Оглавление

Глава 1

Лекция 1.

"ВАРИАЦИОННОЕ ИСЧИСЛЕНИЕ И ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ"
(Осмоловский Н.П.)
Второй поток, 2001/2002 уч. год, осенний семестр.

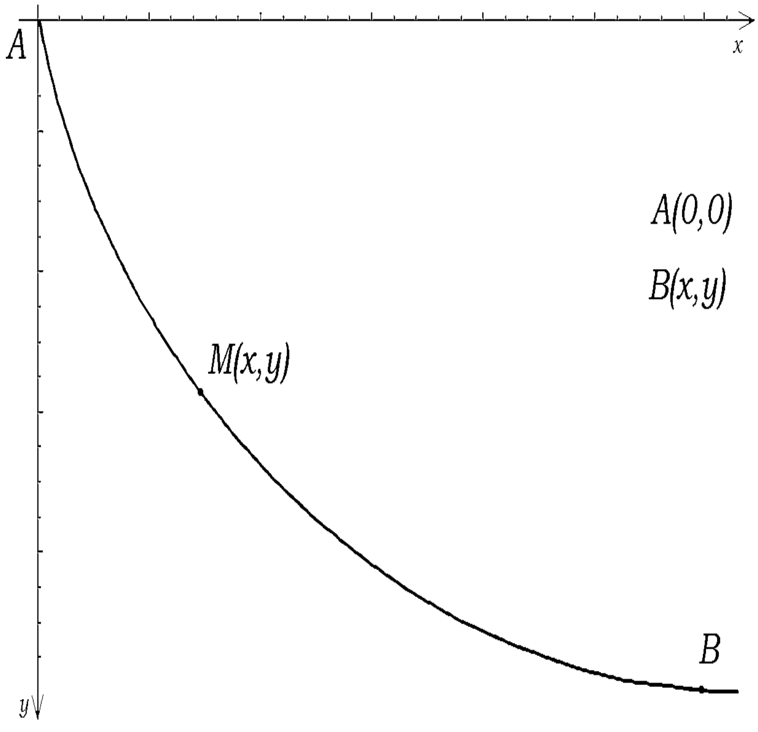
Лекция 1.

1.0.1 Задача о брахистохроне

(1696 г., И. Бернулли). Определить путь, спускаясь по которому под действием собственной тяжести, тело M, начав двигаться из точки A, дойдёт до точки B за кратчайшее время.

Решение задачи - кривая наискорейшего спуска, или *брахистохрона* (циклоида).

Формализация задачи :



В точке $M(x, y)$ по закону сохранения энергии имеем:

$$mgh + \frac{mv^2}{2} = 0$$

(так как в точке A и потенциальная, и кинетическая энергия равны нулю).

$$m = 1,\ h = -y \Rightarrow\ \frac{v^2}{2} = gy;\ v = \sqrt{2gy};$$

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

$$t = \int\limits_0^{x_1} \frac{\sqrt{1+y'^2}dx}{\sqrt{2gy}} \to \min \tag{a}$$

$$y(0) = 0, \quad y(x_1) = y_1. \tag{b}$$

Требуется найти функцию $y = y(x)$, $x \in [0, x_1]$ (x_1 - фиксирован), удовлетворяющую условиям (b) и доставляющую минимум интегралу (a).

Решение (циклоида) было дано самим И.Бернулли, а также Я.Бернулли, Лейбницем и Ньютоном.

1.0.2 Простейшая задача вариационного исчисления.

Какому классу задач принадлежит задача о брахистохроне? Опишем этот класс:

$$\min J(y(\cdot)) = \int\limits_{x_0}^{x_1} F(x, y(x), y'(x))\,dx, \tag{$*$}$$

$$y(x_0) = a, \qquad y(x_1) = b. \tag{$**$}$$

Отрезок $[x_0, x_1]$ фиксирован. Заданы также a, b и функция $F(x, y, z)$. Требуется найти функцию $y(x)$: $[x_0, x_1] \to \mathbb{R}^1$, удовлетворяющую *граничным условиям* ($**$) и доставляющую минимум интегральному функционалу ($*$). Задача ($*$) и ($**$) и есть простейшая задача вариационного исчисления.

В вариационном исчислении принято обозначать независимую переменную через x, а в оптимальном управлении - через t. Мы сразу примем обозначения оптимального управления и переформулируем задачу следующим образом:

$$\min J(x(\cdot)) = \int\limits_{t_0}^{t_1} F(t, x(t), u(t))dt, \tag{1.1}$$

$$\dot{x}(t) = u(t), \tag{1.2}$$

$$x(t_0) = a, \ \ x(t_1) = b, \tag{1.3}$$

$$(t, x(t), u(t)) \in Q. \tag{1.4}$$

Итак, вместо $y(x)$ теперь мы пишем $x(t) : [t_0, t_1] \to \mathbb{R}^1$. Переменную t принято трактовать как время.

Отрезок $[t_0, t_1]$ фиксирован, числа a и b заданы, $Q \in \mathbb{R}^3$ — открытое множество, служащее областью определения функции $F(t, x, u) : Q \to \mathbb{R}$, которая также известна. Пока нам достаточно считать, что *функция F непрерывна на Q вместе со своими производными F_x и F_u.*

Итак, мы имеем задачу на фиксированном отрезке $[t_0, t_1]$ с *закреплёнными концами* (согласно условиям (1.3)). Условия (1.3) называются *ограничениями* задачи. Условие (1.4) также является ограничением, но смысл у него несколько иной, чем у условий (1.3), поскольку множество Q открыто. Q — это, так сказать, «вселенная» данной задачи, где всё разыгрывается. Наконец, к условию (1.2) можно пока относиться как к обозначению для производной $\dot{x}$, но сразу отметим следующий факт: в вариационном исчислении было принято *варьировать* (изменять) функцию $x(t)$, поэтому мы и написали $J(x(\cdot))$. Задав $x(\cdot)$, мы вычисляем $\dot{x}(\cdot)$ и получаем соответствующее значение $J(x(\cdot))$. Но можно посмотреть иначе: задав $u(\cdot)$, мы можем получить $x(t) = a + \int\limits_{t_0}^{t_1} u(\tau)d\tau$ и вычислить соответствующее значение J. Таким образом, мы можем с таким же успехом рассматривать функционал J как функционал, зависящий от $u(\cdot) : J = J(u(\cdot))$, и варьировать не $x(\cdot)$, а $u(\cdot)$ при получении необходимых или достаточных условий минимума. Именно так предпочитают делать в оптимальном управлении, и называют функцию $u(\cdot)$ *управлением.* Оптимальное управление пошло дальше, рассматривая задачу в пространстве пар функций $w(\cdot) = (x(\cdot), u(\cdot))$, но об этом позднее. Сейчас лишь отметим, что правильный выбор переменных и пространства переменных, а также правильная «канонизация» задач обусловили определённый прогресс в оптимальном управлении по сравнению с классическим вариационным исчислением (который проявился даже в рамках вариационного исчисления, т.е на уровне задач, рассматриваемых последним). Мы будем считать себя свободными в выборе вариаций $x(\cdot)$ или $u(\cdot) = \dot{x}(\cdot)$, в зависимости от удобства.

Опишем теперь пространство функций, в котором рассматривается простейшая задача. В вариационном исчислении принято считать, по крайней мере на первом этапе, что $\dot{x}(\cdot) = u(\cdot)$ непрерывна, и, следовательно, $x(\cdot)$ - непрерывно дифференцируема. Таким образом,

$$x(\cdot) \in \mathcal{C}^1([t_0, t_1]; \mathbb{R}),$$

$$u(\cdot) \in \mathcal{C}([t_0, t_1]; \mathbb{R}).$$

Напомним, что

$$\|u(\cdot)\|_C = \max_{t \in [t_0, t_1]} |u(t)|$$

есть норма в пространстве $\mathcal{C}$ непрерывных функций, а

$$\|x(\cdot)\|_{C^1} = \max\{\|x(\cdot)\|_C, \|\dot{x}(\cdot)\|_C\}$$

есть норма в пространстве $\mathcal{C}^1$ непрерывно дифференцируемых функций. Пространства $\mathcal{C}$ и $\mathcal{C}^1$, снабжённые этими нормами, являются *банаховыми*, т.е полными нормированными пространствами. Следуя традиции, мы также будем поначалу рассматривать простейшую задачу в пространстве непрерывно дифференцируемых функций $x(\cdot)$. Затем мы расширим это пространство.

1.0.3 Слабый минимум.

Пусть $x^\circ(\cdot)$ — *допустимая траектория*, т.е траектория, удовлетворяющая ограничениям простейшей задачи и такая, что $x^\circ(\cdot) \in \mathcal{C}^1([t_0, t_1], \mathbb{R})$. Будем говорить, что $x^\circ(\cdot)$ доставляет *слабый минимум* в простейшей задаче, если существует $\varepsilon > 0$ такое, что для любой другой допустимой траектории $x(\cdot) \in \mathcal{C}^1([t_0, t_1], \mathbb{R})$, удовлетворяющей условиям :

$$|x(t) - x^\circ(t)| \leqslant \varepsilon; \qquad |\dot{x}(t) - \dot{x}^\circ(t)| \leqslant \varepsilon, \tag{1.5}$$

выполнено неравенство

$$J(x(\cdot)) \geqslant J(x^\circ(\cdot)).$$

Условия (1.5) равносильны условию

$$\|x(\cdot) - x^\circ(\cdot)\|_{\mathcal{C}^1} \leqslant \varepsilon.$$

Таким образом, слабый минимум есть локальный минимум в пространстве $\mathcal{C}^1([t_0, t_1]; \mathbb{R})$.

Часто для удобства траектории $x^\circ(\cdot)$, $x(\cdot)$ называют *точками пространства.* Далее нас будет интересовать необходимое условие слабого минимума в точке $x^\circ(\cdot)$.

Через $\delta x(\cdot)$ мы будем обозначать произвольную функцию в пространстве $\mathcal{C}^1([t_0, t_1]; \mathbb{R})$ и называть её *вариацией.* Соответственно будем полагать $\delta u(t) = \delta \dot{x}(t),\ t \in [t_0, t_1]$. Вариацию будем называть *допустимой* (в точке $x^\circ(\cdot)$, если точка $x^\circ(\cdot) + \delta x(\cdot)$ является допустимой. При получении условий слабого минимума мы будем использовать вариации $\delta x(\cdot)$ с малой нормой в пространстве $\mathcal{C}^1([t_0, t_1]; \mathbb{R}^n)$. Допустимая вариация должна удовлетворять условиям

$$\delta x(t_0) = \delta x(t_1) = 0. \tag{1.6}$$

Если норма $\|\delta x\|_{\mathcal{C}^1}$ мала, то условие $(t, x^\circ(t) + \delta x(t), u^\circ(t) + \delta u(t)) \in Q$ выполнено автоматически. Обозначим через $\mathcal{C}_0^1$ подпространство функций $\delta x(\cdot) \in \mathcal{C}^1([t_0, t_1]; \mathbb{R})$, удовлетворяющих условиям (6). Слабый минимум означает, что $J(x^\circ + \delta x) - J(x^\circ) \geqslant 0$ для всех достаточно малых по норме $\mathcal{C}^1$ вариаций $\delta x \in \mathcal{C}_0^1$.

1.0.4 Первая вариация функционала

Для краткости будем полагать $(x, u) = w$. Пусть $u^\circ = \dot{x}^\circ$, где x° — допустимая траектория (сейчас нам важно лишь, что $(t, x^\circ(t), u^\circ(t)) \in Q$). Пусть $\delta x(\cdot) \in \mathcal{C}^1([t_0, t_1], \mathbb{R})$ — малая вариация, положим $\delta u = \delta \dot{x}, \quad \delta w = (\delta x, \delta u)$. Согласно теореме о среднем существует функция $\theta(t),\ 0 \leqslant \theta(t) \leqslant 1$ такая, что

$$\delta F = F(t, w^\circ(t) + \delta w(t)) - F(t, w^\circ(t)) = \langle F'_w(t, w^\circ(t) + \theta(t)\delta w(t)), \delta w(t) \rangle$$

(предполагается, что при каждом t отрезок с концами $(t, w^\circ(t))$ и $(t, w^\circ(t) + \theta(t)\delta w(t))$ лежит в Q, что обеспечивается малостью вариации δx).

Следовательно,

$$J(x^\circ+\delta x)-J(x^\circ)=\int_{t_0}^{t_1}\langle F'_w(t,w^\circ(t)),\delta w(t)\rangle dt+r(\delta x), \tag{1.7}$$

где

$$r(\delta x)=\int_{t_0}^{t_1}\langle (F_w^\theta-F_w^\circ),\delta w\rangle\,dt.$$

Здесь $F_w^\theta=F'_w(t,w^\circ(t)+\theta(t)\delta w(t))$, $\quad F_w^0=F'_w(t,w^\circ(t))$.

Ясно, что $\|F_w^\theta-F_w^0\|_{\mathcal{C}}\to 0$ при $\|\delta w\|_{\mathcal{C}}\to 0$. Последнее имеет место, если $\delta u=\delta\dot x$ и $\|\delta x\|_{\mathcal{C}^1}\to 0$.

