \geq 0, \quad (18)$$

в котором хотя бы одно из чисел  $\lambda_i$  положительно. Поскольку  $g_i(\bar{v}) < 0 \forall i$ , то

$$\sum_{i \in I_+} \lambda_i g_i(\bar{v}) < 0.$$

С другой стороны, имеют место оценки

$$\sum_{i \in I_+} \lambda_i g_i(\bar{v}) \geq \sum_{i \in I_+} \lambda_i g_i(\bar{z}) + \sum_{i \in I_+} g'_i(z)(\bar{v} - z) \geq 0,$$

вытекающие из (6), (18). Полученное противоречие доказывает неравенство  $\lambda_0 > 0$ .

**4. Задачи.** Цель предлагаемого цикла задач состоит в обобщении правила множителей Лагранжа на негладкие функции. Напомним, что функционал  $h: \mathbb{R}^n \rightarrow \mathbb{R}$  называют сублинейным, если он выпукл и положительно однороден. Сублинейность функционала  $h: \mathbb{R}^n \rightarrow \mathbb{R}$  эквивалентна соотношениям

$$h(x + y) \leq h(x) + h(y), \quad h(\lambda x) = \lambda h(x),$$

где  $x, y$  – произвольные элементы из  $\mathbb{R}^n$ ,  $\lambda \geq 0$ . Например, если  $M$  – произвольное ограниченное подмножество линейных на  $\mathbb{R}^n$  функционалов, то равенство

$$h(x) \stackrel{\text{def}}{=} \sup_{l \in M} l(x)$$

определяет сублинейный функционал  $h$ .

Задача 1. Линейный функционал  $l: \mathbb{R}^n \rightarrow \mathbb{R}$  называют опорным к сублинейному функционалу  $h$ , если  $l(v) \leq h(v)$  для всех  $v$  из  $\mathbb{R}^n$ . Используя конечномерный вариант теоремы Хана–Банаха [1], [3], доказать следующие утверждения:

- а) множество  $\partial h$  всех опорных функционалов  $l$  к сублинейному функционалу  $h$  есть непустое компактное выпуклое подмножество пространства  $\mathbb{R}_n$ ;
- б) справедливо равенство

$$h(v) = \max_{l \in \partial h} l(v) \quad \forall v \in \mathbb{R}^n.$$

2. Пусть  $g$  – действительная функция, определённая на некоторой окрестности  $U$  точки  $x$  и удовлетворяющая в этой окрестности условию Липшица. Тогда

- а) для любого  $v$  из  $\mathbb{R}^n$  существует конечный верхний предел

$$h(v) = \overline{\lim}_{y \rightarrow x, t \rightarrow +0} \frac{g(y + tv) - g(y)}{t};$$

б) определяемый таким образом функционал  $h: \mathbb{R}^n \rightarrow \mathbb{R}$  сублинеен, соответствующее ему множество  $\partial h$  именуют субдифференциалом функции  $g$  в точке  $x$  и обозначают символом  $\partial g(x)$ .

3. Пусть  $Q$  – выпуклое подмножество  $\mathbb{R}^n$ , функционал  $g$  определён на выпуклой окрестности  $U$  точки  $x \in Q$ , удовлетворяет условию Липшица в этой окрестности и  $g(x) \leq g(y) \forall y \in Q \cap U$ . Тогда

$$0 \in \partial g(x) + N_Q(x). \quad (19)$$

4. Включение (19) очевидным образом аналогично включению (2), а в целом 3. представляет вариант теоремы 1 о необходимых условиях минимума на выпуклом множестве. Сформулировать и доказать для липшицевых функций  $g, g_i$  аналоги всех сформулированных в приложении А утверждений.

5. Применить аналог правила Лагранжа для липшицевых функций к обобщению принципа максимума Понтрягина.

## Приложение В

### О гладкости решений одномерных вариационных задач

**1. Постановка задачи.** В теории вариационных задач имеется определённый разрыв между теоремами существования и условиями минимума. Теоремы существования обычно устанавливают в достаточно широких нормированных пространствах, в которых минимизируемый функционал является растущим и полунепрерывным снизу. С другой стороны, в формулировках условий минимума (и необходимых, и достаточных) фигурирует, как правило, дифференцируемость исследуемого функционала, что заставляет предполагать принадлежность точки минимума некоторому узкому пространству. Один из естественных путей преодоления этого разрыва для интегральных функционалов заключается в доказательстве гладкости слабых решений вариационных задач, существование которых гарантируется прямыми методами. В случае интегральных функционалов со степенным порядком роста интегранта имеется большое количество исследований, связанных с этим вопросом; подробная библиография содержится в [22], [27].

Менее изучен вопрос о гладкости экстремалей с нестепенными нелинейностями. В монографии ([10], с. 394–398) приводится теорема о гладкости слабых решений простейшей вариационной задачи

$$f(x) = \int_{t_0}^{t_1} L(t, x(t), x'(t)) dt \rightarrow \min,$$
$$x(t_0) = x_0, \quad x(t_1) = x_1.$$

В предположениях этой теоремы интегрант  $L(t, x, y)$  может расти по  $y$  нестепенным образом, однако удовлетворяет оценке

$$|L(t, x, y)| \leq c(t, x) \exp(k|y|).$$

Таким образом, указанная теорема не охватывает случая функций  $L$ , растущих по  $y$  быстрее экспоненты.

В этом приложении изучается вопрос о гладкости обобщённых решений вариационной задачи

$$f(x) = \int_{t_0}^{t_1} L(t, x(t), x'(t), \dots, x^{(m)}(t)) dt \rightarrow \min, \quad (1)$$

$$x^{(i)}(t_0) = a_i, \quad x^{(i)}(t_1) = b_i, \quad (i = 0, 1, \dots, m-1), \quad (2)$$

Для точной постановки задачи введем несколько обозначений и терминов. Далее  $\Delta = [t_0, t_1]$ , интеграл

$$\int_{\Delta} h(t) dt$$

от функции  $h: \Delta \rightarrow \overline{\mathbb{R}}$  понимается в смысле Лебега. Предполагается знакомство с одномерным интегралом Лебега; более чем достаточно сведений, содержащихся, например, в [3]. Ниже  $\mathcal{L}(\Delta)$  – пространство суммируемых на отрезке  $\Delta$  функций,  $C^k(\Delta)$  – пространство  $k$  раз непрерывно дифференцируемых на отрезке  $\Delta$  скалярных функций.

Непрерывную на отрезке  $\Delta$  функцию  $x$  будем называть абсолютно непрерывной, если она допускает представление

$$x(t) = x(t_0) + \int_{t_0}^t v(s) ds, \quad (A)$$

в котором  $v \in \mathcal{L}(\Delta)$ . Функция  $x: \Delta \rightarrow \mathbb{R}$  абсолютно непрерывна в том и только в том случае, когда для каждого  $\varepsilon > 0$  существует такое  $\delta > 0$ , что если  $(\alpha_j, \beta_j)$  – конечная система попарно непересекающихся интервалов, суммарная длина которых меньше  $\delta$ , то

$$\sum_j |x(\beta_j) - x(\alpha_j)| < \varepsilon.$$

Любая абсолютно непрерывная функция  $x(t)$  имеет почти всюду суммируемую производную  $x'(t)$ , и функция  $x'(t)$  эквивалентна фигурирующей в равенстве (A) функции  $v(t)$ . Совокупность абсолютно непрерывных на отрезке  $\Delta$  функций обозначим символом  $W^1(\Delta)$ . В класс  $W^1(\Delta)$  входят функции  $x: \Delta \rightarrow \mathbb{R}$ , удовлетворяющие условию Липшица. Функция  $x(t) = \sqrt{t}$  принадлежит классу  $W^1([0, 1])$ , но не удовлетворяет условию Липшица. Для того чтобы функция  $x$  класса  $W^1(\Delta)$  удовлетворяла условию Липшица, необходимо и достаточно, чтобы её производная  $x'$  принадлежала пространству  $\mathcal{L}^\infty(\Delta)$ . Доказательства перечисленных свойств абсолютно непрерывных функций можно найти, например, в [3].

Пусть  $m$  – натуральное число и  $m > 1$ . Обозначим через  $W^m(\Delta)$  часть пространства  $C^{m-1}(\Delta)$ , состоящую из функций  $x$ , для которых производная



$x^{(m-1)}$  порядка  $m-1$  абсолютно непрерывна на отрезке  $\Delta$ . Для любой функции  $x$  класса  $W^m(\Delta)$  справедливо равенство

$$x^{(m-1)}(t) = x^{(m-1)}(t_0) + \int_{t_0}^t u(s) ds. \quad (3)$$

Функция  $u$  в равенстве (3) суммируема; её называют производной порядка  $m$  от функции  $x$ . Для функции  $x$  из  $W^m(\Delta)$  положим

$$\delta x(t) = (x(t), x'(t), \dots, x^{(m-1)}(t)).$$

Далее  $C_0^\infty(\Delta)$  – пространство бесконечно дифференцируемых на  $\Delta$  функций, равных нулю вблизи границы промежутка  $\Delta$ . Если  $u(t) = x^{(m)}(t)$ , то для любой функции  $\varphi$  класса  $C_0^\infty(\Delta)$  имеет место равенство

$$\int_{\Delta} x(t) \varphi^{(m)}(t) dt = (-1)^m \int_{\Delta} u(t) \varphi(t) dt. \quad (4)$$

Весьма часто равенство (4) принимают в качестве определения производной порядка  $m$ . Это новое определение имеет то преимущество, что легко распространяется на функции многих переменных [22]; соответствующие частные производные именуют производными в смысле Соболева. В одномерном случае оба определения по существу эквивалентны.

Если не оговорено противное, в пределах этого параграфа предполагается, что  $L(t, u_0, \dots, u_m)$  – непрерывная на  $\Delta \times \mathbb{R}^{m+1}$  функция. Введем множество  $D_f$  функций  $x$  класса  $W^m(\Delta)$ , удовлетворяющих краевым условиям (2), для которых функция  $L(\cdot, \delta x(\cdot), x^{(m)}(\cdot))$  принадлежит  $\mathcal{L}(\Delta)$ . Множество  $D_f$  является естественной областью определения функционала  $f$ . Для  $x(\cdot)$  из  $W^m(\Delta)$ , не принадлежащих  $D_f$ , положим  $f(x) = \infty$ . Функцию  $x_0(\cdot)$  из  $W^m(\Delta)$  будем называть решением вариационной задачи (1), (2), если  $x_0 \in D_f$  и для всех  $x$  из  $W^m(\Delta)$  имеет место неравенство  $f(x_0) \leq f(x)$ . Ниже существование решения задачи (1), (2) предполагается. Достаточные условия выполнимости этого требования можно найти в [10]. В частности, справедливо

**Предложение 1.** Пусть функция  $L(t, u_0, u_1)$  выпукла по переменной  $u_1$  (при фиксированных значениях остальных переменных) и допускает оценку

$$L(t, u_0, u_1) \geq \psi_1(u_1) - \psi_2(u_0) - k, \quad (5)$$

где  $k$  – положительная константа, а функции  $\psi_1, \psi_2$  удовлетворяют условиям:

1°  $\psi_1$  выпуклая чётная функция на  $[0, \infty)$ ,  $\psi_1(0) = 0$  и  $u^{-1}\psi_1(u) \rightarrow \infty$  при  $u \rightarrow \infty$ ;

2°  $\psi_2$  неотрицательна, непрерывна и неубывает на  $[0, \infty)$ ;

3° при любых положительных числах  $\gamma, \lambda_0$  функция

$$\psi_1(u/\gamma) - \psi_2(u + \lambda_0) \rightarrow \infty \quad \text{при} \quad u \rightarrow \infty.$$

Тогда задача (1) при  $m = 1$  имеет решение.

Обсуждение предложения 1 и других теорем разрешимости можно найти в [10, с. 384–386].

**2. Уравнения Эйлера–Лагранжа в слабой форме.** В этом пункте формулируется и доказывается центральный результат приложения.