Поскольку $|r(\delta x)|\leqslant\|F_w^\theta-F_w^\circ\|_{\mathcal{C}}\|\delta w\|_{\mathcal{C}}(t_1-t_0)$ и $\|\delta w\|_{\mathcal{C}}\leqslant 2\|\delta x\|_{\mathcal{C}^1}$ то

$$|r(\delta x)|=o(\|\delta x\|_{\mathcal{C}^1}). \tag{1.8}$$

Условия (1.7) и (1.8) вместе означают, что выражение

$$\int_{t_0}^{t_1}\langle F'_w(t,w^\circ(t)),\delta w(t)\rangle dt=\int_{t_0}^{t_1}(F'_x(t,x^\circ(t),\dot x^\circ(t))\delta x(t)+F'_{\dot x}(t,x^\circ(t),\dot x^\circ(t))\delta\dot x(t))dt$$

есть производная Фреше функционала $J(x)$ в точке $x^\circ(\cdot)$ (в пространстве $\mathcal{C}^1$). Это выражение называют *первой вариацией* функционала $J(x)$ в точке $x^\circ(\cdot)$. Первую вариацию часто обозначают δJ. Мы будем обозначать её $J'(x^\circ;\delta x)$, а δJ сохраним для обозначения полного приращения функционала:

$$\delta J=J(x^\circ+\delta x)-J(x^\circ).$$

Теорема (необходимое условие слабого минимума). *Если $x^\circ(\cdot)$ — точка слабого минимума в простейшей задаче, то первая вариация функционала в этой точке равна нулю на любой вариации из $\mathcal{C}_0^1$.*

Доказательство. Пусть $\bar x(\cdot)\in\mathcal{C}_0^1$. Положим $\delta x=\varepsilon\bar x$. При всех достаточно малых ε имеем:

$$0\leqslant J(x^\circ+\varepsilon\bar x)-J(x^\circ)=\int_{t_0}^{t_1}\langle F_w^0,\varepsilon\bar w\rangle dt+o(\varepsilon),$$

где $\bar w=(\bar x,\bar u)$, $\bar u=\dot{\bar x}$ (см. (1.7) и (1.8)). Поскольку ε — произвольного знака, то отсюда получаем

$$\int_{t_0}^{t_1}\langle F_w^0,\bar w\rangle\,dt=0,$$

или

$$\int_{t_0}^{t_1}(F_x^0\bar x+F_{\dot x}^\circ\dot{\bar x})\,dt=0, \tag{1.9}$$

где $\bar x(\cdot)\in\mathcal{C}_0^1$ — произвольный элемент. $\square$

1.0.5 Уравнение Эйлера.

Итак, равенство нулю первой вариации на подпространстве $\mathcal{C}_0^1$ есть необходимое условие слабого минимума. Теперь нам предстоит его проанализировать и вывести из него знаменитое «уравнение Эйлера» — важнейшее необходимое условие в вариационном исчислении.

Сначала мы пойдём по более лёгкому, но не вполне «законному» пути, состоящем в интегрировании по частям члена

$$\int_{t_0}^{t_1}F_{\dot x}^\circ\dot{\bar x}\,dt=F_{\dot x}^\circ\bar x|_{t_0}^{t_1}-\int_{t_0}^{t_1}(\frac{d}{dt}F_{\dot x}^\circ)\bar x\,dt=-\int_{t_0}^{t_1}(\frac{d}{dt}F_{\dot x}^\circ)\bar x\,dt.$$

Второе равенство имеет место в силу условий $\bar x(t_0)=\bar x(t_1)=0$. Первое имеет место, если предположить, что $F_{\dot x}^\circ=F_u(t,x^\circ(t),\dot x^\circ(t))$ дифференцируема по t (оказывается, что это предположение предположением не является, а является частью необходимого условия, но для доказательства этого нужен другой путь, который мы проделаем чуть позже). Таким образом, первая вариация преобразовалась к виду

$$\int_{t_0}^{t_1} F_w^\circ \bar{w}\, dt = \int_{t_0}^{t_1} (F_x^\circ - \frac{d}{dt} F_{\dot{x}}^\circ)\bar{x}\, dt. \tag{1.10}$$

Следовательно,

$$\int_{t_0}^{t_1} (F_x^\circ - \frac{d}{dt} F_{\dot{x}}^\circ)\bar{x}\, dt = 0 \quad \forall \bar{x}(\cdot) \in \mathcal{C}_0^1. \tag{1.11}$$

Считаем, что $F_{\dot{x}}^\circ$ непрерывно дифференцируема по t и, следовательно функция $F_x^\circ - \frac{d}{dt} F_{\dot{x}}^\circ$ непрерывна по t. Тогда из (1.11) легко следует, что

$$F_x^\circ - \frac{d}{dt} F_{\dot{x}}^\circ = 0 \tag{1.12}$$

(сформулируйте и докажите самостоятельно лемму, которая здесь нужна; она называется леммой Лагранжа; получите с её помощью уравнение Эйлера). Уравнение (1.12) и есть уравнение Эйлера. Подробнее, оно имеет вид:

$$\frac{d}{dt} F_{\dot{x}}(t, x(t), \dot{x}(t)) = F_x(t, x(t), \dot{x}(t)). \tag{1.13}$$

(мы заменили $x^\circ(t)$ на $x(t)$). Это – уравнение второго порядка, для которого имеется ещё два граничных условия :

$$x(t_0) = a, \quad x(t_1) = b. \tag{1.14}$$

Итак, мы применили интегрирование по частям второго члена в выражении для первой вариации и преобразовали первую вариацию по Лагранжу к виду (1.10).

Теперь мы применим *другой путь*, интегрируя по частям первый член:

$$\int_{t_0}^{t_1} F_x^\circ \bar{x}\, dt = \int_{t_0}^{t_1} (\frac{d}{dt} \int_{t_0}^{t} F_x^\circ[\tau] d\tau)\bar{x}(t) dt = -\int_{t_0}^{t_1} (\int_{t_0}^{t} F_x^\circ[\tau] d\tau)\dot{\bar{x}}(t)\, dt$$

(равенство опять использует условия $\bar{x}(t_0) = \bar{x}(t_1) = 0$). Здесь $F_x^\circ[\tau] = F_x'(\tau, x^\circ(\tau), \dot{x}^\circ(\tau))$. Таким образом,

$$\int_{t_0}^{t_1} \langle F_w^\circ, \bar{w}\rangle dt = \int_{t_0}^{t_1} (-\int_{t_0}^{t} F_x^\circ[\tau] d\tau + F_{\dot{x}}^\circ[t])\bar{u}(t)\, dt, \text{ где } \bar{u} = \dot{\bar{x}}.$$

Это – преобразование первой вариации *по Дюбуа - Реймону.* Окончательно, мы получаем :

$$\int_{t_0}^{t_1} (-\int_{t_0}^{t} F_x^\circ d\tau + F_{\dot{x}}^\circ)\bar{u}(t)\, dt = 0 \quad \forall \bar{u}(t) \in \mathcal{C}_0, \tag{1.15}$$

где

$$\mathcal{C}_0 = \{\bar{u}(t) \in \mathcal{C}([t_0, t_1], \mathbb{R}) \mid \int_{t_0}^{t_1} \bar{u}(t)\, dt = 0\}$$

— подпространство в $\mathcal{C}([t_0, t_1], \mathbb{R})$.

Здесь условие $\int_{t_0}^{t_1} \bar{u}(t)\, dt = 0$ вытекает из условий $\dot{\bar{x}} = \bar{u}$, $\bar{x}(t_0) = \bar{x}(t_1) = 0$.

Для завершения вывода уравнения Эйлера нам понадобится следующая

Лемма (Дюбуа - Реймон). *Пусть $f(t)$ — непрерывная функция на $[t_0, t_1]$. Пусть для любой $\bar{u}(\cdot) \in \mathcal{C}[t_0, t_1]$ из условия $\int_{t_0}^{t_1} \bar{u}(t)\, dt = 0$ следует условие $\int_{t_0}^{t_1} f(t)\bar{u}(t)\, dt = 0$. Тогда $f(t) = const$.*

Абстрактным аналогом леммы Дюбуа-Реймона является следующая

Лемма (о зависимости двух линейных функционалов). *Пусть* X *— линейное пространство,* $l : X \to \mathbb{R}$ *и* $m : X \to \mathbb{R}$ *— два линейных функционала. Если* $m(x) = 0$ *для всех* $x \in X$ *таких, что* $l(x) = 0$, *то существует* $\lambda \in \mathbb{R}$ *такое, что* $m = \lambda l$.

Доказательство. Если $l = 0$, то и $m = 0$ и тогда λ — любое. Пусть $l \neq 0$ и пусть x_1 таков, что $l(x_1) = 1$. Для произвольного $x \in X$ положим $x_0 = x - l(x)x_1$. Тогда $l(x_0) = 0$ и, следовательно $m(x_0) = 0$, откуда $m(x) = l(x)m(x_1)$. Таким образом, $\lambda = m(x_1)$. □

Лемма Дюбуа-Реймона очевидным образом вытекает из этой леммы.

Глава 2

Лекция 2.

Завершим теперь вывод уравнения Эйлера. Вернемся к условию (15). В силу леммы Дюбуа-Реймона из этого условия вытекает, что

$$-\int_{t_0}^{t} F_x^0 d\tau + F_{\dot{x}}^0 = const. \tag{2.1}$$

Функция $F_x^0 = F_x(t, x^0(t), \dot{x}^0(t))$ непрерывна по t. Следовательно, функция $F_{\dot{x}}^0 = F_{\dot{x}}(t, x^0(t), \dot{x}^0(t))$ непрерывно дифференцируема по t (!). Заметим, что этот факт мы получили, анализируя равенство нулю первой вариации, т. е. как следствие из необходимого условия слабого минимума (хотя $\dot{x}^0$ всего лишь непрерывна, и, следовательно, $F_{\dot{x}}^0$ a priori также всего лишь непрерывна).

Дифференцируя условие (2.1) по t мы приходим к уравнению Эйлера. Тем самым вывод уравнения Эйлера завершен.

2.0.6 Пространства функций.

Мы провели все рассмотрения в пространстве непрерывно дифференцируемых функций. Однако имеются задачи, в которых решение реализуется на «ломаных экстремалях», т. е. на функциях с кусочно-непрерывной производной.

В вариационном исчислении рассматривают простейшую задачу в классе $K\mathcal{C}^1$ кусочно гладких функций (с кусочно непрерывной производной $u = \dot{x}$) и получают условия минимума, в том числе и уравнение Эйлера, в этом, более широком классе.

Однако можно пойти дальше, рассматривая задачу в классе абсолютно непрерывных функций $x(\cdot)$.

Напомним, что абсолютно непрерывные функции почти всюду (п.в.) имеют производную и характеризуются тем свойством, что они равны интегралу от своей производной:

$$x(t) = x(t_0) + \int_{t_0}^{t} \dot{x}(\tau) d\tau.$$

Их производная $u = \dot{x}$ является произвольной функцией из L_1, т. е. измеримой и суммируемой с первой степенью. Множество всех абсолютно непрерывных функций $x(t) : [t_0, t_1] \to \mathbb{R}$ обозначается $W_1^1[t_0, t_1]$ (первая производная суммируема с первой степенью).

Этот класс функций $x(t)$ принято рассматривать в оптимальном управлении, и в этом же классе оказывается удобно рассматривать задачи вариационного исчисления и получать соответствующие условия минимума. Напомним, что

$$\|x\|_{W_1^1} = |x(t_0)| + \|\dot{x}\|_{L_1}, \tag{2.2}$$

где $\|u\|_{L_1} = \int\limits_{t_0}^{t_1} |u(t)| dt$ есть норма в пространстве функций, суммируемых с первой степенью. Пространство W_1^1, снабженное нормой (2.2), является банаховым.

Все же нам придется несколько сузить класс абсолютно непрерывных функций. Дело в том, что по некоторым соображениям следует ограничить производную $u = \dot{x}$. Мы будем считать, что $u = \dot{x}$ является произвольной функцией из класса $L_\infty[t_0, t_1]$ ограниченных измеримых функций $u(t) : [t_0, t_1] \to \mathbb{R}$.

Норма в этом пространстве определяется как

$$\|u\|_{L_\infty} = \operatorname{vrai} \max_{t\in[t_0,t_1]} |u(t)|. \tag{2.3}$$

Напомним, что такое vrai max (существенный максимум, или ess sup) измеримой функции $\varphi(t) : [t_0, t_1] \to \mathbb{R}$. Он определяется так:

$$\operatorname{vrai} \max_{[t_0,t_1]} \varphi(t) = \inf_{\mathcal{M}} \sup_{t\in\mathcal{M}} \varphi(t),$$

где infimum берется по всем измеримым множествам $\mathcal{M}$ полной меры Лебега в $[t_0, t_1]$, т.е. таким, что $mes\, \mathcal{M} = t_1 - t_0$. Пространство L_∞ с нормой (2.3) является банаховым. Множество абсолютно непрерывных функций $x(t)$ с существенно ограниченной производной, т.е. удовлетворяющих условию $\dot{x}(\cdot) = u(\cdot) \in L_\infty$, обозначается $W^1_\infty[t_0, t_1]$. Это — пространство с нормой

$$\|x\|_{W^1_\infty} = |x(t_0)| + \|\dot{x}\|_{L_\infty}.$$

Пространство W^1_∞, снабженное этой нормой, также является банаховым. Всякая функция $x(t)$ из этого пространства удовлетворяет условию Липшица на $[t_0, t_1]$ и наоборот, всякая функция, удовлетворяющая условию Липшица, есть элемент пространства W^1_∞ (т. е. она абсолютно непрерывна и ее производная, определенная п.в., является ограниченной измеримой функцией). Таким образом, W^1_∞ есть класс всех липшицевых функций.

Рассматривая задачу для $x(\cdot) \in W^1_\infty$ следует определить слабый минимум как локальный минимум в пространстве W^1_∞. Уравнение Эйлера есть необходимое условие слабого минимума в этом классе. Доказательства практически не меняются (попробуйте их провести). При этом устанавливается, что функция

$$\psi(t) = F_{\dot{x}}(t, x^\circ(t), \dot{x}^\circ(t))$$

абсолютно непрерывна и, следовательно, липшицева (ибо $\dot{\psi}(t) = F_x(t, x^0(t), \dot{x}^0(t))$ принадлежит L_∞). Этот факт есть часть необходимого условия.

Замечание. *Для $x^0(\cdot) \in W^1_\infty$ условие $(t, x, \dot{x}) \in Q$ следует понимать так: существует компакт $\mathcal{C} \subset Q$ такой, что $(t, x^0(t), \dot{x}^0(t)) \in \mathcal{C}$ п.в. на $[t_0, t_1]$.*

В заключение отметим, что все рассмотрения можно было бы провести для вектор-функций $x(t) : [t_0, t_1] \to \mathbb{R}^n$, $x = (x_1, \dots, x_n)$, и мы бы снова пришли к уравнению Эйлера, вид которого не меняется (лишь лемму Дюбуа-Реймона и ее абстрактный аналог пришлось бы несколько обобщить). Позже мы получим все указанные модификации уравнения Эйлера в значительно более широком классе задач. Приведем список пространств, которые мы будем использовать:

$$\begin{array}{ll}
C(\Delta, \mathbb{R}^n) & C^1(\Delta, \mathbb{R}^n) \\
KC(\Delta, \mathbb{R}^n) & KC^1(\Delta, \mathbb{R}^n) \\
L_\infty(\Delta, \mathbb{R}^n) & W^1_\infty(\Delta, \mathbb{R}^n) \\
L_1(\Delta, \mathbb{R}^n) & W^1_1(\Delta, \mathbb{R}^n).
\end{array}$$

Оператор дифференцирования отображает каждое пространство справа на соответствующее пространство слева; сверху - вниз пространства расширяются.

2.0.7 Первые интегралы уравнения Эйлера.

Итак, уравнение Эйлера

$$\frac{d}{dt} F_{\dot{x}}(t, x, \dot{x}) = F_x(t, x, \dot{x})$$

представляет собой дифференциальное уравнение второго порядка. Укажем случаи, когда легко выписываются его первые интегралы.

1. F не зависит от $x \sim F_x = 0 \Rightarrow F_{\dot{x}}(t, \dot{x}) = const.$

2. F не зависит от $\dot{x} \sim F_{\dot{x}} = 0 \Rightarrow F_x(t, x) = 0.$

3. F не зависит от t (самый важный случай) — имеем интеграл энергии:

$$\dot{x}F_{\dot{x}}(x,\dot{x}) - F(x,\dot{x}) = const.$$

Действительно,

$$\frac{d}{dt}(\dot{x}F_{\dot{x}} - F) = \ddot{x}F_{\dot{x}} + \dot{x}\frac{d}{dt}F_{\dot{x}} - F_x\dot{x} - F_{\dot{x}}\ddot{x} = \dot{x}\left(\frac{d}{dt}F_{\dot{x}} - F_x\right) = 0$$

(при условии, что $x(\cdot) \in \mathcal{C}^2$).

2.0.8 Экстремали в задаче о брахистохроне.

Задача имеет вид

$$\int\limits_0^{x_1} \frac{\sqrt{1+y'^2}}{\sqrt{y}}dx \to \min,$$

$$y(0) = 0, \qquad y(x_1) = y_1; \qquad x_1, y_1 \text{ - фиксированы.}$$

Предположения простейшей задачи не выполнены поскольку подинтегральная функция при $y = 0$ имеет особенность.

Положим $Q = \{(x,y,z) \mid y > 0\}$ и изменим начальное условие $y(0) = 0$ на условие $y(0) = \varepsilon$, где $\varepsilon > 0$ мало. Найдем экстремали в этой задаче. Здесь

$$F = \frac{\sqrt{1+y'^2}}{\sqrt{y}}$$

не зависит от x. Поэтому уравнение Эйлера имеет первый интеграл

$$y'F_{y'} - F = const.$$

Поскольку

$$F_{y'} = \frac{y'}{\sqrt{1+y'^2}} \cdot \frac{1}{\sqrt{y}},$$

то

$$y'F_{y'} - F = \frac{y'^2}{\sqrt{1+y'^2}\sqrt{y}} - \frac{\sqrt{1+y'^2}}{\sqrt{y}} = -\frac{1}{\sqrt{1+y'^2}\sqrt{y}}.$$

Следовательно,

$$\sqrt{y}\sqrt{1+y'^2} = c;$$

$$y(1+y'^2) = c^2;$$

$$y'^2 = \frac{c^2-y}{y};$$

$$\frac{dy}{dx} = \pm\frac{\sqrt{c^2-y}}{\sqrt{y}}; \qquad dx = \pm\frac{\sqrt{y}dy}{\sqrt{c^2-y}};$$

$$x = \pm\int \frac{\sqrt{y}dy}{\sqrt{c^2-y}}.$$

Замена: $y = c^2\sin^2 t \quad (y > 0)$

$$dy = 2c^2 \sin t \cos t dt,$$

$$x = \pm\int \frac{c\sin t 2c^2 \sin t \cos t dt}{c\cos t};$$

$$x = \pm 2c^2 \int \sin^2 t dt.$$

Выберем знак $+$ (что соответствует растущей функции). Поскольку

$$\int \sin^2 t dt = \frac{1}{2}\int (1-\cos 2t)dt = \frac{1}{2}(t - \frac{1}{2}\sin 2t) + c_1 = \frac{1}{4}(2t - \sin 2t) + c_1,$$

то

$$\begin{cases} x &= \frac{c^2}{2}(2t - \sin 2t) + c_1 \\ y &= \frac{c^2}{2}(1 - \cos 2t). \end{cases}$$

Положим $2t = \tau$. Тогда

$$\begin{cases} x - c_1 &= \frac{c^2}{2}(\tau - \sin \tau) \\ \\ y &= \frac{c^2}{2}(1 - \cos \tau). \end{cases}$$

Итак, экстремаль есть циклоида.

2.0.9 Сильный минимум. Условие Вейерштрасса.

Мы продолжаем рассматривать простейшую задачу вариационного исчисления (в.и.):

$$\int\limits_{t_0}^{t_1} F(t,x,u)dt \to \min,$$

$$\dot{x} = u,$$

$$x(t_0) = a, \quad x(t_1) = b,$$

$$(t,x,u) \in Q.$$

Предполагается, что F, F_x и F_u непрерывны на открытом множестве Q. Будем говорить, что $x^0(\cdot)$ доставляет *сильный минимум*, если существует $\varepsilon > 0$ такое, что для любой допустимой траектории $x(\cdot)$, удовлетворяющей условию

$$|x(t) - x^0(t)| \leqslant \varepsilon \ \forall t, \tag{2.4}$$

выполнено неравенство:

$$J(x) \geqslant J(x^0).$$

Условие (2.4) равносильно условию

$$\|x(\cdot) - x^0(\cdot)\|_C \leqslant \varepsilon.$$

Поэтому сильный минимум есть локальный минимум в смысле нормы пространства $\mathcal{C}$ (в то время как слабый есть минимум в смысле нормы пространства $\mathcal{C}^1$, если задача рассматривается в $\mathcal{C}^1$, или в смысле нормы W^1_∞, если задача рассматривается в W^1_∞).

Мы не сказали, в каком пространстве мы рассматриваем задачу, определяя понятие сильного минимума. Можно считать, что задача рассматривается в пространстве $x \in W^1_\infty$ (т. е. x — липшицева). В вариационном исчислении принято считать, что $x \in KC^1$, т. е. x — кусочно-гладкая функция.

Нас будет интересовать сейчас получение необходимого условия для сильного минимума, т. е. получение условия Вейерштрасса. Для простоты доказательств мы будем считать, что функция $x^0(t)$, исследуемая на минимум, принадлежит классу KC^1, т. е. ее первая производная $u^0(t) = \dot{x}^0(t)$ кусочно непрерывна.

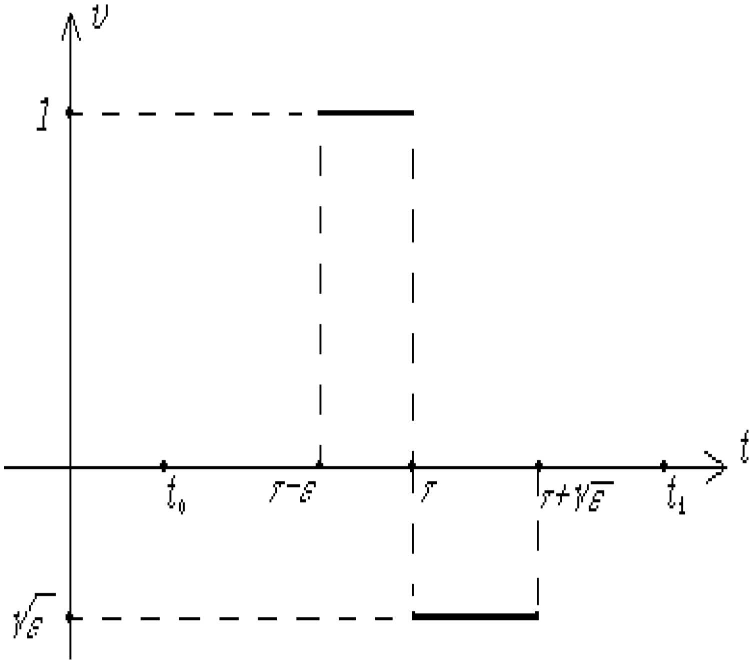
Итак, пусть $x^0 \in KC^1$ — допустимая траектория, доставляющая сильный минимум в KC^1. Пусть $\tau \in (t_0, t_1)$ - точка непрерывности функции $u^0(t)$. В окрестности точки τ мы определим вариацию Вейерштрасса, с помощью которой условие сильного минимума будет получено. Поскольку $x^0(\cdot)$ - допустима по ограничениям, то

$$(\tau, x^0(\tau), u^0(\tau)) \in Q.$$

Пусть Δu — число такое, что

$$(\tau, x^0(\tau), u^0(\tau) + \Delta u) \in Q$$

(если x, а значит и u — вектор, то Δu — также вектор). Рассмотрим функцию, график которой изображен на картинке:



Итак,

$$v = v(\tau, \varepsilon, t) = \begin{cases} 1 \text{ при } t \in (\tau - \varepsilon, \tau) \\ -\sqrt{\varepsilon} \text{ при } t \in (\tau, \tau + \sqrt{\varepsilon}) \\ 0 \text{ на дополнении.} \end{cases}$$

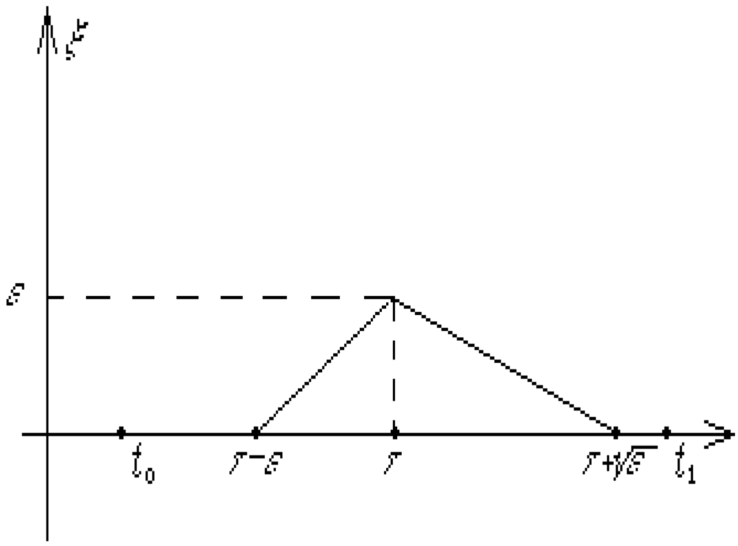
Слева от τ мы имеем «столбик» или «иголку», справа - малую (если ε - мало) «компенсирующую» вариацию, подобранную так, что

$$\int_{t_0}^{t_1} v(\tau, \varepsilon, t) dt = 0.$$

Отметим также, что v сосредоточена на $(\tau - \varepsilon, \tau + \sqrt{\varepsilon}) \overset{def}{=} \mathcal{M}_\varepsilon$ (т.е. отлична от нуля лишь на $\mathcal{M}_\varepsilon$). Рассмотрим также функцию

$$\xi(\tau, \varepsilon, t) = \int_{t_0}^{t} v(\tau, \varepsilon, \tilde{t}) d\tilde{t},$$