**Теорема 1.** Пусть функция  $L: \Delta \times \mathbb{R}^{m+1} \rightarrow \mathbb{R}$  удовлетворяет условиям:

$l_1$ ) существуют и непрерывны по совокупности переменных частные производные  $L_{u_i}$  ( $i = 0, 1, \dots, m$ );

$l_2$ ) функция  $L$  выпукла по переменному  $u_m$ ;

$l_3$ ) для каждой константы  $R$  найдутся выпуклая неотрицательная функция  $\psi_R$  и постоянные  $c_R, k_R$  такие, что  $\lim_{|z| \rightarrow \infty} \psi_R(z) = \infty$  и при

$$|u_0| + \dots + |u_{m-1}| \leq R$$

справедливы оценки

$$L(t, u_0, \dots, u_m) \geq \psi_R(u_m) - c_R, \quad (6)$$

$$\sum_{i=0}^{m-1} |L_{u_i}(t, u_0, \dots, u_m)| \leq k_R \psi(u_m) + c_R. \quad (7)$$

Если  $x_0$  – решение вариационной задачи (1), (2), то функции  $L_{u_i}(\cdot, \delta x_0(\cdot), x_0(\cdot))$  принадлежат  $\mathcal{L}(\Delta)$  и для любой функции  $\varphi$  из  $C_0^\infty(I)$  выполняется равенство

$$\sum_{i=0}^m \int_{\Delta} L_{u_i}(t, \delta x_0(t), x_0^{(m)}(t)) \varphi^{(i)}(t) dt = 0. \quad (8)$$

◀ Для простоты выкладок будем считать, что

$$a_i = b_i = 0 \quad (i = 0, 1, \dots, m-1);$$

общий случай сводится к этому простой заменой переменных. Пусть  $x \in C_0^\infty(\Delta)$ ,  $h = x - x_0$ ,

$$R = \sum_{i=0}^{m-1} \max_{t \in I} |x_0^{(i)}(t)| + \sum_{i=0}^{m-1} \max_{t \in I} |x^{(i)}(t)|,$$

$\psi_R$  – выпуклая функция,  $c_R, k_R$  – постоянные, построенные в соответствии с условием  $l_3$ ) по константе  $R$ . Фиксируем  $\lambda \in [0, 1]$  и установим суммируемость функции  $L(\cdot, \delta x_0 + \lambda \delta h, x_0^{(m)})$ . В силу леммы о промежуточных значениях существует такая измеримая<sup>1</sup> на  $\Delta$  функция  $\theta(t)$ , что  $0 \leq \theta(t) \leq 1$  и

<sup>1</sup>Возможность измеримого выбора  $\theta(t)$  следует обосновать читателю.

$$L(\cdot, \delta x_0 + \lambda \delta h, x_0^{(m)}) - L(\cdot, \delta x_0, x_0^{(m)}) = \sum_{i=0}^{m-1} L_{u_i}(\cdot, \delta x_0 + \lambda \theta \delta h, x_0^{(m)}) \lambda h^{(i)}. \quad (8)$$

Из условия l3) вытекают соотношения

$$\int_{\Delta} \psi_R(x_0^{(m)}(t)) dt < \infty,$$

$$|L_{u_i}(\cdot, \delta x_0 + \lambda \theta \delta h, x_0^{(m)})| \leq k_R \psi_R(x_0^{(m)}) + c_R \quad (i = 0, \dots, m-1),$$

влекущие в силу (8) суммируемость функций  $L(\cdot, \delta x_0 + \lambda \delta h, x_0^{(m)})$ . При  $\lambda = 0$  из этих соотношений следует суммируемость функций

$$L_{u_i}(\cdot, \delta x_0, x_0^{(m)}) \quad (i = 0, \dots, m-1).$$

Поскольку  $x \in C_0^\infty(\Delta)$ , то  $L(\cdot, \delta x_0 + \lambda \delta h, (x_0 + h)^{(m)}) \in \mathcal{L}(I)$ . Функция  $g = L(\cdot, \delta(x_0 + \lambda h), (x_0 + \lambda h)^{(m)})$  удовлетворяет неравенствам

$$\begin{aligned} & \psi_R((x_0 + \lambda h)^{(m)}) - c_R \leq g \leq \\ & \leq \lambda L(\cdot, \delta x_0 + \lambda \delta h, (x_0 + h)^{(m)}) + (1 - \lambda) L(\cdot, \delta(x_0 + \lambda h), x_0^{(m)}). \end{aligned} \quad (9)$$

Как указывалось выше, функции  $L(\cdot, \delta x_0 + \lambda \delta h, (x_0 + h)^{(m)})$ ,  $L(\cdot, \delta(x_0 + \lambda h), x_0^{(m)})$  принадлежат  $\mathcal{L}(\Delta)$ , поэтому правая часть (9) принадлежит  $\mathcal{L}(\Delta)$ . Поскольку  $\psi_R(x_0^{(m)}) \in \mathcal{L}(\Delta)$ ,  $\psi_R((x_0 + \lambda h)^{(m)}) \in \mathcal{L}(\Delta)$ ,  $\psi_R$  – неотрицательная выпуклая функция, то левая часть (9) также суммируема. Следовательно,

$$L(\cdot, \delta(x_0 + \lambda h), (x_0 + \lambda h)^{(m)}) \in \mathcal{L}(\Delta), \quad x_0 + \lambda h \in D_f.$$

Аналогичным образом можно установить суммируемость функции

$$L(\cdot, \delta x_0, (x_0 + h)^{(m)}).$$

Так как  $x_0$  – решение вариационной задачи (1), (2), то

$$f(x_0 + \lambda h) \geq f(x_0) \quad \forall \lambda \in [0, 1].$$

При  $\lambda \in (0, 1]$  очевидны соотношения

$$0 \leq \frac{f(x_0 + \lambda h) - f(x_0)}{\lambda} = f_1(\lambda) + f_2(\lambda), \quad (10)$$

в которых

$$f_1(\lambda) = \lambda^{-1} \left[ f(x_0 + \lambda h) - \int_{\Delta} L(t, \delta x_0(t), (x_0 + \lambda h)^{(m)}(t)) dt \right],$$

$$f_2(\lambda) = \lambda^{-1} \left[ \int_{\Delta} L(t, \delta x_0(t), (x_0 + \lambda h)^{(m)}(t)) dt - f(x_0) \right].$$

Используя вновь лемму о промежуточных значениях, получаем равенство

$$\begin{aligned} & L(t, \delta(x_0 + \lambda h)(t), (x_0 + \lambda h)^{(m)}(t)) - L(t, \delta x_0(t), (x_0 + \lambda h)^{(m)}(t)) = \\ &= \sum_{i=0}^{m-1} L_{u_i}(t, \delta x_0(t) + \theta(t)\lambda\delta h(t), (x_0 + \lambda h)^{(m)}(t))\lambda h^{(i)}(t), \end{aligned} \quad (11)$$

где  $\theta(t)$  – измеримая функция,  $0 \leq \theta(t) \leq 1$ . Из условий  $l_2), l_3)$  следует

$$\begin{aligned} & |L_{u_i}(t, \delta x_0(t) + \theta(t)\lambda\delta h(t), (x_0 + \lambda h)^{(m)}(t))| \leq \\ & \leq k_R \lambda \psi_R((x_0 + \lambda h)^{(m)}(t)) + k_R(1 - \lambda)\psi_R(x_0^{(m)}(t)) + c_R. \end{aligned} \quad (12)$$

Введем последовательность функций  $y_k(t) =$

$$= k \left( L(t, \delta(x_0 + k^{-1}h)(t), (x_0 + k^{-1}h)^{(m)}(t)) - L(t, \delta x_0(t), (x_0 + k^{-1}h)^{(m)}(t)) \right).$$

Очевидно, что

$$f_1(k^{-1}) = \int_{\Delta} y_k(t) dt \quad (k = 1, 2, \dots).$$

Соотношение (11) показывает, что последовательность  $y_k$  при  $k \rightarrow \infty$  почти всюду сходится к функции

$$\sum_{i=0}^{m-1} L_{u_i}(t, \delta x_0(t), (x_0)^{(m)}(t))h^{(i)}(t).$$

Из (12) вытекает абсолютная непрерывность интегралов от функций  $y_k$ , поэтому

$$\lim_{k \rightarrow \infty} f_1(k^{-1}) = \int_{\Delta} \sum_{i=0}^{m-1} L_{u_i}(t, \delta x_0(t), (x_0)^{(m)}(t))h^{(i)}(t) dt < \infty.$$

Отсюда в силу соотношений (10) следует ограниченность снизу последовательности  $f_2\left(\frac{1}{k}\right)$ . Поскольку функция  $L$  выпукла по переменной  $u_n$ , то последовательность функций

$$z_k(t) = k \left( L(t, x_0(t), (x_0 + k^{-1}h)^{(m)}(t)) - L(t, x_0(t), x_0^{(m)}(t)) \right)$$

монотонно убывает с ростом  $k$ . Так как функция  $L$  дифференцируема, то почти всюду существует предел этой функции, равный  $L_{u_m}(t, \delta x(t), x_0^{(m)}(t))h^{(m)}(t)$ . Согласно теореме Бешпо Леви предельная функция суммируема и

$$\lim_{k \rightarrow \infty} \int_{\Delta} z_k(t) dt = \int_{\Delta} L_{u_m}(t, \delta x(t), x_0^{(m)}(t))h^{(m)}(t) dt.$$

Переходя к пределу в неравенстве  $f_1(k^{-1}) + f_2(k^{-1}) \geq 0$  к пределу при  $k \rightarrow \infty$ , получаем, что функция

$$\sum_{i=0}^m L_{u_i}(t, \delta x_0(t), x_0^{(m)}(t))(x - x_0)^{(i)}(t)$$

суммируема и интеграл от неё неотрицателен:

$$\int_{\Delta} \sum_{i=0}^m L_{u_i}(t, \delta x_0(t), x_0^{(m)}(t))(x - x_0)^{(i)}(t) dt \geq 0. \quad (13)$$

Соотношение (13) верно для произвольной функции  $x$  из  $C_0^\infty(\Delta)$ . Полагая  $x(t) \equiv 0$ , получаем включение

$$L_{u_m}(t, \delta x_0(t), x_0^{(m)}(t))x_0^{(m)}(t) \in \mathcal{L}(\Delta).$$

Согласно условиям  $l_2), l_3)$  при  $|u_0| + \dots + |u_{m-1}| \leq R$  справедливы оценки

$$\begin{aligned} L_{u_m}(t, u_0, \dots, u_m)u_m &\geq L(t, u_0, \dots, u_n) - L(t, u_0, \dots, u_{m-1}, 0) \geq \\ &\geq \psi_R(u_m) - c_R - L(t, u_0, \dots, u_{m-1}, 0), \end{aligned}$$

поэтому существует такое  $M > 0$ , что при  $|u_0| + \dots + |u_{m-1}| \leq R, |u_m| \geq M$  имеет место неравенство

$$L_{u_m}(t, u_0, \dots, u_m)u_m > 0.$$

Если  $|x_0^{(m)}(t)| > M + 1$ , то

$$L_{u_m}(t, \delta x_0(t), x_0^{(m)}(t))x_0^{(m)}(t) \geq |L_{u_m}(t, \delta x_0(t), x_0^{(m)}(t))|,$$

поэтому

$$L_{u_m}(; \delta x_0(\cdot), x_0^{(m)}(\cdot)) \in \mathcal{L}(\Delta).$$

Положим теперь в (13)  $x(t) = s\varphi(t)$ , где  $\varphi \in C_0^\infty(\Delta)$ ,  $s$  — действительное число. Тогда имеем

$$\int_{\Delta} \sum_{i=0}^m L_{u_i}(t, \delta x_0(t), x_0^{(m)}(t))(s\varphi - x_0)^{(i)}(t) dt \geq 0$$

для произвольного  $s \in (-\infty, \infty)$ . Это возможно лишь в случае, когда выполнено (8). ►

Остановимся на возможных обобщениях теоремы. Требование непрерывности функций  $L, L_{u_i}$  по совокупности переменных можно заменить соответствующим условием Каратеодори (измеримость по  $t$  и непрерывность по совокупности переменных  $u_0, \dots, u_m$ ). В теореме 1 предполагается выпуклость  $L$  лишь по одному переменному. Усиление условия выпуклости функции  $L$ , например, до выпуклости по совокупности переменных  $u_0, \dots, u_m$ , позволяет ослабить условие  $l_3$ ), основное назначение которого заключается в подавлении влияния невыпуклой зависимости  $L$  от переменных  $u_0, \dots, u_{m-1}$ .

### 3. Гладкость обобщённых решений дифференциальных уравнений.

Соотношение (8) означает, что функция  $x_0$  удовлетворяет в смысле обобщённых функций уравнению Эйлера

$$\sum_{i=0}^m (-1)^i \frac{d^i}{dt^i} L_{u_i}(t, \delta x(t), x^{(m)}(t)) = 0. \quad (14)$$

Изучим свойства обобщённых решений дифференциальных уравнений вида

$$\sum_{i=0}^m (-1)^i \frac{d^i}{dt^i} A_i(t, \delta x(t), x^{(m)}(t)) = 0, \quad (15)$$

где  $A_i(t, u_0, \dots, u_m)$  – непрерывные на  $\Delta \times \mathbb{R}^{m+1}$  функции. Функцию  $x_0$  из пространства  $W^m(\Delta)$  будем называть обобщённым решением уравнения (15), если функции  $A_i(\cdot, \delta x_0(\cdot), x_0^{(m)}(\cdot))$  ( $i = 0, 1, \dots, m$ ) принадлежат  $\mathcal{L}(\Delta)$  и для любой функции  $\varphi$  из  $C_0^\infty(\Delta)$  справедливо равенство

$$\int_{\Delta} \sum_{i=0}^m A_i(t, \delta x_0(t), x_0^{(m)}(t)) \varphi^{(i)}(t) dt = 0. \quad (16)$$

**Теорема 2.** Пусть выполняются условия:

- 1)  $A_i(t, u_0, \dots, u_m)$  есть  $i$  раз непрерывно дифференцируемая функция;
- 2)  $\partial A_n / \partial u_n > 0$  на  $\Delta \times \mathbb{R}^{m+1}$ ;
- 3) для любой константы  $R$  существуют неотрицательная выпуклая функция  $\psi_R$  и константа  $c_R$  такие, что  $|z|^{-1} \psi_R(z) \rightarrow \infty$  и при  $|u_0| + \dots + |u_{m-1}| \leq R$  справедливо неравенство

$$A_m(t, u_0, \dots, u_m) u_m \geq \psi_R(u_m) - c_R.$$

Тогда каждое обобщённое решение уравнения (15) есть функция класса  $C^{2m}(I)$  и является классическим решением уравнения (15).