график которой имеет вид



Отметим, что $\|\xi\|_C \leqslant \varepsilon$ и ξ также сосредоточена на $(\tau-\varepsilon, \tau+\sqrt{\varepsilon}) = \mathcal{M}_\varepsilon$. Положим

$$\delta u_\varepsilon(t) = v(\tau, \varepsilon, t)\Delta u,$$

$$\delta x_\varepsilon(t) = \xi(\tau, \varepsilon, t)\Delta u.$$

Тогда

$$\delta \dot{x}_\varepsilon = \delta u_\varepsilon,$$

$$\delta x_\varepsilon(t_0) = \delta x_\varepsilon(t_1) = 0.$$

Ясно, что при всех достаточно малых $\varepsilon > 0$ траектория $x^0 + \delta x_\varepsilon$ является допустимой (поскольку Q открыто). Поскольку

$$\|\delta x_\varepsilon\|_c \leqslant \varepsilon |\Delta u|,$$

и в точке x^0 имеет место сильный минимум, то

$$J(x^0 + \delta x_\varepsilon) - J(x^0) \geqslant 0$$

для всех достаточно малых $\varepsilon > 0$. Следовательно,

$$\varliminf_{\varepsilon \to +0} \frac{J(x^0 + \delta x_\varepsilon) - J(x^0)}{\varepsilon} \geqslant 0. \tag{2.5}$$

Мы покажем, что нижний предел реализуется *как предел* и равен величине

$$\Delta F - F_u[\tau]\Delta u, \tag{2.6}$$

где

$$\Delta F = F(\tau, x^0(\tau), \dot{x}^0(\tau) + \Delta u) - F(\tau, x^0(\tau), \dot{x}^0),$$
$$F_u[\tau] = F_u(\tau, x^0(\tau), u^0(\tau)).$$

Положим

$$\mathcal{E}(t, x, u, \tilde{u}) = F(t, x, \tilde{u}) - F(t, x, u) - F_u(t, x, u)(\tilde{u} - u).$$

Эту функцию называют *функцией Вейерштрасса*. Ясно, что

$$\Delta F - F_u[\tau]\Delta u = \mathcal{E}(t, x^0(t), \dot{x}^0(t), \dot{x}^0(t) + \Delta u).$$

Таким образом, будет показано, что

$$\mathcal{E}(t, x^0(t), \dot{x}^0(t), \dot{x}^0(t) + \Delta u) \geqslant 0.$$

Это и есть *условие Вейерштрасса* в точке τ непрерывности функции $u^0(\cdot)$. Далее полагаем для краткости $\delta x_\varepsilon = \delta x, \quad \delta u_\varepsilon = \delta u$. Положим также

$$\delta F = F(t, w^0 + \delta w) - F(t, w^0),$$

где $\delta w = (\delta x, \delta u), \quad w^0 = (x^0, u^0)$.

Поскольку

$$\delta w \mathcal{X}_{\mathcal{M}_\varepsilon} = \delta w,$$

где $\mathcal{X}_{\mathcal{M}_\varepsilon}$ — характеристическая функция множества $\mathcal{M}_\varepsilon$, то

$$J(x^0 + \delta x) - J(x^0) = \int\limits_{t_0}^{t_1} \delta F dt = \int\limits_{\mathcal{M}_\varepsilon} \delta F\, dt.$$

Далее

$$\delta F = F(t, x^0 + \delta x, u^0 + \delta u) - F(t, x^0, u^0)$$

можно представить в виде

$$\delta F = \bar{\delta}_x F + \delta_u F,$$

где

$$\bar{\delta}_x F = F(t, x^\circ + \delta x, u^\circ + \delta u) - F(t, x^\circ, u^\circ + \delta u)$$

—приращение по x в «сдвинутой» точке $(t, x^0, u^0 + \delta u)$, а

$$\delta_u F = F(t, x^0, u^0 + \delta u) - F(t, x^0, u^0)$$

— приращение по u в точке (t, x^0, u^0). Поскольку функция F непрерывно дифференцируема по x, то она удовлетворяет условию Липшица по x на каждом ограниченном множестве области определения. Следовательно, существует константа $L > 0$ такая, что

$$\|\bar{\delta}_x F\|_{L_\infty} \leqslant L\|\delta x\|_C = L|\Delta u|\varepsilon,$$

откуда

$$|\int\limits_{\mathcal{M}_\varepsilon} \bar{\delta}_x F dt| \leqslant \|\bar{\delta}_x F\|_{L_\infty} mes\mathcal{M}_\varepsilon \leqslant L|\Delta u|\varepsilon\; mes\mathcal{M}_\varepsilon = o(\varepsilon).$$

Следовательно,

$$J(x^0 + \delta x) - J(x^0) = \int\limits_{\mathcal{M}_\varepsilon} \bar{\delta}_x F dt + \int\limits_{\mathcal{M}_\varepsilon} \delta_u F dt = \int\limits_{\mathcal{M}_\varepsilon} \delta_u F dt + o(\varepsilon). \tag{2.7}$$

Продолжение см. в лекции 3.

Глава 3

Лекция 3.

Продолжим *вывод условия Вейерштрасса.* Имеем

$$\int\limits_{\mathcal{M}_\varepsilon} \delta_u F\,dt = \int\limits_{\mathcal{M}_{\varepsilon-}} \delta_u F\,dt + \int\limits_{\mathcal{M}_{\varepsilon+}} \delta_u F\,dt, \tag{3.1}$$

где

$$\mathcal{M}_{\varepsilon-} = (\tau-\varepsilon,\tau), \qquad \mathcal{M}_{\varepsilon+} = (\tau,\tau+\sqrt{\varepsilon}).$$

Рассмотрим каждое слагаемое в отдельности. Имеем:

$$\int\limits_{\mathcal{M}_{\varepsilon-}} \delta_u F\,dt = \int\limits_{\mathcal{M}_{\varepsilon-}} \Delta F\,dt + \int\limits_{\mathcal{M}_{\varepsilon-}} (\delta_u F - \Delta F)\,dt = \Delta F\cdot\varepsilon + o(\varepsilon), \tag{3.2}$$

поскольку mes $\mathcal{M}_{\varepsilon-} = \varepsilon$, а $\delta_u F - \Delta F$ можно переписать как

$$F(t,x^0(t),u^0(t)+\Delta u) - F(\tau,x^0(\tau),u^0(\tau)+\Delta u) - (F(t,x^0(t),u^0(t)) - F(\tau,x^0(\tau),u^0(\tau))),$$

откуда ясно, что

$$\sup_{\mathcal{M}_{\varepsilon-}} |\delta_u F - \Delta F| \to 0 \text{ при } \varepsilon\to 0,$$

поскольку x° и u° непрерывны в точке τ.

Далее, покажем, что

$$\int\limits_{\mathcal{M}_{\varepsilon+}} \delta_u F\,dt = \int\limits_{\mathcal{M}_{\varepsilon+}} F_u^\circ \delta u\,dt + o(\varepsilon), \tag{3.3}$$

где $F_u^\circ = F_u(t,w^\circ(t))$.

Действительно, по теореме о среднем на $\mathcal{M}_{\varepsilon+}$ имеем:

$$\delta_u F = F_u^\theta \delta u,$$

где

$$F_u^\theta = F_u(t,x^\circ,u^\circ+\theta\delta u), \qquad \theta=\theta(t), \qquad 0\le\theta(t)\leqslant 1$$

и, следовательно, на $\mathcal{M}_{\varepsilon+}$ имеем:

$$\delta_u F = F_u^\circ\delta u + (F_u^\theta - F_u^\theta)\delta u.$$

Поскольку на $\mathcal{M}_{\varepsilon+}$ выполнено условие

$$|\delta u| = \sqrt{\varepsilon}|\Delta u| \Rightarrow \sup_{\mathcal{M}_{\varepsilon+}} |F_u^\theta - F_u^\circ| \to 0 (\varepsilon\to 0)$$

$$\text{mes}\mathcal{M}_{\varepsilon+} = \sqrt{\varepsilon},$$

то

$$|\int\limits_{\mathcal{M}_{\varepsilon+}} (F_u^\theta - F_u^\circ)\,dt| \leqslant \sup_{\mathcal{M}_{\varepsilon+}} |F_u^\theta - F_u^\circ|(\sqrt{\varepsilon}|\Delta u|)\sqrt{\varepsilon} = o(\varepsilon).$$

Следовательно, имеет место оценка (3.3). Наконец,

$$\int\limits_{\mathcal{M}_{\varepsilon+}} F_u^\circ \delta u\,dt = \int\limits_{\mathcal{M}_{\varepsilon+}} F_u[\tau]\delta u\,dt + \int\limits_{\mathcal{M}_{\varepsilon+}} (F_u^\circ - F_u[\tau])\delta u\,dt = -F_u[\tau]\Delta u\sqrt{\varepsilon}\sqrt{\varepsilon} + o(\varepsilon). \tag{3.4}$$

Действительно,

$$\sup_{\mathcal{M}_{\varepsilon+}} |F_u^\circ - F_u[\tau]| \to 0 \text{ при } \varepsilon \to 0,$$

т.к. $w^\circ(t)$ непрерывна в точке τ, следовательно,

$$|\int\limits_{\mathcal{M}_{\varepsilon+}} (F_u^\circ - F_u[\tau])\delta u\,dt| \leqslant \mathcal{M}_{\varepsilon+}|F_u^\circ - F_u[\tau]|(|\Delta u|\sqrt{\varepsilon})\sqrt{\varepsilon} = o(\varepsilon).$$

Из формул (2.7) - (3.4) получаем:

$$J(x^\circ + \delta x) - J(x_0) = \Delta F \cdot \varepsilon - F_u[\tau]\Delta u \cdot \varepsilon + o(\varepsilon).$$

Остается разделить на ε и перейти к пределу. Тогда получаем

$$\Delta F - F_u[\tau]\Delta u \geqslant 0. \tag{3.5}$$

Пусть теперь $t_* \in (t_0, t_1)$ - точка разрыва функции $u^\circ(t)$. Тогда существуют и конечны $u^\circ(t_* - 0)$ и $u^\circ(t_* + 0)$. Переходя в условии (3.5) к пределу при $\tau \to t_* - 0$, получаем

$$F(t_*, x^\circ(t_*), u^\circ(t_* - 0) + \Delta u) - F(t_*, x^\circ(t_*), u^\circ(t_* - 0)) - F_u[\tau]\Delta u \geq 0,$$

где

$$F_u[\tau] = F_u(t_*, x^\circ(t_*), \quad u^\circ(t_* - 0)) = F_u(t, x^\circ(t_*), \quad u^\circ(t_* + 0))$$

(поскольку $\psi(t) = F_u(t, w^\circ(t))$ не рвется!).

Аналогичное условие имеет место справа от t_*:

$$F(t_*, x^\circ(t_*), u^\circ(t_* + 0) + \Delta u) - F(t_*, x^\circ(t_*), u^\circ(t_* + 0)) - F_u[\tau]\Delta u \geq 0.$$

Итак, доказана

Теорема. . *Пусть $x^\circ(\cdot)$ доставляет сильный минимум в простейшей задаче в классе KC^1. Тогда для любой точки t непрерывности функции $\dot{x}^\circ(t)$ имеет место условие Вейерштрасса:*

$$\mathcal{E}(t, x^\circ(t), \dot{x}^\circ(t), \dot{x}^\circ(t) + \Delta u) \geqslant 0$$

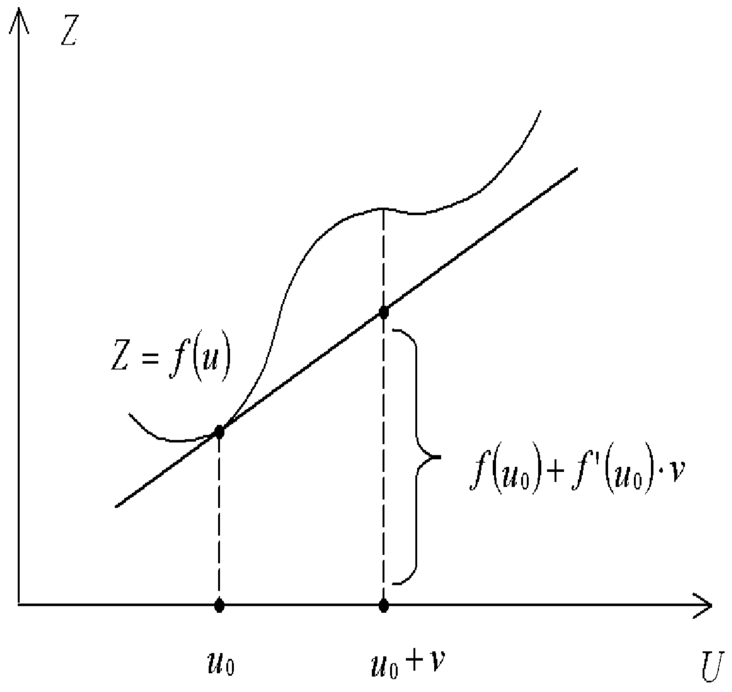
для всех Δu таких, что $(t, x^\circ(t), \dot{x}^\circ(t) + \Delta u) \in Q$, где

$$\mathcal{E}(t, x, u, v) = F(t, x, v) - F(t, x, u) - F_u'(t, x, u)(v - u).$$

Отметим *геометрический смысл* условия Вейерштрасса. Рассмотрим функцию, зависящую только от u:

$$z = f(u)$$

Неравенство $f(u_0 + v) - f(u_0) - f'(u_0)v \geqslant 0$ (фигурирующее в условии Вейерштрасса) означает, что график функции $z = f(u)$ лежит над касательной к графику в точке u_0



(точнее, не ниже касательной). Если функция выпукла, то это условие автоматически выполнено. Поэтому условие Вейерштрасса всегда выполнено, если *интегрант* $F(t,x,u)$ является выпуклой функцией по u. Как уже отмечалось, траекторию $x(t)$, удовлетворяющую уравнению Эйлера

$$\frac{d}{dt}F_{\dot{x}}(t,x,\dot{x}) = F_x(t,x,\dot{x}),$$

называют *экстремалью.*

Однако весьма часто понятие экстремали включает требование, чтобы для $x(t)$ было также выполнено условие Вейерштрасса. Чтобы различать эти два понятия, мы договоримся: траекторию, удовлетворяющую только уравнению Эйлера, называть *слабой* экстремалью, а траекторию, удовлетворяющую уравнению Эйлера и условию Вейерштрасса, называть *сильной* экстремалью. В оптимальном управлении (где условие Вейерштрасса имеет своим аналогом принцип максимума Понтрягина) последнюю называют еще понтрягинской экстремалью (или просто экстремалью).

3.0.10 Канонические переменные. Принцип максимума. Функция Понтрягина.

Итак, «слабая» экстремаль в простейшей задаче определяемая уравнением Эйлера:

$$\frac{d}{dt}F_{\dot{x}}(t,x,\dot{x}) = F_x(t,x,\dot{x}). \tag{3.1}$$

Положим

$$\dot{x} = u, \qquad \psi = F_u(t,x,u). \tag{3.2}$$

Тогда уравнение Эйлера оказывается эквивалентно системе

$$\begin{cases} \dot{x} &= u \\ \dot{\psi} &= F_x(t,x,u) \\ \psi &= F_u(t,x,u) \end{cases} \tag{3.3}$$

Положим

$$H = \psi u - F(t,x,u).$$

Таким образом, $H = H(t,x,u,\psi)$. Тогда система (3.3) может быть переписана в виде

$$\begin{cases} \dot{x} = H_\psi(t,x,u,\psi) \\ -\dot{\psi} = H_x(t,x,u,\psi) \\ H_u(t,x,u,\psi) = 0 \end{cases} \tag{3.4}$$

Переменная ψ называется *двойственной* или *сопряженной*, а переменная u — *управлением.* (Обозначение ψ введено Л.С.Понтрягиным и его сотрудниками для задач оптимального управления. Другое, часто встречающееся обозначение для сопряженной переменной - это p, - связано с тем, что в задачах механики сопряженная переменная есть импульс системы).

Предположим, что выполнены условия (например, $F_{uu} \neq 0$ на Q , или, более того, $F_{uu} > 0$ на Q; последнее гарантирует строгую выпуклость F по переменной u), позволяющие использовать условие

$$H_u(t, x, u, \psi) = 0,$$

или, что тот же самое, условие

$$\psi = F_u(t, x, u),$$

для того, чтобы выразить управление u как функцию от t, x, ψ:

$$u = U(t, x, \psi). \tag{3.5}$$

Пусть

$$H_u(t, x, U(t, x, \psi), \psi) \equiv 0 \ \forall t, x, \psi.$$

Рассмотрим функцию

$$\mathcal{H}(t, x, \psi) = H(t, x, U(t, x, \psi), \psi). \tag{3.6}$$

Имеем:

$$\frac{\partial \mathcal{H}}{\partial \psi} = \left(\frac{\partial H}{\partial u} \frac{\partial U}{\partial \psi} + \frac{\partial H}{\partial \psi} \right) |_{u=U(t,x,\psi)}$$

Но $U(t, x, \psi)$ найдена из условия $H_u = 0$. Следовательно,

$$\frac{\partial \mathcal{H}}{\partial \psi} = \frac{\partial H}{\partial \psi}$$

Аналогично,

$$\frac{\partial \mathcal{H}}{\partial x} = \frac{\partial H}{\partial x}.$$

Итак, мы получаем: если $x(\cdot)$ — экстремаль, а $\psi(\cdot)$ — соответствующая двойственная переменная ($\psi(t) = F_u(t, x(t), \dot{x}(t))$), то пара функций $x(t), \psi(t)$ удовлетворяет системе уравнений

$$\begin{cases} \dot{x} = \mathcal{H}_\psi(t, x, \psi) \\ -\dot{\psi} = \mathcal{H}_x(t, x, \psi) \end{cases} \tag{3.7}$$

Системы такого вида называют *гамильтоновыми*, а переменные x, ψ — *каноническими.*

Отметим, что система (3.4) гамильтоновой не является, поскольку содержит параметр u, связанный с остальными переменными через последнее соотношение $H_u = 0$. Тем не менее, основной для нас все-таки будет система (3.4), поскольку переход к гамильтоновой системе связан с дополнительными требованиями (которыми мы не хотим себя заранее обременять).

Указанный прием перехода к гамильтоновой системе и к каноническим переменным на уровне слабых экстремалей (определяемых уравнением Эйлера) имеет прямое отношение к так называемому классическому *преобразованию Лежандра.* Пусть $f = f(u)$ — гладкая функция ($u \in \mathbb{R}^n$). Ее классическим преобразованием Лежандра называется новая функция $h = h(\psi)$, определяемая равенством $h = \psi u - f(u)$, где $u = U(\psi)$ найдено из условия $\psi = f'(u)$. При этом предполагается, что условие $\psi = f'(u)$ позволяет однозначно найти u как функцию от $\psi : u = U(\psi)$ (например $f''(u) > 0 \ \forall u$). Ясно, что переход от $F(t, x, u)$ к $\mathcal{H}(t, x, \psi)$ есть преобразование Лежандра (классическое) функции F по аргументу u (при фиксированных t, x, рассматриваемых в преобразовании Лежандра в качестве параметров). Подробнее на этом не останавливаемся (см. АТФ, стр. 226).

Обратимся теперь к условию Вейерштрасса. Функция Понтрягина H позволяет сформулировать условие Вейерштрасса в следующем эквивалентном виде (проверьте эквивалентность): для любой точки t непрерывности «управления» $u^\circ(t)$ имеет место неравенство

$$H(t, x^\circ(t), u, \psi(t)) \leqslant H(t, x^\circ(t), u^\circ(t), \psi(t)) \tag{3.8}$$

для всех значений u таких, что $(t, x^\circ(t), u) \in Q$, где

$$\begin{aligned} &\psi(t) = F_u(t, x^\circ(t), u^\circ(t)) \\ &u^\circ(t) = \dot{x}^\circ(t). \end{aligned} \tag{3.9}$$

Условие (3.8) носит название *принципа максимума*. Оно утверждает, что максимум функции Понтрягина H по переменной u достигается (при каждом t) на «оптимальном управлении» $u^{\circ}(t)$. Это условие также может быть использовано для перехода к гамильтоновой системе. А именно, полагают

$$\mathcal{H}(t, x, \psi) = \sup_{u} H(t, x, u, \psi), \tag{3.10}$$

где supremum берется по u таким, что $(t, x, u) \in Q$. Эту функцию называют *гамильтонианом*. (Мы используем для нее то же обозначение, что и ранее использовали для функции (3.6). Докажите, что если $F_{uu} > 0$ всюду, то обе функции совпадают; воспользуйтесь при этом тем обстоятельством, что для выпуклой функции условие стационарности не только необходимо, но и достаточно для минимума. В общем случае функции (3.6) и (3.10) могут не совпадать). При определенных условиях доказывается, что сильные экстремали (т.е. удовлетворяющие уравнению Эйлера и условию Вейерштрасса) могут быть найдены как решения гамильтоновой системы с гамильтонианом (3.10). Переход к этому гамильтониану связан с преобразованием, известном в выпуклом анализе под названием преобразования Юнга - Фенхеля. Дадим определение.

Пусть $f = f(u)$ — некоторая функция. Ее преобразованием Юнга-Фенхеля называется новая функция $h = h(\psi)$, определенная равенством:

$$h(\psi) = \sup_{u} \{\psi u - f(u)\}$$

Полагают $h(\psi) = f^{*}(\psi)$ и называют f^{*} функцией, сопряженной к f. Сопряженная функция является выпуклой. Ясно, что гамильтониан $\mathcal{H}$, определенный равенством (3.10), есть преобразование Юнга - Фенхеля по переменной u (при фиксированных t и x) от функции П онтрягина H. Подробнее об этом преобразовании см. АТФ, с.224.

В дальнейшем мы получим принцип максимума для класса задач, охватывающего не только все вариационное исчисление, но и многие задачи оптимального управления.

Глава 4

Лекция 4.

4.1 Тема: Аппарат теории экстремума.

4.1.1 Теорема Банаха об обратном операторе, лемма о правом обратном, теорема Банаха–Шаудера об открытом отображении.

Пусть X, Y — банаховы пространства, $A : X \to Y$ — линейный непрерывный оператор. Непрерывность $A \sim$ ограниченность.

Теорема Банаха (об обратном операторе). *Пусть* $A : X \xrightarrow[1:1]{HA} Y$ *(т. е.* $AX = Y$ *и* $KerA = \{0\}$*). Тогда* $A^{-1} : Y \to X$ *— ограниченный линейный оператор.*

Доказательство. (см. книги по функциональному анализу). □

Лемма (о правом обратном). *Пусть* $A : X \to Y$ *— ограниченный линейный оператор, причем,* $AX = Y$*. Тогда существует отображение* $T : Y \to X$ *(вообще говоря, нелинейное и разрывное) и константа* $c > 0$ *такие, что*

$$A(T(y)) = y \quad \forall y \in Y$$

(т. е. $AT = I_Y : Y \to Y \Rightarrow T$ *— правый обратный), и*

$$\|T(y)\| \leqslant C\|y\| \quad \forall y \in Y \Rightarrow T$$

непрерывен в нуле.

Для доказательства мы используем понятия фактор-пространства и фактор-оператора. Напомним их.

Пусть $L \subset X$ — замкнутое подпространство. Фактор-простоpанство $\hat{X} = X/_L$ определяется следующим образом. Элементами $X/_L$ являются *классы смежности*

$$\hat{x} = x + L = \{x + x_0 \,|x_0 \in L\}.$$

Норма элемента $\hat{x}$ определяется так:

$$\|\hat{x}\| = \inf\{\|x\| | x \in \hat{x}\}$$

Пространство $\widehat{X}$, снабженное этой нормой, является банаховым. Нулевым элементом является L. Далее, пусть $A : X \to Y$ — линейный ограниченный оператор. Положим $L = KerA$. Рассмотрим $\widehat{X} = X/KerA$.

Определим *фактор-оператор* $\hat{A} : \widehat{X} \to Y$ следующим образом:

$$\hat{A}\hat{x} = Ax,$$

где $x \in \hat{x}$ — произвольный представитель класса. Это определение корректно: если $x_1 \in \hat{x}$ и $x_2 \in \hat{x}$, то $x_1 - x_2 \in KerA$ и, следовательно, $Ax_1 = Ax_2$.

Это отображение линейно и непрерывно, причем,$\|\hat{A}\| \leqslant \|A\|$. Кроме того, это отображение инъективно (проверьте!).

Доказательство леммы о правом обратном. Пусть $\hat{A} : \widehat{X} \xrightarrow{HA} Y$. Тогда $\hat{A} : \widehat{X} \xrightarrow[1:1]{HA} Y$. По теореме Банаха $\hat{A}^{-1} : Y \to X$ — ограниченный оператор. Положим $C > \|\hat{A}^{-1}\|$. Пусть $y \in Y$ — произвольный элемент, $y \neq 0$. Пусть

$$\hat{x} = \hat{A}^{-1}y.$$

Тогда

$$0 < \|\hat{x}\| \leqslant \|\hat{A}^{-1}\| \cdot \|y\| < C\|y\|.$$

Тогда $\exists\, x \in \hat{x} : \|x\| < C\|y\|$. Положим $T(y) = x$. Такой выбор возможен $\forall y \in Y$, $y \neq 0$. Если $y = 0$, то полагаем $T(y) = 0$ и, следовательно,

$$\|T(y)\| \leqslant C\|y\| \qquad \forall\, y,$$

$$A(T(y)) = Ax = y \qquad \forall\, y.$$

□

Итак, годится любое $C > \|\hat{A}^{-1}\|$.

Пусть B_x, B_Y — единичные шары в X и Y соответственно.

Теорема Банаха–Шаудера (об обратном отображении). *Пусть* $A : X \xrightarrow{HA} Y$ *— линейный ограниченный оператор. Тогда* $\exists\, a > 0$*:*

$$A(B_X) \supset aB_Y, \tag{4.1}$$

т. е. образ единичного шара в X накрывает *шар радиуса* a *в* Y *(с центром в нуле).*

Доказательство. Пусть $C > \|\hat{A}^{-1}\|$. Пусть $T : Y \to X$ — правый обратный к A такой, что

$$\|T(y)\| \leqslant C\|y\| \quad \forall\, y.$$

Тогда

$$T(B_Y) \subset C \cdot B_X.$$

Поскольку $AT = I_Y$, $I_Y : Y \to Y$ — тождественное отображение, то

$$B_Y \subset C \cdot A(B_X).$$

Следовательно,

$$A(B_X) \supset \frac{1}{C} B_Y, \text{ следовательно, } a = \frac{1}{C}\ .$$

□

Итак, годится любое положительное

$$a < \frac{1}{\|\hat{A}^{-1}\|}.$$

Условие (4.1) называется *свойством накрывания* с константой $a > 0$ для линейного оператора.