◀ Пусть  $m = 2$ . Введём функции

$$g_0(t) = \int_{t_0}^t (t-s) A_0(s, \delta x_0(s), x_0^{(2)}(s)) ds,$$

$$g_1(t) = \int_{t_0}^t A_1(s, \delta x_0(s), x_0^{(2)}(s)) ds,$$

$$g_2(t) = A_2(t, \delta x_0(t), x_0^{(2)}(t)).$$

Интегральные соотношения, определяющие функции  $g_0, g_1$ , эквивалентны условиям

$$g_0^{(2)} = A_0(t, \delta x_0(t), x_0^{(2)}(t)), \quad g_0(t_0) = 0, \quad g_0'(t_0) = 0,$$

$$g_1'(t) = A_1(t, \delta x_0(t), x_0^{(2)}(t)), \quad g_1(t_0) = 0.$$

Используя эти соотношения, равенство (16) и интегрирование по частям, получаем, что

$$\int_{\Delta} (g_2(t) - g_1(t) + g_0(t)) \varphi^{(2)}(t) dt = 0$$

для любой функции  $\varphi$  из  $C_0^\infty(\Delta)$ . Это означает, что вторая производная в смысле Соболева функции  $g_2(t) - g_1(t) + g_0(t)$  равна 0, поэтому данная функция эквивалентна некоторому многочлену  $P(t)$  первой степени. Таким образом, почти всюду в  $\Delta$  выполнено равенство

$$A_2(t, \delta x_0(t), x_0^{(2)}(t)) = \int_{t_0}^t A_1(s, \delta x_0(s), x_0^{(2)}(s)) ds -$$

$$- \int_{t_0}^t (t-s) A_0(s, \delta x_0(s), x_0^{(2)}(s)) + P(t) \equiv Q(t). \quad (17)$$

Введём функцию  $H(t, z)$ , определённую равенством  $H(t, z) = A_2(t, \delta x_0(t), z)$  ( $t \in \Delta, z \in \mathbb{R}$ ). Пусть  $R \geq |x_0(t)| + |x_0'(t)|$  при всех  $t \in \Delta$ . Из условия 3 получим

$$(H(t, z) - Q(t))z \geq \psi_R(z) - c_r - Q(t)z,$$

а поскольку  $\psi_R(z)|z|^{-1} \rightarrow \infty$  при  $|z| \rightarrow \infty$ , то при  $|z| \geq M$ , где  $M$  достаточно велико, правая часть последнего неравенства положительна. Отсюда вытекает существование и ограниченность неявной функции  $y(t)$ , определяемой из уравнения

$$H(t, y(t)) = Q(t). \quad (18)$$

Согласно условию 2 функция  $H(t, z) = A_2(t, \delta x_0(t), z)$  монотонно возрастает по  $z$ , поэтому уравнение (18) имеет лишь одно решение. Из единственности и ограниченности неявной функции  $y(t)$ , определяемой из (18), в силу непрерывности  $H(t, z)$  и  $Q(t)$ , следует непрерывность функции  $y(t)$ . В соответствии с (17)  $x_0^{(2)}(t) = y(t)$  почти всюду, отсюда следует включение  $x_0 \in C^2(\Delta)$  и равенство  $x_0^{(2)}(t) = y(t)$  при всех  $t$  из  $\Delta$ .

Равенство (17) показывает, что  $Q \in C^1(\Delta)$ . По теореме о неявной функции, применённой к (18),  $y \in C^1(\Delta)$ , т.е.  $x_0 \in C^3(\Delta)$ . Используя снова (17), имеем:  $Q \in C^2(\Delta)$ . Применяя ещё раз теорему о неявной функции, приходим к включению  $y \in C^2(\Delta)$ , т.е.  $x_0 \in C^4(\Delta)$ . Функция  $x_0$  является классическим решением уравнения (15). Для того чтобы убедиться в этом, достаточно дважды дифференцировать (17) по  $t$ .

Теорема в случае  $m = 2$  доказана. Общий случай рассматривается аналогично. ►

Следствием теорем 1 и 2 является

**Теорема 3.** Пусть функция  $L$  удовлетворяет условиям:

1° частная производная  $L_{u_i}$  ( $i = 0, 1, \dots, m$ ) есть  $i$  раз непрерывно дифференцируемая функция;

2° производная  $\partial^2 / \partial u_m^2 L(t, t, u_0, \dots, u_m) > 0$  для всех допустимых  $t, u_j$ ;

3° имеют место неравенства (6), (7), в которых

$$\lim_{|z| \rightarrow \infty} \frac{\psi_R(z)}{|z|} = \infty.$$

Если  $x_0$  есть обобщённое решение вариационной задачи (1), (2), то  $x_0$  принадлежит  $C^{2m}(\Delta)$  и является классическим решением уравнения (14).

◀ Каждое условие  $i^\circ$  влечёт за собой условие  $l_i$  теоремы 1, поэтому согласно теореме 1  $x_0$  является обобщённым решением уравнения (14). Применим теорему 2 к уравнению (14), полагая  $A_i = L_{u_i}$  ( $i = 0, 1, \dots, m$ ). В силу выпуклости  $L$  по  $u_m$  при  $|u_0| + \dots + |u_{m-1}| \leq R$  имеем соотношения:

$$\begin{aligned} L_{u_m}(t, u_0, \dots, u_m) u_m &\geq L(t, u_0, \dots, u_m) - L(t, u_0, \dots, u_{m-1}, 0) \geq \\ &\geq \psi_R(u_m) - c_R - L(t, u_0, \dots, u_{m-1}, 0), \end{aligned}$$

из которых следует выполнимость условия 3 теоремы 2. Поскольку условия 1 и 2 теоремы 2 выполнены очевидным образом, то доказываемый результат вытекает из теоремы 2. ►

Пример вариационной задачи, удовлетворяющей условиям теоремы 3:

$$f(x) = \int_0^1 [(2 + \sin x) \exp(x')^2 - tx(t)] dt \rightarrow \min,$$

$$x(0) = x(1) = 0.$$

Существование решения данной вариационной задачи вытекает из предложения 1. В рассматриваемом случае интегрант

$$L(t, x, y) = (2 + \sin x) \exp y^2 - tx$$

растёт по  $y$  быстрее любой экспоненты. В качестве  $\psi_R$  можно брать функцию  $\psi(z) = \exp z^2$ .



Теорема легко обобщается на случай, когда  $x$  – вектор-функция со значениями в конечномерном пространстве  $\mathbb{R}^n$ . Естественными аналогами пространств  $\mathcal{L}(\Delta)$ ,  $C^m(\Delta)$ ,  $W^m(\Delta)$  являются пространства вектор-функций со значениями в  $\mathbb{R}^n$ , обозначаемые символами  $\mathcal{L}(\Delta, \mathbb{R}^n)$ ,  $C^m(\Delta, \mathbb{R}^n)$ ,  $W^m(\Delta, \mathbb{R}^n)$ . Они могут рассматриваться как прямые произведения  $n$  экземпляров соответствующих пространств скалярных функций, поэтому наследуют их свойства. Необходимо признать, что перечисленные пространства плохо приспособлены для исследования вариационных задач.

#### 4. Задачи и комментарии.

1. Если в условиях теоремы 2 функции  $A_i$  обладают повышенной гладкостью, то и гладкость обобщённого решения  $x_0$  соответствующего дифференциального уравнения также повышается. Например, если  $A_i \in C^{i+k}$ , то  $x_0 \in C^{2m+k}$ ; в случае бесконечной дифференцируемости  $A_i$  функция  $x_0$  оказывается также бесконечно дифференцируемой. Аналогичное утверждение верно и для решений вариационных задач. *Аккуратно сформулировать и доказать теоремы о повышенной гладкости решений дифференциальных уравнений.* Такого рода результаты полезны при анализе достаточных условий минимума.

2. В общем случае гладкости решений вариационных задач может и не быть. Аналогичные примеры уже приводились в связи с условием Вейерштрасса–Эрдмана (см. п. 2.6). Вот ещё один пример, принадлежащий Д. Гильберту.

$$f(x) = \int_0^1 t^{2/3} (x'(t))^2 dt \rightarrow \min; \quad x(0) = 0, \quad x(1) = 1.$$

*Доказать, что экстремаль  $\hat{x}$  этого функционала, удовлетворяющая краевым условиям в точках 0 и 1, единственна, доставляет глобальный минимум, но не принадлежит классу  $C^1([0, 1])$ .*

3. Показать, что у функционала

$$f(x) = \int_0^1 t^2 (x'(t))^2 dt, \quad x(0) = 0, \quad x(1) = 1$$

*не существует допустимой экстремали даже среди абсолютно непрерывных функций (пример Вейерштрасса).*

4. Рассмотреть вариационную задачу

$$f(x) = \int_0^1 ((1 - x'(t))^2 + x^2(t)) dt \rightarrow \inf, \quad x(0) = 0, \quad x(1) = 0.$$

*Показать, что точная нижняя грань функционала равна нулю, но не достигается ни на одной допустимой функции класса  $KS^1$ .*

Приведённые примеры подводят к восходящей к Д. Гильберту мысли, что каждую вариационную задачу следует рассматривать в своём пространстве.

При этом возникающие пространства и решения могут существенно отличаться от классических образцов (насколько сильно это различие можно понять, заглянув в книгу Л. Янга [30]). Гильберту же принадлежит гипотеза, что при определённых предположениях регулярности каждое обобщённое решение вариационной задачи является на самом деле гладким и степень гладкости определяется лишь степенью гладкости интегранта. Если для одномерных вариационных задач предположение Гильберта справедливо, то для многомерных вариационных задач это неверно; обсуждение данных вопросов и точные результаты можно найти в [22], [27].

5. В этом приложении речь шла о гладкости решений простейшей вариационной задачи. В частности, отсутствовали дополнительные ограничения изопериметрического типа. *Следует разобраться, насколько дополнительные условия могут повлиять на гладкость обобщённых решений.* Трудность заключается в том, что привычный аппарат анализа экстремальных задач в данной ситуации может не сработать ввиду недифференцируемости исследуемых функционалов; эта негладкость может возникнуть даже для функционалов, порожаемых сколь угодно гладкими интегрантами (из-за плохих свойств пространства  $\mathcal{L}^1$  суммируемых функций).

6. Интересен вопрос об экстремальных функционала

$$f(x) = \int_0^1 ((x'(t))^2 + a|x(t)|^p) dt, \quad (18)$$

удовлетворяющих граничным условиям

$$x(0) = 0, \quad x(1) = 0. \quad (19).$$

Здесь  $p > 1$ ,  $a \in \mathbb{R}$ . Уравнение Эйлера для функционала (18) имеет вид

$$-x'' + \frac{ap}{2}|x|^{p-2}x = 0. \quad (20)$$

Допустимые экстремали в данном случае это решения краевой задачи (19, (20)). Доказать следующие утверждения.

$\alpha$ ) при любом  $a \geq 0$  краевая задача (19), (20) имеет единственное решение  $\hat{x} = 0$ , реализующее абсолютный минимум функционала (18) при граничных условиях (19);

$\beta$ ) тот же вывод верен, если  $p = 2$  и  $a > -\pi^2$ ; продумать самостоятельно, что произойдёт в случае  $p = 2$  и  $a = -\pi^2$  или  $a < -\pi^2$ ;

$\gamma$ ) если  $p > 2$ , то при любом  $a < 0$  существует такая последовательность решений  $x_n$  краевой задачи (19), (20), что  $f(x_n) \rightarrow \infty$  при  $n \rightarrow \infty$ ; функционал  $f$  неограничен и сверху, и снизу;

$\delta$ ) если  $1 < p < 2$ , то при любом  $a < 0$  существует такая последовательность решений  $x_n$  краевой задачи (19), (20), что  $f(x_n) < 0$  и  $f(x_n) \rightarrow 0$  при  $n \rightarrow \infty$ .

Данный пример показывает, что для «вылавливания» критических значений функционала следует использовать и отличные от максимального или минимального его значений. Каким образом – это ещё тот вопрос!



# Приложение С

## Метод условного градиента

В этом разделе изучается сходимость метода условного градиента в задаче приближённого нахождения точек минимума функционалов, определённых на ограниченных замкнутых и выпуклых подмножествах рефлексивного пространства. Устанавливаются достаточные условия сходимости, применимые, в частности, к невыпуклым функционалам. Приводятся приложения к задачам оптимального управления и вариационного исчисления. Включают раздел сведения о методе Рунге и несколько посвящённых этому методу задач.

### С.1 Сходимость метода условного градиента

**1. Сведения из функционального анализа.** Напомним некоторые определения функционального анализа [3], [10], [19]. Пусть  $X$  – действительное банахово пространство с нормой  $\|x\|$ . Его можно рассматривать как полное метрическое пространство с метрикой  $\rho(x_1, x_2) = \|x_1 - x_2\|$ . Совокупность  $X^*$  всех линейных непрерывных функционалов на  $X$  (сопряжённое к  $X$  пространство) является банаховым, если задать в  $X^*$  норму

$$\|x^*\|_* = \sup_{\|x\| \leq 1} \langle x, x^* \rangle,$$

где  $\langle x, x^* \rangle$  означает результат применения к  $x$  функционала  $x^*$ . Любой элемент  $x$  из  $X$  порождает линейный функционал на пространстве  $X^*$ . Таким образом, возникает естественное отображение  $\kappa$  пространства  $X$  во второе сопряжённое к  $X$  пространство  $X^{**} = (X^*)^*$ . Если отображение  $\kappa: X \rightarrow X^{**}$  есть биекция, то пространство  $X$  называют рефлексивным.

В банаховом пространстве  $X$  можно рассматривать сильную и слабую сходимости последовательности элементов  $x_n$  этого пространства. Будем говорить, что последовательность  $x_n$  сильно сходится к элементу  $x$  и писать  $x_n \rightarrow x$ , если  $\rho(x_n, x) = \|x_n - x\| \rightarrow 0$  при  $n \rightarrow \infty$ . Скажем, что последовательность  $x_n$  слабо сходится к элементу  $x$ , и станем писать  $x_n \rightharpoonup x$ , если для любого функционала  $x^* \in X^*$  последовательность  $\langle x_n, x^* \rangle$  сходится к  $\langle x, x^* \rangle$  при  $n \rightarrow \infty$ . Всякая

сильно сходящаяся последовательность сходится слабо. Обратное, вообще говоря, неверно.

Сильная топология на пространстве  $X$  – это топология, порождаемая метрикой  $\rho$ . Слабая топология в  $X$  – это слабейшая из топологий, в которой непрерывны все функционалы  $x^*$  из  $X^*$ . Ниже  $Cv(X)$  – совокупность всех непустых выпуклых замкнутых подмножеств пространства  $X$ . В этом определении имеется в виду замкнутость относительно сильной топологии. Однако если  $X$  – рефлексивное пространство, то множество  $Q$  класса  $Cv(X)$  замкнуто относительно и слабой топологии пространства  $X$ .

Стандартным образом определяются замыкания множества  $M \subset X$  относительно слабой и сильной топологии. Например, сильное замыкание множества  $M$  – это множество  $\bar{M}$ , состоящее из элементов  $x$ , для которых существует последовательность элементов  $x_n$ , сильно сходящаяся к элементу  $x$ . Пространство  $X$  называют сепарабельным, если существует такое его счётное подмножество  $M$ , что  $\bar{M} = X$ .