Всякий линейный сюръективный оператор накрывает с некоторой константой $a > 0$.

В силу линейности это свойство можно распространить на любые шары:

$$A(B_X(x,r)) \supset B_Y(Ax, ar), \quad \forall\, x \in X,\ \forall\, r > 0,$$

где $B_X(x,r)$ — шар в X радиуса r с центром в x (проверьте!!).

Именно это свойство принимается за определение накрывания для нелинейных операторов.

4.1.2 Накрывание для нелинейных операторов.

Пусть X, Y — метрические пространства, $F : X \to Y$ — оператор.

$B_X(x,r)$ — шар радиуса r с центром в точке x в X.

$B_Y(y,r)$ — шар радиуса r с центром в точке y в Y.

Пусть $U \subset X$ — открытое множество.

Определение. Мы скажем, что оператор F *накрывает на* U с константой $a > 0$, если $\forall B_X(x,r) \subset U$ имеем:

$$F(B_X(x,r)) \supset B_Y(F(x), ar) \tag{4.2}$$

Если есть накрывание с константой $a > 0$, то есть накрывание и с любой $a' < a$, $a' > 0$. (*Supremum* a, для которого есть накрывание называют *константой накрывания* F на U. Но мы не будем использовать этот термин).

Оператор $S : X \to Y$ назовем *стягивающим (сжимающим)* на U, с константой $b > 0$, если

$$d(S(x_1), S(x_2)) \leqslant bd(x_1, x_2) \quad \forall\, x_1, x_2 \in U \tag{4.3}$$

Условие (4.3) есть условие Липшица, а b — константа Липшица. Таким образом, стягивающий на U — это липшицев на U.

Оказывается, если накрывающий оператор «возмутить» стягивающим с малой константой, то свойство накрывания не нарушится. Это утверждается в следующей теореме о накрывании, принадлежащей А. А. Милютину и родственной теореме Люстерника о касательном многообразии (с последней нам еще предстоит познакомиться).

Теорема (о накрывании). *Пусть (X, d) — полное метрическое пространство, $(Y, \|\cdot\|)$ — банахово пространство, $U \subset X$ — открытое множество.*

Пусть $T : X \to Y$ накрывает на U с константой $a > 0$ и непрерывен на U, $S : X \to Y$ стягивает на U с константой $b > 0$.

Пусть $a > b$. Тогда оператор

$$F = T + S : X \to Y \qquad (F(x) := T(x) + S(x) \quad \forall\, x \in X)$$

накрывает на U с константой $a - b$.

Для доказательства нам понадобится лемма. Пусть $B_X(x_0, R)$ — фиксированный шар радиуса $R > 0$ с центром в x_0 в X (индекс X будем опускать). Накрывание и стягивание на $B_X(x_0, R)$ определим как и на U. По-прежнему X— полное метрическое, Y — банахово.

Лемма (о накрывании). *Пусть оператор $T : X \to Y$ непрерывен и накрывает на $B_X(x_0, R)$ с константой $a > 0$, а оператор $S : X \to Y$ стягивает на $B_X(x_0, R)$ с константой $b > 0$. Пусть $a > b$. Тогда для оператора $F = T + S$ выполнено включение*

$$F(B_X(x_0, R)) \supset B_Y(F(x_0), (a - b)R). \tag{4.4}$$

Доказательство. Пусть $y \in B_Y(F(x_0), (a - b)R)$ — произвольный элемент. Покажем, что $\exists\, x \in B_X(x_0, R)$: $F(x) = y$. Тем самым (4.4) будет установлено и лемма будет доказана.

Элемент x найдем как предел фундаментальной последовательности $x_0, x_1, \ldots, x_n, \ldots$ в X, обладающей при $n = 1, 2, \ldots$ свойствами:

(i) $T(x_n) + S(x_{n-1}) = y$,

(ii) $d(x_{n-1}, x_n) \leqslant (\frac{b}{a})^{n-1}(1 - \frac{b}{a})R$.

Из (ii) вытекает, что при $m \geqslant 0$, $p \geqslant 1$

$$\begin{aligned} d(x_m, x_{m+p}) &\leqslant d(x_m, x_{m+1}) + \ldots + d(x_{m+p-1}, x_{m+p}) \leqslant \\ &\leqslant \left(\left(\frac{b}{a}\right)^m + \ldots + \left(\frac{b}{a}\right)^{m+p-1}\right)\left(1 - \frac{b}{a}\right)R = \\ &= \left(\frac{b}{a}\right)^m \left(1 - \left(\frac{b}{a}\right)^p\right) R \to 0 \text{ при } m \to \infty. \end{aligned} \tag{4.5}$$

Следовательно, x_n — фундаментальна, и тогда $x_n \to x$.

Кроме того, из (4.5) вытекает, что (при $m = 0$, $p = n$)

$$d(x_0, x_n) \leqslant \left(1 - \left(\frac{b}{a}\right)^n\right) R < R \quad \forall\, n, \tag{4.6}$$

и следовательно, имеет место свойство:

(iii) $x_n \in B_X(x_0, \left(1 - (\frac{b}{a})^n\right) R) \subset B_X(x_0, R)$.

Переходя к пределу в условиях (i) и (iii) и пользуясь непрерывностью T и S (непрерывность S вытекает из липшицевости), получаем:

$$T(x) + S(x) = y \sim F(x) = y,\ x \in B_X(x_0, R).$$

Итак, требуется построить последовательность $\{x\}$ с начальным членом x_0, обладающую при $n = 1, 2, \dots$ свойствами (i) и (ii). Найдем x_1 из условия

$$T(x_1) + S(x_0) = y,$$

(расположенный возможно ближе к x_0) т. е.

$$T(x_1) = y - S(x_0),$$

пользуясь a-накрыванием оператора T.

Поскольку

$$y \in B_Y(T(x_0) + S(x_0), (a - b)R),$$

то

$$\|(y - S(x_0)) - T(x_0)\| \leq (a - b)R)$$

и, следовательно,

$$y - S(x_0) \in B_Y(T(x_0), (a - b)R)$$

По условию накрывания

$$T\left(B_X\left(x_0, (1 - \frac{b}{a})R\right)\right) \supset B_Y(T(x_0), (a - b)R).$$

Следовательно найдется

$$x_1 \in B_X(x_0, \left(1 - \frac{b}{a}\right) R)$$

такой, что

$$T(x_1) = y - S(x_0),$$

или

$$T(x_1) + S(x_0) = y,$$

При этом

$$d(x_0, x_1) \leqslant \left(1 - \frac{b}{a}\right) R$$

Таким образом, для x_1 $(n = 1)$ условия (i), (ii) и, следовательно, (iii) выполнены. Пусть построены первые n членов последовательности

$$x_0, x_1, \dots, x_n$$

Найдем x_{n+1} как решение уравнения

$$T(x_{n+1}) + S(x_n) = y.$$

По предположению индукции

$$T(x_n) + S(x_{n-1}) = y.$$

Следовательно, уравнение на x_{n+1} можно записать в виде:

$$T(x_{n+1}) = T(x_n) + (S(x_{n-1}) - S(x_n)). \tag{4.7}$$

Поскольку S липшицев с константой b, то

$$\|S(x_{n-1}) - S(x_n)\| \leqslant b\, d(x_{n-1}, x_n) \overset{\text{(ii)}}{\leqslant} \overset{\text{(ii)}}{\leqslant} b\,(\tfrac{b}{a})^{n-1}(1 - \tfrac{b}{a})R = a(\tfrac{b}{a})^n(1 - \tfrac{b}{a})R.$$

Следовательно,

$$T(x_n) + S(x_{n-1}) - S(x_n) \in B_Y(T(x_n), a\left(\frac{b}{a}\right)^n \left(1 - \frac{b}{a}\right) R). \tag{4.8}$$

Рассмотрим шар:

$$B_n := B_X(x_n, \left(\frac{b}{a}\right)^n \left(1-\frac{b}{a}\right) R).$$

Имеем: $B_n \subset B_X(x_0, R)$, так как

$$d(x_0,x_n) + \left(\frac{b}{a}\right)^n \left(1-\frac{b}{a}\right) R \overset{(6)}{\leqslant} R\left(1-\left(\frac{b}{a}\right)^n\right) R + \left(\frac{b}{a}\right)^n \left(1-\frac{b}{a}\right) R =$$

$$= R\left(1-\left(\frac{b}{a}\right)^n + \left(\frac{b}{a}\right)^n - \left(\frac{b}{a}\right)^{n+1}\right) =$$

$$= R\left(1-\left(\frac{b}{a}\right)^{n+1}\right) < R.$$

Следовательно, в силу a-накрывания оператора T имеем:

$$T(B_n) \supset B_Y(T(x_n), a\left(\frac{b}{a}\right)^n \left(1-\frac{b}{a}\right) R).$$

Отсюда и из условия (4.8) вытекает существование $x_{n+1} \in B_n$, удовлетворяющего уравнению (4.7), а значит и условию (i). Из условия $x_{n+1} \in B_n$ вытекает, что

$$d(x_n, x_{n+1}) \leqslant \left(\frac{b}{a}\right)^n \left(1-\frac{b}{a}\right) R,$$

т. е. выполнено условие (ii) → (iii). Индукция завершена.

Лемма доказана. □

Доказательство теоремы о накрывании получается применением леммы к каждому шару, содержащемуся в U.□

Глава 5

Лекция 5.

5.1 Тема: Аппарат теории экстремума (продолжение).

5.1.1 Накрывание и метрическая регулярность.

Пусть (X, d) и (Y, d) — метрические пространства, $F : X \to Y$ — оператор.

Определение. $F : X \to Y$ *накрывает* с константой $a > 0$ на X, если

$$F\big(B(x,r)\big) > B\big(F(x), ar\big) \qquad \forall\, B(x,r) \subset X, \tag{5.1}$$

где $B(X, r)$ — замкнутый шар радиуса $r \geqslant 0$ с центром в точке x в X.

Определение. Отображение $F : X \to Y$ *метрически регулярно* с константой $K > 0$ на $X \times Y$, если

$$d\big(x, F^{-1}(y)\big) \leqslant K d\big(F(x), y\big) \qquad \forall\, x \in X, \quad \forall\, y \in Y \tag{5.2}$$

Утверждение. *Накрывание на X с константой $a > 0$ влечет метрическую регулярность на $X \times Y$ с константой $k = \frac{1}{a}$.*

Доказательство. Пусть $F : X \to Y$ накрывает на X с константой $a > 0$. Пусть $x \in X$, $y \in Y$. По условию накрывания для

$$r = d\big(F(x), y\big)$$

имеем:

$$F\big(B(x, \frac{r}{a})\big) \supset B\big(F(x), r\big).$$

При этом $y \in B\big(F(x), r\big)$. Следовательно, найдется $x_y \in B\big(F(x), r\big)$ такой, что $F(X_y) = y$.
При этом

$$d(x, x_y) \leqslant \frac{r}{a} = \frac{1}{a}\, d\big(F(x), y\big).$$

Поскольку $x_y \in F^{-1}(y)$, то

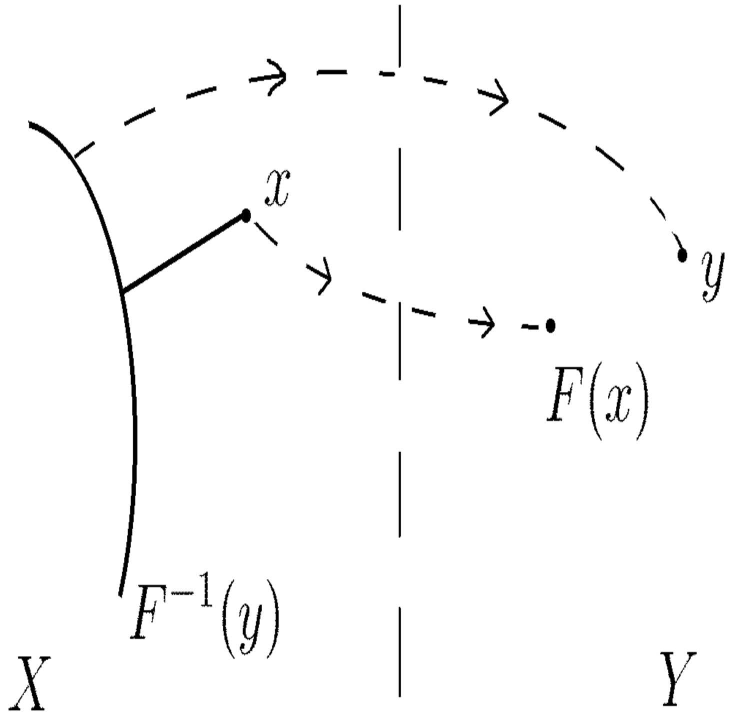
$$d\big(x, F^{-1}(y)\big) \leqslant d(x, x_y).$$

Таким образом

$$d\big(x, F^{-1}(y)\big) \leqslant \frac{1}{a} d\big(F(x), y\big).$$

□

Метрическая регулярность означает, что расстояние от точки x до уровня y функции F оценивается первой степенью близости образа $F(x)$ к y.



Очень часто эта оценка применяется для $y = 0$ — расстояния до нулевого уровня оператора. Наличие оценки (5.2) предполагает непустоту множества $F^{-1}(y)$ — полного прообраза точки y. Если $F^{-1}(y) = \varnothing$, то по определению $d\big(x, F^{-1}(y)\big) = +\infty$ — расстояние от точки до пустого множества равно $+\infty$. У понятий «накрывание» и «метрическая регулярность» имеется «локальный вариант». Пусть $x_0 \in X$ — фиксированная точка, $y_0 = F(x_0)$.