Первостепенное значение имеет для нас следующий критерий слабой компактности.

**Предложение 1.** Пусть  $Q$  – ограниченное непустое выпуклое замкнутое подмножество рефлексивного сепарабельного пространства  $X$ . Тогда из всякой последовательности  $u_n$  элементов множества  $Q$  можно извлечь подпоследовательность  $u_{i_n}$ , слабо сходящуюся к некоторому элементу  $u$  множества  $Q$ .

Отображение  $F: M_1 \rightarrow M_2$  ( $M_1, M_2$  – подмножества банаховых пространств  $X_1, X_2$  соответственно) именуют ограниченным, если для каждого ограниченного в  $X_1$  множества  $M \subset M_1$  область значений  $F(M)$  есть ограниченное подмножество  $X_2$ .

Рассматриваемые ниже функционалы предполагаются действительными. Если  $f: M \rightarrow \mathbb{R}$  – функционал на множестве  $M$ , то  $\arg \min_M f$  – множество точек глобального минимума  $f$  на  $M$ : включение  $x_* \in \arg \min_M f$  эквивалентно неравенству  $f(x_*) \leq f(x) \forall x \in M$ . Стандартным образом определяется производная  $f'(x)$  функционала  $f$ , заданного на некоторой окрестности  $O(x)$  точки  $x$ . Именно, пусть существует такой линейный функционал  $x^*$  класса  $X^*$ , что для любого элемента  $v$  из  $X$  справедливо равенство

$$\lim_{\xi \rightarrow 0} \frac{f(x + \xi v) - f(x)}{\xi} = \langle v, x^* \rangle .$$

Тогда функционал  $f$  называют дифференцируемым в точке  $x$ , а функционал  $x^*$  – производной  $f$  в точке  $x$  и обозначают символом  $f'(x)$ . Через  $C^1(U)$  обозначается совокупность непрерывно дифференцируемых на множестве  $U \subset X$  функционалов,  $C^{1,\nu}(U)$  ( $0 < \nu \leq 1$ ) – часть  $C^1(U)$ , состоящая из функционалов  $f$ , производные которых удовлетворяют неравенству

$$\|f'(u) - f'(v)\|_{X^*} \leq K \|u - v\|_X^\nu ,$$

постоянная  $K$  не зависит от  $u, v$  из  $U$ .

**2. Специальные операторы и функционалы.** В этом и следующем пунктах  $X$  – сепарабельное рефлексивное банахово пространство,  $\|\cdot\|$  и  $\|\cdot\|_*$  – нормы в  $X$  и  $X^*$  соответственно, символы  $\rightarrow$  и  $\rightharpoonup$  означают слабую и сильную сходимости соответственно,  $Q \in Cv(X)$ . Обозначим через  $S(Q)$  класс ограниченных отображений  $F: Q \rightarrow X^*$ , удовлетворяющих следующему условию: для произвольной последовательности  $x_n \in X$ , обладающей свойствами

$$x_n \rightharpoonup x, \quad F(x_n) \rightharpoonup x^*, \quad \overline{\lim}_{n \rightarrow \infty} \langle x_n, F(x_n) \rangle \leq \langle x, x^* \rangle,$$

имеют место соотношения  $x_n \rightarrow x, x^* = F(x)$ . Близкие к  $S(Q)$  классы операторов изучались многими авторами (см., например, [16], [27] и приведённую там литературу). Отметим, что сужение оператора  $F$  класса  $S(X)$  на множество  $Q$  принадлежит классу  $S(Q)$ .

Элемент  $x \in Q$  назовём особой точкой отображения  $F: Q \rightarrow X^*$ , если

$$\langle v - x, F(x) \rangle \geq 0 \quad \forall v \in Q. \quad (1)$$

Соотношения типа (1) называют вариационными неравенствами. Они неоднократно появлялись и в нашем курсе для конечномерного пространства  $X$ , например, как необходимые (а иногда и достаточные) условия минимума функции на выпуклом множестве  $Q$ . В настоящее время теория вариационных неравенств представляет бурно развивающийся раздел математики (библиография может быть найдена в [16]).

Если не оговорено противное, ниже предполагается, что  $Q$  – ограниченное подмножество  $X$ . В этом случае множество  $\mathfrak{K}(F)$  всех особых точек отображения  $F$  класса  $S(Q)$  есть непустой компакт. Непустота  $\mathfrak{K}(F)$  установлена для более широкого, чем  $S(Q)$ , класса операторов. Докажем компактность  $\mathfrak{K}(F)$ . Действительно, пусть  $x_n \in \mathfrak{K}(F), n = 1, 2, \dots$ . Поскольку  $Q$  – ограниченное множество, то последовательности  $x_n, F(x_n)$  ограничены в пространствах  $X$  и  $X^*$  соответственно. Пространство  $X$  рефлексивно, поэтому, не нарушая общности, можно считать, что

$$x_n \rightharpoonup x, \quad F(x_n) \rightharpoonup x^*.$$

Так как  $x_n \in \mathfrak{K}(F)$ , то в силу (1) имеет место неравенство

$$\langle v - x_n, F(x_n) \rangle \geq 0 \quad \forall v \in Q. \quad (2)$$

В частности, неравенство (2) справедливо при  $v = x$ . Это влечёт за собой оценку

$$\overline{\lim}_{n \rightarrow \infty} \langle x_n, F(x_n) \rangle \leq \langle x, x^* \rangle.$$

Поскольку  $F \in S(Q)$ , то  $x_n \rightarrow x, x^* = F(x)$ . Переходя в (2) к пределу при  $n \rightarrow \infty$ , получаем, что  $x \in \mathfrak{K}(F)$ . Отсюда и вытекает компактность множества  $\mathfrak{K}(F)$ .

При любом  $x$  из  $Q$  линейный функционал  $z \rightarrow \langle z, F(x) \rangle$  достигает своего минимума на множестве  $Q$ ; имеет смысл функция  $\zeta(x) = \min_{z \in Q} \{ \langle z - x, F(x) \rangle \}$ .

Очевидно, что  $\zeta(x) \leq 0 \forall x \in Q$ , при этом  $\{x \in Q, \zeta(x) = 0\} = \mathfrak{K}(F)$ .

Последовательность  $Q_n$  ( $n = 0, 1, \dots$ ) множеств класса  $Cv(X)$  назовём исчерпывающей множество  $Q$ , если  $Q_n \subset Q_{n+1}$  и

$$\overline{\bigcup_N Q_N} = Q. \quad (3)$$

Положим  $\zeta_n(x) = \min_{z \in Q_{n+1}} \{ \langle z - x, F(x) \rangle \}$  ( $n = 0, 1, \dots, x \in Q_n$ ).

**Лемма 1.** Пусть  $F \in S(Q)$ . Тогда всякая последовательность  $x_n \in Q_n$  ( $n = 0, 1, \dots$ ), для которой  $\zeta_n(x_n) \rightarrow 0$  при  $n \rightarrow \infty$ , содержит подпоследовательность, сходящуюся к элементу из  $\mathfrak{K}(F)$ .

◀ Будем считать, что  $x_n \rightharpoonup x, F(x_n) \rightharpoonup x^*$ . Фиксируем  $z$  из множества  $Q_N$  ( $N = 1, 2, \dots$ ). Из определения функции  $\zeta_n$  вытекает неравенство

$$\zeta_n(x_n) + \langle x_n, F(x_n) \rangle \leq \langle z, F(x_n) \rangle \quad (n > N),$$

влекущее за собой оценку

$$\overline{\lim}_{n \rightarrow \infty} \langle x_n, F(x_n) \rangle \leq \langle z, x^* \rangle.$$

В силу (3) последняя оценка верна для  $z = x$ . Из установленных свойств последовательности  $x_n$  и включения  $F \in S(Q)$  вытекает сходимость  $x_n \rightarrow x$ , равенство  $x^* = F(x)$  и соотношение (1). В частности,  $x \in \mathfrak{K}(F)$ . ▶

Далее основное внимание уделяется потенциальным отображениям. Пусть  $f$  – дифференцируемый в каждой точке множества  $Q$  функционал. Элемент  $x$  из  $Q$  назовём стационарной точкой функционала  $f$ , если выполняется соотношение (1) с  $F = f'$ . Множество стационарных точек  $f$  обозначим символом  $\mathfrak{K}(f')$ . Справедливо включение  $\arg \min_Q f \subset \mathfrak{K}(f')$ . Элементы множества  $Q \setminus \mathfrak{K}(f')$  назовём регулярными точками  $f$ . Положим  $\zeta(x) = \min_{v \in Q} \langle v - x, f'(x) \rangle$ . Вычисление функции  $\zeta$  в точке  $x$  сводится к решению экстремальной задачи

$$\langle z, f'(x) \rangle \rightarrow \min, \quad z \in Q.$$

Существование решения  $z(x)$  этой задачи очевидно; решение, вообще говоря, может быть неединственным.

Пусть  $Q_n$  ( $n = 0, 1, \dots$ ) – исчерпывающая множество  $Q$  последовательность множеств класса  $Cv(X)$ . Введём функции  $\zeta_n(x) = \min_{z \in Q_{n+1}} \langle z - x, f'(x) \rangle$  ( $x \in Q_n$ ) и отображения  $z_n: Q_n \rightarrow Q_{n+1}$  ( $n = 0, 1, \dots$ ), обладающие свойствами

$$z_n(x) \in Q_{n+1}, \quad \langle z_n(x), f'(x) \rangle \leq \langle z, f'(x) \rangle \quad \forall z \in Q_{n+1}.$$



Указанными свойствами отображения  $z_n: Q_n \rightarrow Q_{n+1}$  ( $n = 0, 1, \dots$ ) определяются, вообще говоря, неоднозначно.

Исзуемая версия метода условного градиента для решения задачи

$$f(x) \rightarrow \min, \quad x \in Q \quad (4)$$

состоит в построении последовательности  $x_n \in Q_n$  ( $n = 0, 1, \dots$ ) по следующим формулам:

$$x_{n+1} = x_n + \alpha_n(z_n(x_n) - x_n), \quad \alpha_n \in \arg \min_{[0,1]} f_n, \quad (5)$$

где  $f_n(\alpha) = f(x_n + \alpha(z_n(x_n) - x_n))$  ( $0 \leq \alpha \leq 1$ ). Элемент  $x_0$  из  $Q_0$  предполагается заданным. На каждом шаге необходимо решать две вспомогательные задачи: 1) задачу о минимизации линейного функционала  $z \rightarrow \langle z, f'(x) \rangle$  на множестве  $Q_{n+1}$ , 2) задачу о минимизации функции одного переменного  $f_n(\alpha)$  на отрезке  $[0, 1]$ . Обе задачи имеют решения. В предшествующих построениях не исключается случай постоянной последовательности  $Q_n = Q$  ( $n = 0, 1, \dots$ ).

Вопрос о сходимости метода условных градиентов будет исследован при дополнительных ограничениях на функционал  $f$  и множество  $Q$ . Обозначим через  $\Lambda_1(Q)$  совокупность дифференцируемых на множестве  $Q$  функционалов, для которых отображение  $F(x) = f'(x)$  принадлежит классу  $S(Q)$ . Для конечномерного пространства  $X$  класс  $\Lambda_1(Q)$  совпадает с классом  $C^1(Q)$ . В бесконечномерном пространстве это уже не так; примеры функционалов класса  $\Lambda_1$  на пространствах Рисса приводятся далее.

Ниже будет использоваться следующий вариант формулы конечных приращений

$$f(v) - f(u) = \langle v - u, f'(w) \rangle, \quad (6)$$

где  $f \in \Lambda_1(Q)$ ,  $u \in Q$ ,  $v \in Q$ ,  $w = u + \theta(v - u)$ ,  $0 < \theta < 1$ . В качестве  $\theta$  можно взять точку из  $(0, 1)$ , в которой реализуется экстремум функции

$$\varphi(t) = f(u + t(v - u)) - t(f(v) - f(u))$$

на отрезке  $[0, 1]$ . Для функционала  $f$  класса  $\Lambda_1(Q)$  задача (1) корректна. Точное определение корректности приводится ниже, а сейчас будет установлено важное само по себе свойство рассматриваемого класса функционалов. Напомним, что функционал  $f: Q \rightarrow \mathbb{R}$  называют слабо полунепрерывным снизу, если для произвольной слабо сходящейся к элементу  $x$  последовательности  $x_n \in Q$  справедливо неравенство

$$\liminf_{n \rightarrow \infty} f(x_n) \geq f(x).$$

**Лемма 2.** Функционал класса  $\Lambda_1(Q)$  слабо полунепрерывен снизу.

◀ В предположении противного существует функционал  $f$  класса  $\Lambda_1(Q)$ , точка  $x$  из  $Q$ , числа  $\varepsilon > 0, \delta > 0$  и последовательность  $x_n \in Q$ , обладающие свойствами

$$x_n \rightharpoonup x, \quad f(x_n) < f(x) - 2\varepsilon, \quad f(y) > f(x) - \varepsilon \quad \forall y \in Q, \|y - x\| < \delta;$$

в частности, если  $y_n = x_n + t(y_n)(x_n - x)$ ,  $\|y_n - x\| = \delta$ ,  $0 < t(y_n) < 1$ , то  $f(y_n) \geq f(x) - \varepsilon > f(x_n) + \varepsilon$ ,  $t(y_n) > \delta_0 > 0$ .