Определение. *F накрывает с* константой $a > 0$ в окрестности точки x_0, если существует окрестность U точки x_0, в которой F накрывает с константой a, т. е.

$$F(B(x,r)) \supset B(F(x), ar) \quad \forall\, B(x,r) \subset U.$$

Определение. F метрически регулярно с константой $k > 0$ в окрестности точки x_0, если существуют окрестность U точки $x_0 = F(x_0)$ в X и окрестность V точки $x_0 = F(x_0)$ в Y, такие, что

$$d(x, F^{-1}(y)) \leqslant kd(F(x), y) \quad \forall\, x \in U,\ \forall\, y \in V.$$

Замечание. *Правильнее было бы сказать: в окрестности точки (x_0, y_0) в $X \times Y$.*

Теорема. *Если F накрывает в окрестности точки x_0 с константой $a > 0$, то F метрически регулярно в окрестности той же точки с константой $k = \frac{1}{a}$.*

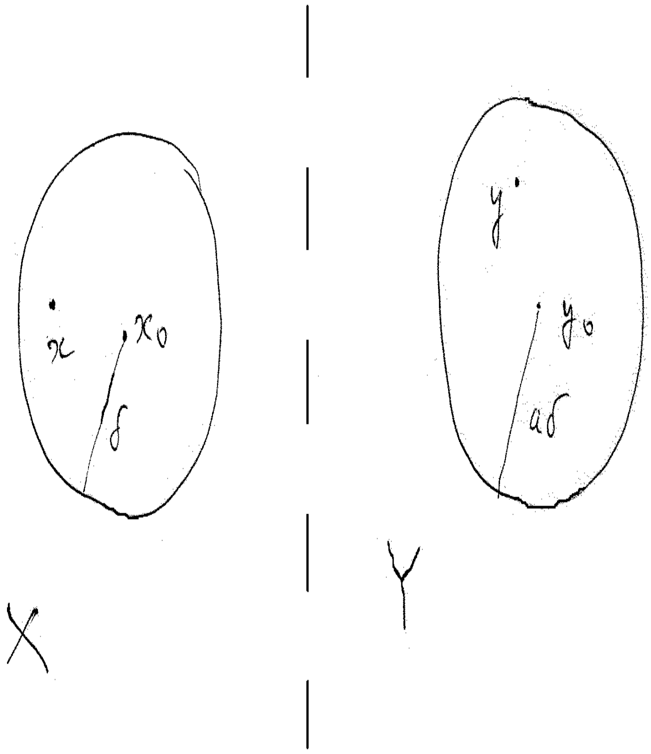
Доказательство. Пусть $\varepsilon > 0$ таково, что

$$F(B(x,r)) \supset B(F(x), ar) \quad \forall B(x,r) \subset B(x_0, \varepsilon) \tag{5.3}$$

Положим $\delta = \frac{\varepsilon}{3}$ и покажем, что

$$d(x, F^{-1}(y)) \leqslant \frac{1}{a} d(F(x), y) \quad \forall x \in B(x_0, \delta)\ \forall y \in B(y_0, a\delta) \tag{5.4}$$

Тем самым теорема будет доказана. Итак, пусть $x \in B(x_0, \delta)$, $y \in B(y_0, a\delta)$. Установим оценку (5.4).



Рассмотрим два случая.

а) Пусть $d(F(x), y) \geqslant 2\delta a$ ($F(x)$ и y «далеки» друг от друга).

Поскольку $\delta < \varepsilon$, то по условию накрывания

$$F(B(x_0, \delta)) \supset B(y_0, a\delta)$$

Так как $y \in B(y_0, a\delta)$, то найдется $x_y \in B(x_0, \delta)$ такой, что $F(x_y) = y$, т. е. $x_y \in F^{-1}(y)$. При этом

$$d(x, F^{-1}(y)) \leqslant d(x, x_y) \leqslant d(x, x_0) + d(x_0, x_y) \leqslant \delta + \delta \leqslant \frac{1}{a}\, d(F(x), y),$$

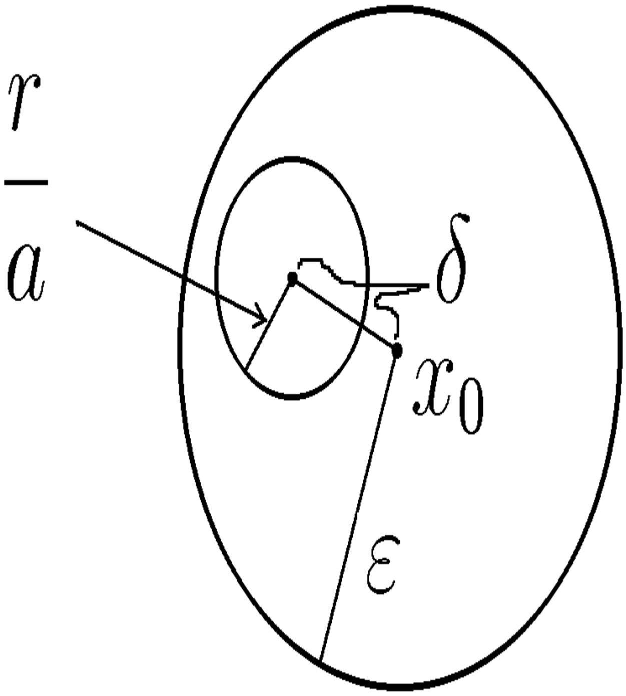
что и требовалось.

б) Пусть теперь $r := d(F(x), y) < 2\delta a$ (т. е. $F(x)$ и y «близки»). Тогда

$$\frac{r}{a} + \delta < \varepsilon.$$

Следовательно, по неравенству треугольника

$$B(x, \frac{r}{a}) \subset B(x_0, \varepsilon).$$



Тогда по условию накрывания

$$F(B(x, \frac{r}{a})) \supset B(F(x), r).$$

Но $y \in B(F(x), r)$ (см. определение r). Следовательно, найдется $x_y \in B(x, \frac{r}{a})$, такой что $F(x_y) = y$. При этом

$$d(x, F^{-1}(y)) \leqslant d(x, x_y) \leqslant \frac{r}{a} = \frac{1}{a}\, d(F(x), y),$$

что и требовалось.

□

5.1.2 Дифференцируемость по Фреше и строгая дифференцируемость для оператора.

Пусть X,Y — банаховы пространства, U — окрестность точки $x_0 \in X$, $g : U \to Y$ — оператор.

Определение. Оператор g *дифференцируем по Фреше* в точке x_0, если существует ограниченный линейный оператор $A : X \to Y$, такой что

$$g(x_0 + h) = g(x_0) + Ah + r(x_0, h),$$

где $\|r(x_0, h)\| = o(\|h\|)$, т. е. $\frac{\|r(x_0,h)\|}{\|h\|} \to 0$ при $\|h\| \to 0$.

Иначе $r(x_0, h) = \tilde{y}(x_0, h)\|h\|$, где $\|\tilde{y}(x_0, h)\| \to 0$ при $\|h\| \to 0$. Оператор A называется *производной Фреше* оператора g в точке x_0. Полагают

$$A = g'(x_0).$$

В частности, при $Y = \mathbb{R}^1$ мы получаем определение функционала, дифференцируемого по Фреше в точке x_0.

Пример. *Рассмотрим*

$$y(x) = \int_{t_0}^{t_1} F(t, x, \dot{x})\, dt,$$

где $x(\cdot) \in \mathbb{C}^1(\Delta, \mathbb{R}^n)$ $\quad \Delta = [t_0, t_1]$. *Пусть* $F, F_x, F_{\dot{x}}$ *непрерывны по всем переменным. Тогда*

$$J(x + h) = J(x) + \int_{t_0}^{t_1} (F_x h + F_{\dot{x}} \dot{g}\, dt + o(\|h\|_{e^1})),$$

где

$$F_x = F_x(t, x(t), \dot{x}(t)),$$
$$F_{\dot{x}} = F_x(t, x(t), \dot{x}(t)).$$

Эту оценку мы получали ранее. Из нее вытекает, что y дифференцируем по Фреше в точке x и выражение $\int_{t_0}^{t_1} F_x h + F_{\dot{x}} \dot{h}\, dt$ есть производная Фреше с этого функционала в точке x в пространстве $\mathbb{C}^1$. Аналогичное утверждение имеет место в пространстве $\mathbb{W}_1^1$ (покажите!!).

Оператор g называют *непрерывно дифференцируемым* по Фреше в точке x_0, если он имеет производную Фреше $g'(x)$ в окрестности точки x_0 и при этом

$$\|g'(x) - g'(x_0)\| \to 0 \quad \text{при } \|x - x_0\| \to 0.$$

Оператор g называют *строго дифференцируемым* в точке x_0, если существует ограниченный линейный оператор $A : X \to Y$ такой, что $\forall\, \varepsilon > 0\ \exists\, \delta > 0$ такое, что для любых x_1, x_2, удовлетворяющих неравенствам

$$\|x_1 - x_0\| < \delta,$$
$$\|x_2 - x_0\| < \delta$$

выполнено неравенство

$$\|g(x_2) - g(x_1) - A(x_2 - x_1)\| \leqslant \varepsilon \|x_2 - x_1\|. \tag{5.2}$$

Строгая дифференцируемость в точке сильнее, чем дифференцируемость по Фреше в этой точке. Действительно на языке $\varepsilon - \delta$ дифференцируемость по Фреше можно сформулировать так: $\forall\, \varepsilon > 0\ \exists\, \delta > 0$ такое, что из условия $\|x - x_0\| < \delta$ вытекает, что

$$\|g(x) - g(x_0) - A(x - x_0)\| \leqslant \varepsilon \|x - x_0\|.$$

Отсюда ясно, что строгая дифференцируемость $\Rightarrow$ дифференцируемость по Фреше, и что в определении строгой дифференцируемости $A = g'(x_0)$.

Покажем, что из непрерывной дифференцируемости в точке вытекает строгая дифференцируемость в этой точке. Для этого нам понадобится теорема о среднем для операторов.

Теорема (о среднем). *Пусть X, Y — банаховы, $U \subset X$ — открытое множество, $[a, b] \subset U$ — отрезок, $f : U \to Y$ — оператор, дифференцируемый по Фреше в каждой точке $x \in [a, b]$. Тогда*

$$\|f(b) - f(a)\| \leqslant \sup_{x \in [a,b]} \|f'(x)\|\, \|b - a\|.$$

Замечание. *Условия этой теоремы можно ослабить: считать, что X, Y — нормированные пространства, а f — дифференцируем по «Гато» в каждой точке отрезка $[a, b]$ (т. е. в более слабом смысле). Подробности см. в АТФ стр. 148, там же приведено и доказательство.*

Теперь покажем, что непрерывная дифференцируемость в точке сильнее, чем строгая. Пусть $g(x)$ непрерывно дифференцируем по Фреше в точке x_0. Рассмотрим оператор

$$\varphi(x) = g(x) - g'(x_0)x.$$

Для него при x близком к x_0 имеем:

$$\varphi'(x) = g'(x) - g'(x_0),$$

и следовательно

$$\|\varphi'(x)\| \to 0 \quad \text{при } \|x - x_0\| \to 0.$$

Тогда

$$\begin{aligned} \|g(x_2) - g(x_1) - g'(x_0)(x_2 - x_1)\| &= \|\varphi(x_2) - \varphi(x_1)\| \leqslant \\ &\leqslant \sup_{x \in [x_1, x_2]} \|\varphi'(x)\|\, \|x_2 - x_1\|. \end{aligned} \tag{5.3}$$

Поскольку $\sup_{x \in [x_1, x_2]} \|\varphi'(x)\| \to 0$ при $\|x_1 - x_0\| + \|x_2 - x_0\| \to 0$, то g строго дифференцируем в точке x_0. Итак, доказана

Теорема. *Непрерывная дифференцируемость по Фреше влечет строгую дифференцируемость в этой точке.*

5.1.3 Накрывание и метрическая регулярность для строго дифференцируемого оператора.

Пусть $g : U \to Y$ строго дифференцируем в точке $x_0 \in U$. Представим его в виде

$$g(x) = g'(x_0)x + (g(x) - g'(x_0)x). \tag{5.4}$$

Рассмотрим оператор

$$S(x) = g(x) - g'(x_0)x. \tag{5.5}$$

Для $x_1, x_2 \in U$ имеем

$$S(x_2) - S(x_1) = g(x_2) - g(x_1) - g'(x_0)(x_2 - x_1).$$

Из определения строгой дифференцируемости оператора g в точке x_0 вытекает, что $\forall\, \varepsilon > 0\ \exists\, \delta > 0$, такое, что

$$\|S(x_2) - S(x_1)\| \leqslant \varepsilon\|x_2 - x_1\| \qquad \forall x_1, x_2 \in B(x_0, \delta),$$

т. е. оператор S удовлетворяет условию Липшица (является стягивающим) с константой $\varepsilon > 0$ на шаре $B(x_0, \delta)$. При этом ε можно сделать сколь угодно малым за счет выбора достаточно малого δ.

Теперь мы легко докажем следующий критерий накрывания в окрестности точки x_0 для нелинейного оператора $g : Y \to Y$.