В силу формулы конечных приращений имеем

$$\varepsilon > f(x_n) - f(y_n) = \langle x_n - y_n, f'(z_n) \rangle,$$

где

$$\begin{aligned} z_n &= x + t(z_n)(x_n - x), \quad t(y_n) < t(z_n) < 1, \\ x_n - y_n &= (1 - t(y_n))(x_n - x) = \frac{1 - t(y_n)}{t(z_n)}(z_n - x). \end{aligned}$$

Поэтому  $\langle z_n - x, f'(x_n) \rangle < -\varepsilon t(z_n) < -\varepsilon \delta_0$ . Не нарушая общности, можно считать, что  $f'(z_n) \rightharpoonup z^*$ . Последовательность  $z_n$  обладает свойствами

$$z_n \rightharpoonup x, \quad f'(z_n) \rightharpoonup z^*, \quad \overline{\lim}_{n \rightarrow \infty} \langle z_n - x, f'(z_n) \rangle < -\varepsilon \delta,$$

из которых вытекает оценка

$$\overline{\lim}_{n \rightarrow \infty} \langle z_n, f'(z_n) \rangle < \langle x, z^* \rangle.$$

Последняя оценка противоречит включению  $f \in \Lambda_1(Q)$ . ►

Задачу (4) называют корректной, если 1) множество  $M_f = \arg \min_Q f$  непусто; 2) любая минимизирующая функционал  $f$  последовательность  $x_n$  сходится к множеству  $M_f$ , т.е. если  $x_n \in Q$  и  $f(x_n) \rightarrow a = \min_Q f$ , то  $d_X(x_n, M_f) \rightarrow 0$  при  $n \rightarrow \infty$ ; здесь и далее  $d_X(z, M) = \inf_{y \in M} \|z - y\|$  — расстояние от точки  $z$  до множества  $M$  в метрике пространства  $X$ .

**Теорема 1.** Пусть  $Q$  — непустое ограниченное выпуклое замкнутое подмножество рефлексивного пространства  $X$ , а  $f$  — функционал класса  $\Lambda_1(Q)$ . Тогда задача (4) корректна.

◄ Пусть  $a = \inf_{x \in Q} f(x)$  — точная нижняя грань функционала  $f$  на множестве  $Q$ , последовательность  $x_n$  минимизирует  $f$ , т.е.

$$x_n \in Q, \quad f(x_n) \rightarrow a \quad \text{при } n \rightarrow \infty. \quad (7)$$

Не нарушая общности, можно считать, что  $x_n \rightharpoonup x$  при  $n \rightarrow \infty$ . Так как функционал  $f$  слабо полунепрерывен снизу, то

$$\underline{\lim}_{n \rightarrow \infty} f(x_n) \geq f(x),$$

в частности,  $f(x) \leq a$ . Поскольку по самому определению  $a \leq f(x)$ , то  $f(x) = a$ . Конечность  $a$  и непустота множества  $M_f$  доказаны.

Пусть  $x_n$  — последовательность, обладающая свойствами (7). Будем считать, что  $a < f(x_n)$  и  $x_n \rightharpoonup x$ . Последовательность  $(x_n + x)/2$  также слабо сходится к  $x$ , поэтому

$$\underline{\lim}_{n \rightarrow \infty} f\left(\frac{x_n + x}{2}\right) \geq a = f(x).$$

В силу формулы конечных приращений (6) справедливо равенство

$$f(x_n) - f\left(\frac{x_n + x}{2}\right) = 2^{-1} \langle x_n - x, f'(z_n) \rangle,$$

где  $z_n = x + \theta_n(x_n - x)$ ,  $1/2 < \theta_n < 1$ .

Объединяя полученные соотношения, приходим к неравенству

$$\overline{\lim}_{n \rightarrow \infty} \langle z_n - x, f'(z_n) \rangle \leq 0.$$

Поскольку  $f \in \Lambda_1(Q)$ , то последнее неравенство влечёт сильную сходимость последовательностей  $x_n, z_n$  к  $x$ . Очевидно, что  $d_X(x_n, M_f) \leq \|x_n - x\|$ , следовательно,  $d_X(x_n, M_f) \rightarrow 0$ . ►

Вернёмся к методу условного градиента.

**Теорема 2.** Пусть  $f \in \Lambda_1(Q)$  и справедлива оценка

$$f(x + h) - f(x) \leq \langle h, f'(x) \rangle + \omega(\|h\|), \quad (8)$$

где  $x, x + h \in Q$ , функция  $\omega: [0, \infty) \rightarrow \mathbb{R}$  не зависит от  $x$  из  $Q$  и  $\omega(s) = o(s)$  при  $s \rightarrow 0$ . Пусть  $x_n \in Q_n$  ( $n = 0, 1, \dots$ ) – последовательность элементов, определяемая из соотношений (5), причём  $\zeta_n(x_n) < 0$ .

Тогда  $f(x_{n+1}) < f(x_n)$ ,  $\zeta_n(x_n) \rightarrow 0$  при  $n \rightarrow \infty$ , последовательность  $x_n$  компактна и всякая предельная точка этой последовательности принадлежит  $\mathfrak{K}(f')$ .

◄ Полагая в (8)  $x = x_n, h = \alpha(z_n(x_n) - x_n)$  ( $0 \leq \alpha \leq 1$ ), получаем неравенство

$$f(x_n + \alpha(z_n(x_n) - x_n)) - f(x_n) \leq \alpha \zeta_n(x_n) + \omega(\alpha \|z_n(x_n) - x_n\|). \quad (9)$$

Так как  $\omega(s) = o(s)$ , то для любого  $\delta > 0$  найдётся такое число  $\tau > 0$ , что  $|\omega(\alpha \|z_n(x_n) - x_n\|)| < \frac{\delta \alpha}{2}$  ( $0 \leq \alpha \leq \tau$ ), число  $\tau$  не зависит от выбора начального приближения  $x_0$  и числа  $n$ .

Фиксируем  $\delta > 0$  и соответствующее ему число  $\tau > 0$ . Обозначим через  $\mathbb{N}(\delta)$  совокупность таких натуральных чисел  $n$ , что  $\zeta_n(x_n) < -\delta$ . Если  $n \in \mathbb{N}(\delta)$ , то из (9) вытекает оценка

$$f(x_n + \alpha(z_n(x_n) - x_n)) - f(x_n) \leq -\frac{\delta \alpha}{2} \quad (0 \leq \alpha \leq \tau),$$

в частности,  $f(x_{n+1}) - f(x_n) < \frac{\delta \alpha}{2}$ . Функционал  $f$  ограничен снизу, поэтому множество  $\mathbb{N}(\delta)$  конечно при любом  $\delta > 0$ . Неравенство  $f(x_{n+1}) < f(x_n)$  и сходимость  $\zeta_n(x_n) \rightarrow 0$  при  $n \rightarrow \infty$  доказаны, остальные утверждения вытекают из леммы 1. ►

**Следствие.** Если в условиях теоремы 2 стационарная точка  $x_*$  функционала  $f$  единственна, то  $x_n \rightarrow x_*$  при  $n \rightarrow \infty$  и  $x_* \in M_f$ .

Скорость сходимости  $x_n$  к  $x$  может зависеть и от свойств функционала  $f$ , и от скорости приближения множества  $Q$  последовательностью множеств  $Q_n$ . Для выполнения оценки (8) достаточно, чтобы отображение  $F = f': Q \rightarrow X^*$  было равномерно непрерывным. В теореме 2 исключается случай  $\zeta_n(x_n) = 0$  при некотором  $n$ . В этом случае метод условного градиента конечен, т.е. останавливается после конечного числа шагов. Обсуждение данного вопроса можно найти в [13, с. 189–190].

**3. О скорости сходимости метода условного градиента.** В этом разделе метод условного градиента применяется к дифференцируемым и выпуклым на множестве  $Q$  функционалам, производные которых ограничены на множестве  $Q$ . Функционалы данного класса удовлетворяют условию Липшица и слабо полунепрерывны снизу, множества точек минимума и стационарных точек совпадают. Как и в предыдущем разделе,  $\zeta(x) = \min_{v \in Q} \langle v - x, f'(x) \rangle$ ,  $a = \min_Q f$ . Выпуклость функции  $f$  влечёт неравенство

$$f(v) - f(x) \geq \langle v - x, f'(x) \rangle \quad \forall v \in Q,$$

непосредственным следствием которого является оценка

$$-\zeta(x) \geq f(x) - a \quad \forall x \in Q.$$

Для анализа сходимости метода условного градиента используется

**Предложение 1. [13, с. 52].** Пусть  $w_n > 0$  и

$$w_{n+1} \leq w_n - cw_n^{1+\beta} \quad (c > 0, \quad \beta > 0, \quad n = 0, 1, \dots). \quad (10)$$

Тогда

$$w_n \leq w_0(1 + nc\beta w_0^\beta)^{-1/\beta}. \quad (11)$$

Обозначим через  $\Lambda_\gamma(Q)$  ( $1 < \gamma < 2$ ) совокупность выпуклых и дифференцируемых на множестве  $Q$  функционалов, удовлетворяющих оценкам

$$\|f'(x)\|_* \leq K_1, \quad f(x+h) - f(x) \leq \langle h, f'(x) \rangle + K_2\|h\|^\gamma \quad (12)$$

с некоторыми постоянными  $K_1, K_2$ , не зависящими от  $x$  и  $x+h$  из  $Q$ . Оценки (12) имеют место для функционалов  $f$  класса  $C^{1,\nu}(Q)$  ( $\nu \geq \gamma - 1$ ).

Сопоставим регулярной точке  $x$  из  $Q$  элемент  $z(x)$  из  $\arg \min_Q f'(x)$ . Таким образом,  $\zeta(x) = \langle z(x) - x, f'(x) \rangle < 0$ . Полагая в (12)  $h = \alpha(z(x) - x)$ , приходим к соотношению

$$f(x + \alpha(z(x) - x)) - f(x) \leq \alpha\zeta(x) + K_2\alpha^\gamma\|z(x) - x\|^\gamma \leq \alpha\zeta(x) + K_2D^\gamma\alpha^\gamma,$$

где  $D = \sup_{u,v \in Q} \|u - v\|$  — диаметр множества  $Q$ . Из этого соотношения вытекает оценка

$$\min_{0 \leq \alpha \leq 1} [f(x + \alpha(z(x) - x)) - f(x)] \leq \min_{0 \leq \alpha \leq 1} (\alpha\zeta(x) + K_2D^\gamma\alpha^\gamma) \leq -c_1|\zeta(x)|^{\gamma/(\gamma-1)}, \quad (13)$$

постоянная  $c_1 > 0$  не зависит от выбора регулярной точки  $x$  из  $Q$ .

**Теорема 3.** Пусть  $f \in \Lambda_\gamma(Q)$ ,  $x_0 \in Q$ ,  $x_n (n = 1, 3, \dots)$  – последовательность из соотношений (5), причём  $Q_n = Q$ ,  $\zeta(x_n) < 0 (n = 0, 1, \dots)$ . Тогда для последовательности  $w_n = f(x_n)$  – а верна оценка (11), где  $\beta(\gamma - 1) = 1, c = c_1$  из оценки (10).

◀ Из оценки (13) следует неравенство

$$f(x_{n+1}) - f(x_n) \leq -c_1 |\zeta(x_n)|^{\gamma\beta}.$$

С учётом (10) получаем, что последовательность  $w_n = f(x_n)$  – а удовлетворяет оценке (11) с  $\beta(\gamma - 1) = 1, c = c_1$ . Теперь доказываемое утверждение вытекает из предложения 1. ▶

Для более узкого класса выпуклых множеств  $Q$  и выпуклых функционалов  $f$  теорема 3 допускает некоторое усиление. Множество  $Q \subset X$  назовём  $r$ -выпуклым ( $2 \leq r < \infty$ ), если существует такое  $\delta > 0$ , что для любых  $x, y$  из  $Q$  и  $z$  из  $X$  ( $\|z\| \leq \delta \|x - y\|^r$ ) точка  $\frac{x+y}{2} + z$  принадлежит множеству  $Q$ . Например, если  $Q$  – шар в пространстве  $L^p$  ( $1 < p < \infty$ ), то можно положить  $r = \max\{2, p\}$ .

**Предложение 2.** Пусть  $Q$  –  $r$ -выпуклое подмножество  $X$ ,  $f$  – выпуклый функционал  $\|f'(x)\|_* \geq \varepsilon_0 > 0 \quad \forall x \in Q$  и  $f(x_*) = a = \min_Q f$ . Тогда

$$-\zeta(x) \geq 2\delta\varepsilon_0 \|z(x) - x\|^r, \quad f(x) - a \geq 2\delta\varepsilon_0 \|z(x) - x\|^r. \quad (14)$$

Доказательство предоставляется провести читателю; его можно найти, например, в [20], [26].

Для функционалов класса  $\Lambda_1(Q)$  неравенство  $\|f'(x)\|_* \geq \varepsilon_0 \quad \forall x \in Q$  с некоторым  $\varepsilon_0 > 0$  следует из условия  $f'(x) \neq 0 \quad \forall x \in Q$ . Достаточно заметить, что если  $x_n \rightarrow x$ ,  $f'(x_n) \rightarrow 0$ , то  $x \in Q$  и  $f'(x) = 0$ .

**Теорема 4.** Пусть функционал  $f$ , последовательность  $x_n (n = 0, 1, \dots)$  удовлетворяют условиям теоремы 3 и предложения 2. Тогда

1) если  $r > \gamma$ , то для последовательности  $w_n = f(x_n)$  – а имеет место оценка (11), где  $\beta = \frac{r - \gamma}{r(\gamma - 1)}$ ,  $c$  – положительная постоянная;

2) если  $r = \gamma = 2$ , то  $w_n \leq c_0 q_0^n$  ( $c_0 < \infty, 0 < q_0 < 1$ ).

◀ Пусть вначале  $r > \gamma$ . Установим аналог оценки (13). Применяя (12), (14), имеем последовательно

$$f(x + \alpha(z(x) - x)) - f(x) \leq \alpha\eta(x) + K_2\alpha^\gamma \|z(x) - x\|^\gamma \leq \alpha\eta(x) + K_3\alpha^\gamma |\eta(x)|^{\gamma/r}$$

с некоторой постоянной  $K_3$ . Отсюда вытекает соотношение

$$\min_{0 \leq \alpha \leq 1} [f(x + \alpha(z(x) - x)) - f(x)] \leq$$

$$\leq \min_{0 \leq \alpha \leq 1} [\alpha \zeta(x) + K_3 \alpha^\gamma |\eta(x)|^{\gamma/r}] \leq -c_2 |\eta(x)|^{1+\beta}, \quad (15)$$

здесь  $\beta = \frac{r\gamma}{r(\gamma-1)}$ ,  $c_2$  – положительная постоянная, не зависящая от  $x$ .

Соотношение (15) влечёт за собой неравенство

$$f(x_{n+1}) - f(x_n) \leq -c_2 |\zeta(x_n)|^{1+\beta}.$$

Так как  $|\eta(x_n)| \geq f(x_n) - a$ , то последовательность  $w_n = f(x_n) - a$  удовлетворяет оценке (11) с  $\beta = \frac{r\gamma}{r(\gamma-1)}$ ,  $c = c_2$ . Теперь первое утверждение теоремы вытекает из предложения 1.

Если  $\gamma = r = 2$ , то оценка (15) имеет место с  $\beta = 0$ . Отсюда получаем, что последовательность  $w_n = f(x_n) - a$  удовлетворяет неравенству  $w_{n+1} \leq w_n - c_2 w_n$ . Из этой оценки очевидным образом следует второе утверждение теоремы. ►

Комбинируя теорему 4 с неравенством (14), можно получить оценки скорости сходимости последовательности  $x_n$  к точке  $x_*$  минимума функционала  $f$  на множестве  $Q$ . В частности, при  $\gamma = r = 2$  метод условного градиента сходится со скоростью геометрической прогрессии, т.е.  $\|x_n - x_*\| \leq \text{const } q^n$  ( $0 < q < 1$ ).

Теоремы о сходимости сохраняются и при отличных от (5) способах выбора параметров  $\alpha_n$ . Обсуждение такого рода вопросов можно найти в [13, с. 77–80]; [20, с. 89–93].

## С.2 Приложения к задачам оптимального управления

**1. Оператор суперпозиции.** Для перехода от результатов общего характера к задачам оптимального управления нам потребуется ряд сведений об операторе суперпозиции. Большинство формулируемых в этом пункте утверждений хорошо известно, поэтому приводится без доказательства. Вместе с тем центральные результаты – леммы 1, 2 – доказываются полностью ввиду их принципиальной важности.

Ниже  $\Delta = [t_0, t_1]$  – отрезок действительной оси  $\mathbb{R}$ ,  $mes_1 e$  – одномерная мера Лебега множества  $e \in \mathbb{R}$ ,  $\mathbb{R}^k$  –  $k$ -мерное арифметическое пространство, стандартным образом наделённое структурой евклидова пространства. Обозначим через  $\mathcal{L}_0(\Delta, \mathbb{R}^k)$  совокупность измеримых почти всюду конечных функций  $\varphi: \Delta \rightarrow \mathbb{R}^k$ . Как обычно, эквивалентные функции отождествляются. В  $\mathcal{L}_0(\Delta, \mathbb{R}^k)$  вводится сходимость по мере. Напомним, что последовательность  $\varphi_n$  из  $\mathcal{L}_0(\Delta, \mathbb{R}^k)$  сходится по мере к функции  $\varphi$  того же класса, если для любого положительного  $\sigma$  лебегова мера множества  $E_n(\sigma) = \{t \in \Delta, |\varphi_n(t) - \varphi(t)| > \sigma\}$  стремится к нулю при  $n \rightarrow \infty$ . В этом случае, краткости ради, будем писать  $\varphi_n \Rightarrow \varphi$ .

В пространстве  $\mathcal{L}_0(\Delta, \mathbb{R}^k)$  рассматривается сходимость почти всюду. Из сходимости почти всюду вытекает её сходимость по мере. Обратное неверно. Однако из каждой последовательности, сходящейся по мере, можно извлечь подпоследовательность, сходящуюся почти всюду (теорема Ф. Рисса).

Пусть  $p \in [1, \infty)$ . Через  $\mathcal{L}^p(\Delta, \mathbb{R}^k)$  обозначается часть класса  $\mathcal{L}_0(\Delta, \mathbb{R}^k)$ , состоящая из функций  $\varphi$ , для которых имеет смысл и конечна норма

$$\|\varphi\|_p = \left( \int_{\Delta} |\varphi(t)|^p dt \right)^{1/p}.$$

Вместе с пространством  $\mathcal{L}^p(\Delta, \mathbb{R}^k)$  рассматривают пространство  $\mathcal{L}^\infty(\Delta, \mathbb{R}^k)$  измеримых функций. Норма в этом пространстве определяется равенством

$$\|\varphi\|_\infty = \text{vrai} \max_{t \in \Delta} |\varphi(t)| = \lim_{p \rightarrow \infty} \|\varphi\|_p.$$

Пространства  $\mathcal{L}^p(\Delta, \mathbb{R}^k)$  с введёнными таким образом нормами полны. Сопряжённое к пространству  $\mathcal{L}^p(\Delta, \mathbb{R}^k)$  при  $p \in [1, \infty)$  можно отождествить с пространством  $\mathcal{L}^q(\Delta, \mathbb{R}^k)$ , где

$$q = \begin{cases} \frac{p}{p-1} & , \quad \text{если } 1 < p < \infty \\ \infty & , \quad \text{если } p = 1. \end{cases}$$

При  $p \in [1, \infty)$  пространства  $\mathcal{L}^p(\Delta, \mathbb{R}^k)$  сепарабельны, а для  $p \in (1, \infty)$  – и рефлексивны.

Отметим известное неравенство Гёльдера

$$\left| \int_{\Delta} (\varphi(t), \psi(t)) dt \right| \leq \|\varphi\|_p \|\psi\|_q.$$

Сходимость в  $\mathcal{L}^p(\Delta, \mathbb{R}^k)$  влечёт за собой сходимость по мере. Обратное неверно. Справедлив

**Критерий Валле–Пуссена.** Для того чтобы последовательность  $\varphi_n$  сходилась к  $\varphi$  в метрике пространства  $\mathcal{L}^p(\Delta, \mathbb{R}^k)$  ( $1 \leq p < \infty$ ), необходимо и достаточно, чтобы выполнялись условия: 1°  $\varphi_n \Rightarrow \varphi$ ; 2° интегралы от функций  $|\varphi_n(t)|^p$  равномерно абсолютно непрерывны, т.е. справедлива оценка

$$\int_e |\varphi_n(t)|^p dt \leq \omega_1(\text{mes}_1 e),$$

в которой функция  $\omega_1$  не зависит от  $n$  и  $\omega_1(s) \rightarrow 0$  при  $s \rightarrow 0$ .

Пусть  $\mathbb{R}^l$  –  $l$ -мерное арифметическое пространство,  $\Delta \times \mathbb{R}^l \times \mathbb{R}^k$  – прямое произведение  $\Delta, \mathbb{R}^l, \mathbb{R}^k$ ,  $\Phi: \Delta \times \mathbb{R}^l \times \mathbb{R}^k \rightarrow \mathbb{R}^k$  – непрерывное отображение, удовлетворяющее оценке роста

$$|\Phi(t, \eta, \xi)| \leq \psi_0(|\eta|)(1 + |\xi|^{p-1}), \quad (1)$$

в которой  $\psi_0$  – непрерывная положительная функция действительного переменного. Из оценки (1) вытекает, что если  $1 < p < \infty$ ,  $v \in C(\Delta, \mathbb{R}^l)$ ,  $u \in \mathcal{L}^p(\Delta, \mathbb{R}^k)$ , то функция  $w(t) = \Phi(t, v(t), u(t))$  принадлежит пространству  $\mathcal{L}^q(\Delta)$ . Соответствие, сопоставляющее паре  $(v, u) \in C(\Delta, \mathbb{R}^l) \times \mathcal{L}^p(\Delta, \mathbb{R}^k)$  функцию  $w = \Phi(\cdot, v(\cdot), u(\cdot))$  из  $\mathcal{L}^q(\Delta, \mathbb{R}^k)$ , обозначим символом  $\mathcal{F}$  и будем называть оператором суперпозиции.

**Лемма 1.** *Оператор  $\mathcal{F}$  действует и непрерывен из  $C(\Delta, \mathbb{R}^l) \times \mathcal{L}^p(\Delta, \mathbb{R}^k)$  в  $\mathcal{L}^q(\Delta, \mathbb{R}^k)$ .*

◀ Пусть  $v_n \in C(\Delta, \mathbb{R}^l)$ ,  $u_n \in \mathcal{L}^p(\Delta, \mathbb{R}^k)$ ,  $w_n = \mathcal{F}(v_n, u_n)$ ,  $w = \mathcal{F}(v, u)$  и

$$v_n \rightarrow v \quad \text{в } C(\Delta, \mathbb{R}^l), \quad u_n \rightarrow u \quad \text{в } \mathcal{L}^p(\Delta, \mathbb{R}^k).$$

Последовательность  $u_n$  сходится к  $u$  по мере. В силу теоремы Рисса из каждой последовательности  $u_{n_j}$  ( $j = 1, 2, \dots$ ) можно выбрать подпоследовательность  $u_{n_{j'}}$ , сходящуюся к  $u$  почти всюду. Из непрерывности функции  $\Phi$  вытекает, что последовательность  $\mathcal{F}(u_{n_{j'}})$  также сходится почти всюду к функции  $w$ . Это значит, что из каждой последовательности  $w_n$  можно выбрать подпоследовательность, сходящуюся к  $w$  почти всюду. Поэтому последовательность  $w_n$  сходится к  $w$  по мере.

Из неравенства (1) вытекает оценка

$$|w_n(t)|^q \leq c_0(1 + |u_n(t)|^p) \quad (2)$$

с не зависящей от  $n$  постоянной  $c_0$ . В условиях доказываемой леммы  $u_n \rightarrow u$  в пространстве  $\mathcal{L}^p(\Delta, \mathbb{R}^k)$ , поэтому интегралы от функций  $|u_n|^p$  равномерно абсолютно непрерывны. В силу оценки (2) аналогичным свойством обладают и интегралы от функций  $|w_n|^q$ . Теперь сходимость  $w_n \rightarrow w$  в пространстве  $\mathcal{L}^q(\Delta, \mathbb{R}^k)$  следует из критерия Валле–Пуссена. ▶

Отображение  $\Phi(t, x, u)$  ( $t \in \Delta, x \in \mathbb{R}^l, u \in \mathbb{R}^k$ ) назовём строго монотонным по  $u$ , если

$$(\Phi(t, x, u) - \Phi(t, x, \bar{u}), u - \bar{u}) > 0$$

для произвольных  $t \in \Delta, x \in \mathbb{R}^l$  и любых не равных между собой  $u, \bar{u}$  из  $\mathbb{R}^k$ . Рассмотрим отображение  $\Phi$ , удовлетворяющее оценке снизу

$$(\Phi(t, x, u), u) \geq \psi_1(|x|)|u|^p - \psi_2(|x|), \quad (3)$$

в котором  $\psi_1, \psi_2$  – непрерывные положительные функции одного переменного.

**Лемма 2.** *Пусть непрерывное отображение*

$$\Phi: \Delta \times \mathbb{R}^l \times \mathbb{R}^k \rightarrow \mathbb{R}^k$$

*удовлетворяет условиям роста (1), (3) и строго монотонно по  $u \in \mathbb{R}^k$ . Пусть  $u_n \in \mathcal{L}^p(\Delta, \mathbb{R}^k)$ ,  $v_n \in C(\Delta, \mathbb{R}^l)$ ,  $w_n = \mathcal{F}(v_n, u_n) \in \mathcal{L}^q(\Delta, \mathbb{R}^k)$  и*

$$u_n \rightharpoonup u, \quad v_n \rightarrow v, \quad w_n \rightharpoonup w, \quad (4)$$



$$\overline{\lim}_{n \rightarrow \infty} \int_I (u_n(t), w_n(t)) dt \leq \int_{\Delta} (u(t), w(t)) dt. \quad (5)$$

Тогда  $u_n \rightarrow u$  и  $w = \mathcal{F}(v, u)$ .

► Положим  $z_n = \mathcal{F}(v_n, u)$ . В силу леммы 1 и (4) последовательность  $z_n$  сходится к  $z = \mathcal{F}(v, u)$  в метрике пространства  $\mathcal{L}^q(\Delta, \mathbb{R}^k)$ .

Вначале установим сходимость  $u_n \Rightarrow u$ . Введём в рассмотрение функции

$$h_n(t) = (u_n(t) - u(t)) \cdot (w_n(t) - z_n(t)), \quad r_n(t) = (u_n(t) - u(t)) \cdot w_n(t),$$

$$s_n(t) = (u(t) - u_n(t)) \cdot z_n(t) \quad (n \in \mathbb{N}, \quad t \in \Delta).$$

Очевидно, что функции  $h_n(t)$  неотрицательны и  $h_n = r_n + s_n$ . Поскольку  $z_n \rightarrow z$  в  $\mathcal{L}^q(\Delta, \mathbb{R}^k)$ , а  $u_n \rightarrow u$ , то

$$\lim_{n \rightarrow \infty} \int_{\Delta} s_n(t) dt = 0.$$

Справедливы соотношения

$$\begin{aligned} \overline{\lim}_{n \rightarrow \infty} \int_{\Delta} r_n(t) dt &= \overline{\lim}_{n \rightarrow \infty} \int_{\Delta} (u_n(t) - u(t)) \cdot w_n(t) dt = \overline{\lim}_{n \rightarrow \infty} \int_{\Delta} u_n(t) \cdot w_n(t) dt - \\ &- \lim_{n \rightarrow \infty} \int_{\Delta} u(t) \cdot w_n(t) dt \leq \int_{\Delta} u(t) \cdot w(t) dt - \int_{\Delta} u(t) \cdot w(t) dt = 0. \end{aligned}$$

Здесь используется определение последовательности  $r_n$ , элементарные свойства пределов последовательностей, неравенство (5) и сходимость  $w_n \rightarrow w$ . Теперь ясно, что  $h_n \rightarrow 0$  в  $\mathcal{L}(\Delta)$ . В частности,  $h_n \Rightarrow 0$  и интегралы от функций  $h_n$  равномерно абсолютно непрерывны.

Некоторая подпоследовательность последовательности  $h_n$  сходится к 0 почти всюду. Производя перенумерацию, можно указать такое множество  $\Delta_1$ , что  $\Delta_1 \subset \Delta$ ,  $mes_1 \Delta_1 = mes_1 \Delta$  и для всех  $t \in \Delta_1$  выполняются условия

$$h_n(t) \rightarrow 0, \quad |u(t)| < \infty.$$

Из этих условий и (3) вытекает ограниченность последовательности  $u_n(t)$  для каждого  $t$  из множества  $\Delta_1$ :  $|u_n(t)| < R(t) < \infty \forall t \in \Delta_1$ . Пусть  $u^*(t)$  — частичный предел последовательности  $u_n(t)$  ( $t \in \Delta_1$ ). Так как  $h_n(t) \rightarrow 0$  при  $t \in \Delta_1$ , то

$$\begin{aligned} &(u^*(t) - u(t)) \cdot (\Phi(t, v(t), u^*(t)) - \Phi(t, v(t), u(t))) = \\ &= \lim_{n \rightarrow \infty} (u_n(t) - u(t)) \cdot (\Phi(t, v_n(t), u_n(t)) - \Phi(t, v_n(t), u(t))) = 0. \end{aligned}$$

Поскольку отображение  $\Phi$  строго монотонно по  $u$ , то  $u^*(t) = u(t)$ . Таким образом,  $u_n(t) \rightarrow u(t)$  для  $t$  из  $\Delta_1$ .

Проведённые рассуждения показывают, что для любого бесконечного множества  $\mathbb{N}_0$  натуральных чисел существует такая возрастающая последовательность  $i_n \in \mathbb{N}_0$ , для которой  $u_{i_n}(t)$  почти всюду сходится к  $u(t)$ . Это возможно, когда последовательность  $u_n$  сходится к  $u$  по мере.

Как отмечалось выше,  $z_n \rightarrow z$  в  $\mathcal{L}^q(\Delta, \mathbb{R}^k)$ , а последовательность  $u - u_n$  ограничена в пространстве  $\mathcal{L}^p(\Delta, \mathbb{R}^k)$ . В силу неравенства Гёльдера справедлива оценка

$$\int_e |s_n(t)| dt \leq \left( \int_e |z_n(t)|^q dt \right)^{1/q} \|u_n - u\|_p,$$

из которой вытекает, что последовательность  $s_n$  имеет равностепенные интегралы. Поскольку  $h_n \rightarrow 0$  в  $\mathcal{L}(\Delta)$ , а  $h_n = r_n + s_n$ , то и последовательности  $h_n, r_n$  также имеют равностепенно абсолютно непрерывные интегралы. Из (17) вытекает оценка

$$|r_n(t)| \geq c_1 |u_n(t)|^p - c_2,$$

где  $c_1, c_2$  – положительные постоянные. Поэтому интегралы от функций  $|u_n(t)|^p$  равностепенно абсолютно непрерывны.

В силу критерия Валле–Пуссена  $u_n \rightarrow u$  в пространстве  $\mathcal{L}^p(\Delta, \mathbb{R}^k)$ . Теперь очевидны равенства  $w = \mathcal{F}(v, u) = z$ . ►

Внимательный читатель заметит, что требование непрерывности отображения  $\Phi$  по совокупности переменных можно заменить менее ограничительным условием Каратеодори: при любых  $(x, u)$  из  $\mathbb{R}^l \times \mathbb{R}^k$  отображение  $\Phi(\cdot, x, u): \Delta \rightarrow \mathbb{R}^k$  измеримо, а почти при всех  $t$  из  $\Delta$  отображение  $\Phi(t, \cdot, \cdot): \mathbb{R}^l \times \mathbb{R}^k \rightarrow \mathbb{R}^k$  непрерывно.

**2. Производные специальных функционалов.** Рассматривается управляемый объект, эволюция которого описывается соотношениями

$$\frac{dx}{dt} = A(t)x + B(t)u, \quad x(t_0) = x_0. \quad (6)$$

Здесь  $x = (x_i)$  ( $i = 1, \dots, N$ ) – фазовый вектор,  $u = (u_j)$  ( $j = 1, \dots, m$ ) – управление,  $A(t), B(t)$  – кусочно-непрерывные на отрезке  $\Delta = [t_0, t_1]$  матрицы размеров  $N \times N, N \times m$  соответственно,  $x_0$  – фиксированный элемент из  $\mathbb{R}^N$ .

Каждому управлению  $u$  класса  $U_p = \mathcal{L}^p(\Delta, \mathbb{R}^m)$  ( $1 < p < \infty$ ) соответствует единственное решение  $x = W(u)$  задачи Коши (6), называемое фазовой траекторией. Это решение определяется равенством

$$x(t) = V(t)x_0 + \int_{t_0}^t V(t)V^{-1}(s)B(s)u(s) ds. \quad (7)$$

Здесь  $V(t)$  – фундаментальная матрица решений однородной системы

$$\frac{dy}{dt} = A(t)y,$$

совпадающая при  $t = t_0$  с единичной матрицей  $\mathcal{E}_n$ . Из равенства (7) вытекает усиленная непрерывность отображения  $W: U_p \rightarrow C(\Delta, \mathbb{R}^N)$ : если последовательность  $u_n$  из  $U_p$  слабо сходится к  $u$ , то  $W(u_n) \rightarrow W(u)$  в пространстве  $C(\Delta, \mathbb{R}^N)$ . Критерий качества управления задаётся в форме Больца

$$J(u) = \int_{t_0}^{t_1} L(t, x(t), u(t)) dt + g(x(t_1));$$

здесь и далее  $u \in U_p, x = W(u)$ .

**Предложение 1** [17]. с. 93–95. Пусть функции  $L, L'_x, L'_u$  непрерывны по совокупности переменных ( $t \in \Delta, x \in \mathbb{R}^N, u \in \mathbb{R}^m$ ) и удовлетворяют неравенствам

$$|L(t, x, u)| + |L'_x(t, x, u)| + (1 + |u|)|L'_u(t, \eta, \xi)| \leq L_0(|x|)(1 + |u|^p),$$

где  $L_0: [0, \infty) \rightarrow (0, \infty)$  – непрерывная функция.

Тогда  $J \in C^1(U_p)$  и справедливо равенство

$$\langle v, J'(u) \rangle = \int_{t_0}^{t_1} (a(t)h(t) + b(t)v(t)) dt + g'(x(t_1))h(t_1), \quad (8)$$

в котором  $v$  – произвольный элемент из  $U_p, g \in C^1(\mathbb{R}^N)$ ,

$$a(t) = L'_x(t, x(t), u(t)), \quad b(t) = L'_u(t, x(t), u(t)),$$

$h(t)$  – решение задачи Коши

$$\frac{dh}{dt} = A(t)h + B(t)v, \quad h(t_0) = 0. \quad (9)$$

Если функции  $L(t, \cdot, \cdot): \mathbb{R}^N \times \mathbb{R}^m \rightarrow \mathbb{R}$  ( $t \in I$ ),  $g: \mathbb{R}^N \rightarrow \mathbb{R}$  выпуклы, то функционал  $J: U_p \rightarrow \mathbb{R}$  также выпуклый (см. [17], с. 101). При дополнительных предположениях о гладкости функций  $L, g$  можно установить включение  $J \in C^{1,\nu}(U_p)$  (см., например, [17, с. 97–99]), простой пример такого рода приводится ниже.

Сформулируем ограничения на функции  $L, g$ , обеспечивающие включение  $J \in \Lambda_1(U_p)$ .

**Лемма 3.** Пусть выполнены условия предложения 1. Пусть при любых  $(t, x)$  из  $\Delta \times \mathbb{R}^N$  функция  $L(t, x, \cdot): \mathbb{R}^m \rightarrow \mathbb{R}$  строго выпукла и  $L(t, 0, u) \geq \kappa_0|u|^p - \kappa_1$  ( $\kappa_0 > 0, \kappa_1 > 0$ ).

Тогда  $J \in \Lambda_1(U_p)$ .

◀ Достаточно установить, что отображение  $F(u) = J'(u)$  ( $u \in U_p$ ) принадлежит классу  $S(U_p)$ . Пусть последовательность  $u_n \in U_p$  обладает свойствами

$$u_n \rightharpoonup u, \quad F(u_n) \rightharpoonup u^*, \quad \overline{\lim}_{n \rightarrow \infty} \langle u_n, F(u_n) \rangle \leq \langle u, u^* \rangle. \quad (10)$$

Положим  $x_n = W(u_n)$ ,  $a_n(t) = L'_x(t, x_n(t), u_n(t))$ ,  $b_n(t) = L'_u(t, x_n(t), u_n(t))$ . Из формулы (8) вытекает равенство

$$\langle v, F(u_n) \rangle = \int_{t_0}^{t_1} (a_n(t)h(t) + b_n(t)v(t)) dt + g'(x_n(t_1))h(t_1),$$

где  $v \in U_p$ ,  $h$  – решение задачи Коши (9). Последовательность  $x_n = W(u_n)$  сходится к  $x = W(u)$  в пространстве  $C(\Delta, \mathbb{R}^N)$ , последовательности  $a_n, b_n$  ограничены в пространствах  $\mathcal{L}(\Delta, \mathbb{R}^N)$ ,  $\mathcal{L}^q(I, \mathbb{R}^m)$  соответственно. Без ограничения общности можно считать, что последовательность  $b_n$  слабо сходится в  $\mathcal{L}^q(\Delta, \mathbb{R}^m)$  к некоторой функции  $b_0$ . Из (24) следует неравенство

$$\overline{\lim}_{n \rightarrow \infty} \langle u_n - u, F(u_n) \rangle \leq 0,$$

эквивалентное в силу формул (8), (9) неравенству

$$\overline{\lim}_{n \rightarrow \infty} \left( \int_{t_0}^{t_1} (a_n(t)h_n(t) + b_n(t)(u_n(t) - u(t))) dt + g'(x_n(t_1))h_n(t_1) \right) \leq 0,$$

где  $h_n$  – решение задачи Коши (9) с  $v = u_n - u$ . Так как  $h_n \rightarrow 0$  в  $C(I, \mathbb{R}^N)$ , последовательность  $a_n$  ограничена в  $\mathcal{L}(\Delta, \mathbb{R}^N)$ , а  $b_n \rightharpoonup b_0$ , то последнее неравенство влечёт за собой оценку

$$\overline{\lim}_{n \rightarrow \infty} \int_{t_0}^{t_1} b_n(t)u_n(t) dt \leq \int_{t_0}^{t_1} b_0(t)u(t) dt.$$

Согласно лемме 2 из этой оценки вытекает сходимость  $u_n \rightarrow u$  в метрике пространства  $U_p$  и равенство  $b_0(t) = L'_u(t, x(t), u(t)) = b(t)$ . Поскольку  $x_n \rightarrow x$  в  $C(\Delta, \mathbb{R}^N)$ ,  $u_n \rightarrow u$  в  $U_p$ , то  $a_n \rightarrow a = L'_x(\cdot, x, u)$  в метрике  $\mathcal{L}(\Delta, \mathbb{R}^N)$ . Установленные свойства последовательностей  $a_n, b_n$  и их пределов  $a, b$  означают, что  $u^* = F(u)$ . ►

Условиям леммы 3 удовлетворяют функции  $L(t, x, u) = |u|^p$  ( $1 < p < \infty$ ),  $g \in C^{1,\nu}(\mathbb{R}^N)$ . Функционал качества, соответствующий такой паре  $L, g$ , принадлежит  $C^{1,\nu}(U_p)$ , если  $\nu + 1 \leq \min\{2, p\}$ ,  $g \in C^{1,\nu}(\mathbb{R}^N)$ .

Так как сопряжённое к пространству  $U_p = \mathcal{L}^p(\Delta, \mathbb{R}^m)$  можно отождествить с пространством  $\mathcal{L}^q(\Delta, \mathbb{R}^m)$  ( $q = \frac{p}{p-1}$ ), то равенство (8), определяющее производную  $J'(u)$ , можно записать в виде

$$\langle v, J'(u) \rangle = \int_{t_0}^{t_1} (\mathbf{F}(u)(t), v(t)) dt,$$

где  $\mathbf{F}(u)$  – элемент пространства  $U_p^* = \mathcal{L}^q(\Delta, \mathbb{R}^m)$ ,  $(\cdot, \cdot)$  – скалярное произведение в  $\mathbb{R}^m$ . Иногда производную  $J'(u)$  отождествляют с вектор-функцией  $\mathbf{F}(u)$  (см. [16, с. 298]; [17, с. 93]; [20, с. 129]); там же описаны способы нахождения вектор-функции  $\mathbf{F}(u)$ .

Задачи минимизации функционала Больца на множестве  $Q \subset U_p$  записывают в виде

$$J(u) = \int_{t_0}^{t_1} L(t, x(t), u(t)) dt + g(x(t_1)) \rightarrow \min, u \in Q, \quad (11)$$

при этом подразумевается, что  $x = W(u)$ , т.е.  $x, u$  связаны соотношениями (7). В условиях леммы 3 задача (7), (11) корректна для любого ограниченного множества  $Q$  класса  $Cv(U_p)$ .

Для нахождения её решения по методу условного градиента конструируется последовательность  $u_n \in Q_n$  ( $n = 0, 1, \dots$ ),  $Q_n$  – исчерпывающая множество  $Q$  последовательность множеств класса  $Cv(U_p)$ ). В качестве исходного приближения выбирается любой элемент  $u_0$  из  $Q_0$ . Если известны начальные приближения  $u_0, u_1, \dots, u_n$ , то для отыскания следующего приближения находится  $\mathbf{F}(u_n)$ , затем определяется  $v_n$  как решение вспомогательной экстремальной задачи

$$\langle v, J'(u_n) \rangle = \int_I (\mathbf{F}(u_n)(t), v(t)) dt \rightarrow \min, \quad v \in Q_{n+1},$$

далее составляется выпуклая комбинация  $u_{n\alpha} = (1 - \alpha)u_n + \alpha v_n$  и вычисляется  $\alpha_n$  из  $[0, 1]$ , минимизирующее функцию  $\alpha \rightarrow J(u_{n\alpha})$  на отрезке  $[0, 1]$ , и, наконец, полагается  $u_{n+1} = (1 - \alpha_n)u_n + \alpha_n v_n$ . Чаще всего ограничиваются постоянной последовательностью  $Q_n = Q$ . Если  $Q$  – достаточно простое множество в пространстве  $U_p$  (например, шар), то вспомогательные задачи легко решаются.

Приведённый выше анализ свойств функционала  $J$  показывает, что в достаточно общих относительно функций  $L, g$  предположениях к задаче (7), (11) применимы результаты предшествующих разделов. В частности, из теоремы 1 и леммы 2 следует компактность последовательности  $u_n$  в пространстве  $U_p$ , всякая предельная точка  $u$  этой последовательности является стационарным управлением, т.е. удовлетворяет линейаризованному принципу минимума

$$\int_{t_0}^{t_1} (\mathbf{F}(u)(t), v(t) - u(t)) dt \geq 0 \forall v \in Q.$$

Если стационарное управление единственно, то оно является решением задачи (7), (11). Требование единственности можно заменить предположением о выпуклости функционала  $J$ , в этой ситуации можно применить теоремы 3, 4.

В качестве частного случая рассмотрим вариационную задачу со свободным

правым концом

$$f(x) = \int_{t_0}^{t_1} L(t, x(t), x'(t)) dt + g(x(t_1)) \rightarrow \min, \quad x(t_0) = x_0. \quad (12).$$

Эта задача эквивалентна задаче (7), (11), если считать  $m = N$ ,  $A(t)$  – нулевая матрица,  $B(t) = \mathcal{E}_N$  – единичная матрица размеров  $N \times N$ . Соответствующая (12) задача оптимального управления может быть записана в следующей форме

$$J(u) = \int_{t_0}^{t_1} L(t, x(t), u(t)) + g(x(t_1)) dt \rightarrow \min, u \in U_p, \quad (13)$$

$$\frac{dx}{dt} = u, \quad x(t_0) = x_0. \quad (14)$$

Основное отличие (13), (14) от аналогичной ей задачи (11), (7) связано с тем, что в (13) множество  $Q$  – всё пространство  $U_p$ , а не его ограниченная часть, как в (11).

Справедливы следующие утверждения (доказать самостоятельно).

**Лемма 4.** *Интегральный оператор*

$$\mathfrak{A}(u) = \int_{t_0}^t u(s) ds$$

действует и непрерывен в пространстве  $U_p$ .

**Лемма 5.** *Пусть функция  $g$  ограничена снизу, а функция  $L(t, x, u)$  удовлетворяет предположениям леммы 3 и условию роста*

$$L(t, x, u) \geq \alpha |u|^p - \beta |x|^p - r,$$

где  $\alpha, \beta, r$  – положительные константы и

$$\alpha > \beta \|\mathfrak{A}\|_{U_p}^p.$$

Тогда для всех функций  $u$  класса  $U_p$  имеет место оценка

$$J(u) \geq \alpha_1 \|u\|_p^p - \beta_1, \quad (15)$$

в которой  $\alpha_1, \beta_1$  – положительные постоянные.

**Следствие.** *Пусть  $u_0 \in U_p$ . Тогда 1) множество  $\{u \in U_p, J(u) \leq J(u_0)\}$  ограничено в пространстве  $U_p$ , т.е. принадлежит некоторому шару  $Q$  пространства  $U_p$ ; 2) функция  $u_*$ , минимизирующая функционал  $J$  на шаре  $Q$ , реализует абсолютный минимум функционала  $J$  на пространстве  $U_p$ .*

В условиях леммы 5 функционал  $J$  принадлежит классу  $\Lambda_1(U_p)$  и является растущим на этом пространстве. Задача (13), (14) корректна и для её решения можно использовать метод условного градиента. В качестве  $Q$  можно взять шар  $\|u\| \leq R$  с достаточно большим радиусом  $R$ ; величину  $R$  можно найти, применяя оценку (15).

Укажем способ отыскания производной функционала  $J$  в точке  $u \in U_p$  по направлению  $v$ . Положим

$$a(t) = L_x(t, x(t), u(t)), \quad \mathcal{A}(t) = \int_t^{t_1} a(s) ds, \quad b(t) = L_u(t, x(t), u(t)),$$

где

$$x(t) = x_0 + \int_{t_0}^t u(s) ds.$$

Тогда

$$J'(u)v = \int_{t_0}^{t_1} (\mathcal{A}(t) + b(t) + \nabla g(x(t_1))v(t)) dt.$$

Данное равенство (в других терминах) уже встречалось нам при анализе задачи Больца. Если  $J'(u)v = 0$  для всех  $v$  из  $U_p$ , то

$$\mathcal{A}(t) + b(t) + \nabla g(x(t_1)) = 0.$$

Из этого равенства можно ещё раз вывести уравнение Эйлера и условие трансверсальности для решения исходной вариационной задачи (12).

**3. Метод Ритца.** В вариационном исчислении издавна используется метод Ритца. Кратко опишем его суть. Пусть требуется минимизировать действительный функционал  $J$  на некотором множестве  $\mathcal{U}$ . В методе Ритца задаче

$$J(u) \rightarrow \min, \quad u \in \mathcal{U} \tag{16}$$

сопоставляется последовательность аналогичных задач

$$J(u) \rightarrow \min, \quad u \in \mathcal{U}^n. \tag{16^n}$$

При этом предполагается, что задача (16<sup>n</sup>) проще исходной задачи (16) и её решение  $u^n$  не только существует, но и достаточно легко находится. Анализ  $u^n$  – приближения Ритца – оказывается полезным и для отыскания, и для характеристики решения  $\hat{u}$  исходной задачи (16). Такого рода подход фактически уже использовался нами. Например, при доказательстве принципа максимума Понтрягина исходная бесконечномерная задача редуцировалась с помощью игольчатых вариаций к некоторой конечномерной экстремальной задаче.

При определённых предположениях относительно функционала  $J$  и множеств  $\mathcal{U}, \mathcal{U}^n$  удаётся установить равенство

$$\lim_{n \rightarrow \infty} J(u^n) = \min_{u \in \mathcal{U}} J(u) = J(\hat{u}).$$

В этом случае говорят о сходимости метода Ритца по функционалу. Если же последовательность  $u^n$  сходится к  $\hat{u}$  в смысле какой-нибудь метрики  $\rho$ , то говорят о сходимости метода Ритца в смысле данной метрики. Вниманию читателя предлагается несколько задач, связанных с методом Ритца. Ниже  $U$  – непустое ограниченное замкнутое выпуклое подмножество рефлексивного пространства  $X$ ,  $f$  – функционал класса  $\Lambda_1(U)$ .

Задача 1. Последовательность  $U^n$  непустых замкнутых выпуклых подмножеств  $Q$  назовём исчерпывающей множество  $U$ , если

$$U^n \subset U^{n+1} \quad \text{и} \quad \overline{\bigcup_n U^n} = U.$$

Доказать, что

1° задачи (16) и  $(16)^n$  имеют решения  $\hat{u}$  и  $u^n$  соответственно;

2°  $J(u^{n+1}) \leq J(u^n)$  и  $J(u^n) \rightarrow J(\hat{u})$  при  $n \rightarrow \infty$ , т. е. метод Ритца сходится по функционалу;

3° если решение  $\hat{u}$  задачи (16) единственно, то  $u^n \rightarrow \hat{u}$  – сильная сходимость метода Ритца.

Задача 2. Для сепарабельного банахова пространства  $X$  существует последовательность  $X^n$  конечномерных подпространств  $X$ , обладающая свойствами

$$X^n \subset X^{n+1} \quad \text{и} \quad \overline{\bigcup_n X^n} = X. \quad (17)$$

Задача 3. Пусть  $U$  – замкнутое выпуклое подмножество пространства  $X$  с непустой внутренней точкой  $\overset{\circ}{U}$ , а последовательность конечномерных пространств  $X^n$  обладает свойствами (17). Доказать, что последовательность множеств

$$U^n := U \cap X^n$$

исчерпывает множество  $U$ .

Задача 4. Построить последовательность конечномерных пространств  $X^n$ , исчерпывающую пространство  $X = L^p(\Delta, \mathbb{R}^m)$  ( $1 < p < \infty$ ).

Задача 5. Реализуйте метод Ритца для задачи (7), (11).

Задача 6. Как избавиться от предположения  $\overset{\circ}{U} \neq \emptyset$ ?

Задача 7. Опишите несколько реализаций метода Ритца для простейшей вариационной задачи.



## Послесловие

Автор вынужден посетовать, что далеко не всё из задуманного удалось воплотить в представляемом суду читателя пособии. Ситуация вполне типичная: рассказано меньше, чем предполагалось, а объём оказался больше, чем хотелось.

Первоначально автор планировал посвятить отдельную главу постановке, а иногда и разбору решений ряда конкретных экстремальных задач. Назову несколько причин, в силу которых подобного раздела нет. Во-первых, по вариационному исчислению имеется большое число руководств, содержащих не только тексты задач, но и подробный их анализ с разнообразными комментариями (см., например, [5], [6], [11], [14], [28]). Вторая и не менее важная причина: включение подобного материала в пособие удвоило, а может быть, и утроило его размер. Это совершенно противоречит основной методической установке автора о разумной краткости учебных руководств. В-третьих, при всей важности и полезности задачников и решебников, в настоящее время более востребованы их электронные версии. Автора убедили в этом многочисленные беседы со студентами, мнение которых я всегда считал решающим. Мои ученики, бывшие и настоящие, в значительной мере – соавторы данного пособия. И если я обучаю и воспитываю студента около пяти лет, то студенты обучают и воспитывают меня более сорока лет.

Кто добился успехов в своей деятельности – я, студенты или те и другие, – пусть решит время!

## Список литературы

### I. Основные учебники

1. Алексеев В. М., Тихомиров В. М., Фомин С. В. Оптимальное управление. М.: Наука, 1979.
2. Болтянский В. Г. Математические методы оптимального управления. М.: Наука, 1969.
3. Колмогоров А. Н., Фомин С. В. Элементы теории функций и функционального анализа. М.: Наука, 1976.
4. Понтрягин Л. С., Болтянский В. Г., Гамкрелидзе Р. В., Мищенко Е. Ф. Математическая теория оптимальных процессов. М.: Наука, 1969.

### II. Дополнительная литература

5. Алексеев В. М., Галеев Э. М., Тихомиров В. М. Сборник задач по оптимизации. М.: Наука, 1984.
6. Андреева Е. А., Цирулёва В. М. Вариационное исчисление и методы оптимизации. Тверь: Твер. гос. ун-т, 2001.
7. Ахиезер Н. И. Лекции по вариационному исчислению. Харьков: Вища школа, 1981.
8. Буслаев В. С. Вариационное исчисление. Л.: Изд-во ЛГУ, 1980.
9. Гельфанд И. М., Фомин С. В. Вариационное исчисление. М.: Физматгиз, 1961.
10. Иоффе А. Д., Тихомиров В. М. Теория экстремальных задач. М.: Наука, 1974.
11. Краснов М. Л., Макаренко Г. И., Кисилёв А. И. Вариационное исчисление. М.: Наука, 1973.
12. Ли Э. Б., Маркус Л. Основы теории оптимального управления. М.: Наука, 1972.
13. Поляк Б. Т. Введение в оптимизацию. М.: Наука, 1983.
14. Пантелеев А. В., Лётова Т. А. Методы оптимизации в примерах и задачах. М.: Высшая школа, 2005.
15. Эльсгольц Л. Э. Дифференциальные уравнения и вариационное исчисление. М.: Физматгиз, 1965.

### Специальная литература

16. Бобылёв Н. А., Емельянов С. В., Коровин С. К. Геометрические методы в вариационных задачах. М.: Магистр, 1998.
17. Васильев Ф. П. Методы решения экстремальных задач. М.: Наука, 1981.
18. Васильев Ф. П. Численные методы решения экстремальных задач. М.: Наука, 1980.

19. Гирсанов И. В. Лекции по математической теории экстремальных задач. М.: Изд-во Моск. гос. ун-та, 1970.
20. Демьянов В. Ф., Рубинов А. М. Приближённые методы решения экстремальных задач. Л.: Изд-во Ленинградского гос. ун-та, 1968.
21. Лаврентьев М. А., Люстерник Л. А. Основы вариационного исчисления. М.: ОНТИ, 1935.
22. Ладыженская О. А., Уральцева Н. Н. Линейные и квазилинейные уравнения эллиптического типа. М.: Наука, 1973.
23. Ланцош К. Вариационные принципы механики. М.: Мир, 1965.
24. Милютин А. А. Принцип максимума в общей задаче оптимального управления. М.: Физматлит, 2001.
25. Мойсеев Н. Н., Иванюков Ю. П., Столярова Е. М. Методы оптимизации. М.: Наука, 1978.
26. Пшеничный Б. Н., Данилин Ю. М. Численные методы в экстремальных задачах. М.: Наука, 1975.
27. Скрышник И. В. Методы исследования нелинейных эллиптических граничных задач. М.: Наука, 1990.
28. Тихомиров В. М. Рассказы о максимумах и минимумах. М.: Наука, 1986.
29. Экланд И., Темам Р. Выпуклый анализ и вариационные проблемы. М.: Мир, 1979.
30. Янг Л. Лекции по вариационному исчислению и теории оптимального управления. М.: Мир, 1974.

Учебное издание

Климов Владимир Степанович

# Одномерные вариационные задачи

*Учебное пособие*

Редактор, корректор М. В. Никулина

Компьютерная верстка А. Ю. Ухалов

Подписано в печать 27.01.2011. Формат 60 x 84/16.

Бумага тип. Усл. печ. л. 16,27. Уч.-изд. л. 8,0.

Тираж 150 экз. Заказ .

Оригинал-макет подготовлен в редакционно-издательском отделе  
Ярославского государственного университета им. П. Г. Демидова  
150000 Ярославль, ул. Советская, 14.

Отпечатано на ризографе.

Ярославский государственный университет им. П. Г. Демидова  
150000 Ярославль, ул. Советская, 14.