Теорема (достаточное условие для накрывания). *Пусть оператор g строго дифференцируем в точке x_0 и пусть $g'(x_0)X = Y$. Тогда существуют $a > 0$ и окрестность точки x_0 такие, что оператор g накрывает с константой a в этой окрестности, а значит g метрически регулярен в окрестности точки x_0.*

Доказательство. Имеем

$$g(x) = g'(x_0)x + S(x),$$

где $S(x) = g(x) - g'(x_0)x$. Поскольку $g'(x_0)X{=}$Y, то оператор $T(x) = g'(x_0)x$ накрывает на X с некоторой положительной константой $a_0 > 0$. Оператор $S(x)$ стягивает на $B(x_0, \delta)$ с константой $\delta > 0$, причем, ε можно выбрать сколь угодно малым за счет выбора достаточно малого $\delta > 0$. Пусть $a_0 - \varepsilon > 0$. По теореме Милютина о накрывании оператор $g = T + S$ накрывает с константой $a_0 - \varepsilon$ в окрестности точки x_0. □

Замечание. *Из доказательства видно, что годится любое $a < a_0$, где a_0 — константа накрывания для $g'(x_0)$. В свою очередь*

$$a_0 < \frac{1}{\|\hat{A}^{-1}\|}\ ,$$

где $A = g'(x_0)$, $\hat{A}$ — фактор-оператор, а a_0 может быть выбрано сколь угодно блитзкой к величине $\frac{1}{\|\hat{A}^{-1}\|}$. Значит, и a может быть выбрано сколь угодно близким к той же величине.

Пусть X,Y — банаховы, $U \subset X$ — открытое множество, $g : U \to Y$, $x_0 \in U$. Пусть $g(x_0) = 0$.

Определение. Будем говорить, что в окрестности точки x_0 оператор g обладает *оценкой расстояния до нулевого уровня*, если существуют $\delta > 0$, $k > 0$, такие что $\forall\, x \in B(x_0, \delta)\ \ \exists\, \bar{x} \in X$, такой что

$$g(x + \bar{x}) = 0, \qquad \|\bar{x}\| \leqslant k\|f(x)\|.$$

Глава 6

Лекция 6.

6.1 Аппарат теории экстремума (продолжение)

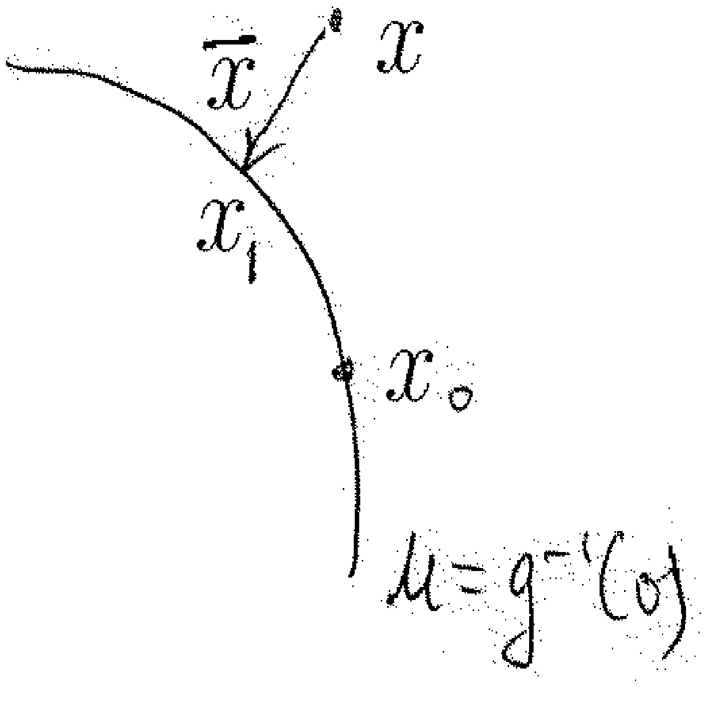
6.1.1 Оценка расстояния до нулевого уровня нелинейного оператора.

Пусть X,Y — банаховы пространства, $U \subset X$ — открытое множество, $g : U \to Y$ — оператор,

$$M = \{x \in U \mid g(x) = 0\} = g^{-1}(0)$$

— нулевой уровень этого оператора.

Теорема (о поправке). *Пусть $x_0 \in U$ такова, что $g(x_0) = 0$ (т. е. $x_0 \in M$), g строго дифференцируем в точке x_0 и $g'(x_0)X = Y$. Тогда существуют окрестность $B(x_0, \delta)$ ($\delta > 0$) точки x_0 и $K > 0$ такие, что $\forall\, x \in B(x_0, \delta)$ $\exists\, \bar{x} \in X$ такой, что $g(x + \bar{x}) = 0$; $\|\bar{x}\| \leqslant K\|g(x)\|$ и, следовательно, $d(x, M) \leqslant K\|g(x)\|$.*



Доказательство. Пусть в точке x_0 выполнены условия теоремы. Тогда g метрически регулярен в окрестности точки x_0. Следовательно, существуют $\delta > 0$ и $K_0 > 0$ такие, что

$$d(x, g^{-1}(0)) \leqslant K_0\|g(x)\| \qquad \forall x \in B(x_0, \delta)$$

(в определении метрической регулярности следует положить $y = 0$).

Пусть $\varepsilon > 0$. Положим $K = (1 + \varepsilon)K_0$. Покажем, что с таким K поправка существует.

а) Если $g(x) = 0$, то полагаем $\bar{x} = 0$.

б) Пусть $\|g(x)\| > 0$, но $d(x, g^{-1}(0)) = 0$. Тогда найдется x_1 такой, что

$$g(x_1) = 0, \qquad \|x - x_1\| \leqslant K\|g(x)\|.$$

Для $\bar{x} = x - x_1$ оценка имеет место.

в) Пусть $\|g(x)\| > 0$ и $d(x, g^{-1}(0)) > 0$. Тогда найдется $x_1 \in g^{-1}(0)$ такой, что

$$\|x_1 - x\| \leqslant (1 + \varepsilon) d(x, g^{-1}(0)).$$

Для $\bar{x} = x_1 - x$ оценка имеет место. Теорема доказана. □

6.1.2 Теорема Люстерника о касательном многообразии.

Пусть X — банахово, $M \subset X$ — множество, $x_0 \in M$ — точка.

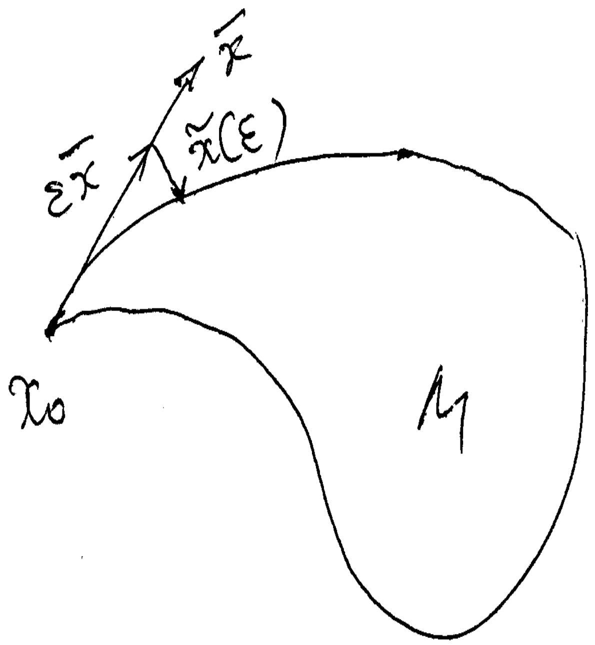
Определение. Вектор $\bar{x}$ называется (односторонним) *касательным* вектором к M в точке x_0, если существует «кривая»

$$\tilde{x}(\varepsilon) : (0, \varepsilon_0) \to X \qquad (\varepsilon_0 > 0)$$

такая, что

1) $x_0 + \varepsilon\bar{x} + \tilde{x}(\varepsilon) \in M \quad \forall \varepsilon \in (0, \varepsilon_0)$,

2) $\frac{\|\tilde{x}(\varepsilon)\|}{\varepsilon} \to 0$ при $\varepsilon \to +0$, т. е. $\|\tilde{x}(\varepsilon)\| = o(\varepsilon)$,

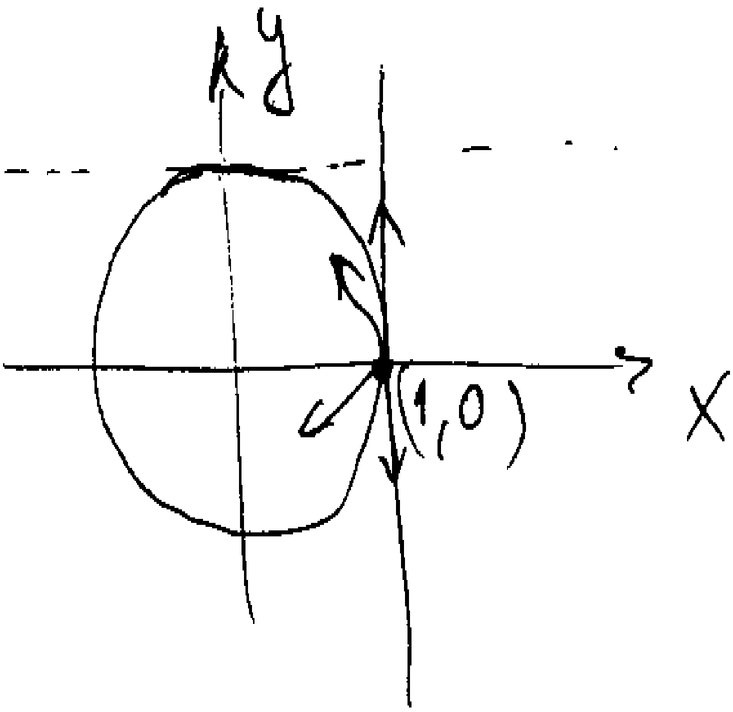


Обозначим через $TM(x_0)$ множество всех касательных векторов в точке x_0.
Если $\bar{x} \in TM(x_0)$ и $\lambda > 0$, то $\lambda\bar{x} \in TM(x_0)$. Следовательно, $TM(x_0)$ — конус (с вершиной с нуле).

Примеры. 1^0. *Если $x_0 \in \inf M$, то $TM(x_0) = X$.*

2^0. *Пусть $M = \{x \in \mathbb{R}^n \mid x > 0\} \bigcup \{0\}$. Тогда касательный конус в нуле есть замыкание множества M.*

3^0. *Пусть* $M = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leqslant 1\}$. *Касательный конус в точке* $(1, 0)$ *есть полуплоскость* $\{(x, y) \in \mathbb{R}^2 \mid x \leqslant 0\}$,



4^0. *Пусть* $M = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$. *Касательный конус в точке* $(1, 0)$ *есть подпространство* $\{(x, y) \in \mathbb{R}^2 \mid x = 0\}$ *— касательное подпространство.*

Теорема (Л. А. Люстерник). *Пусть* X,Y *— банаховы,* $U \subset X$ *— открытое множество,* $x_0 \in U$ *— точка,* $g(x_0) = 0$. *Пусть*

$$M = \{x \in U \mid g(x) = 0\}.$$

Пусть оператор g *строго дифференцируем в точке* x_0 *и* $g'(x_0)X = Y$ *(условие Люстерника). Тогда*

$$TM(x_0) = \{\bar{x} \in X \mid g'(x_0)\bar{x} = 0\},$$

т. е. касательный конус к множеству нулей оператора есть подпространство, совпадающее с ядром его производной.

Предложение. *Независимо от выполнения условия Люстерника имеет место включение*

$$TM(x_0) \subset Ker\, g'(x_0).$$

Доказательство. Пусть $\bar{x} \in TM(x_0)$.Тогда $\exists$ функция

$$\tilde{x}(\varepsilon) : (0, \varepsilon_0) \to X \qquad (\varepsilon_0 > 0)$$

такая, что

$$g(x_0 + \varepsilon\bar{x} + \tilde{x}(\varepsilon)) = 0 \qquad \forall \varepsilon \in (0, \varepsilon_0) \qquad \|\tilde{x}(\varepsilon)\| = o(\varepsilon)$$

Следовательно,

$$0 = g(x_0) + g'(x_0)(\varepsilon\bar{x} + \tilde{x}(\varepsilon)) + r(\varepsilon),$$

где $\|r(\varepsilon)\| = o(\varepsilon)$. Но $g(x_0) = 0$, следовательно,

$$g'(x_0)(\varepsilon\bar{x}) + r_1(\varepsilon) = 0,$$

где $\|r_1(\varepsilon)\| = o(\varepsilon)$.

Деля это равенство на ε и переходя к пределу по $\varepsilon \to +0$, получаем, что $g'(x_0)\bar{x} = 0$. Следовательно, $\bar{x} \in Ker\, g'(x_0)$. $\square$

Доказательство теоремы Люстерника. Пусть выполнены условия теоремы. Требуется показать, что

$$Ker\, g'(x_0) \subset TM(x_0).$$

Пусть $g'(x_0)\bar{x} = 0$. Тогда

$$g(x_0 + \varepsilon\bar{x}) = g(x_0) + g'(x_0)\varepsilon\bar{x} + r(\varepsilon) = r(\varepsilon),$$

где $\|r(\varepsilon)\| = o(\varepsilon)$.

Поскольку g обладает оценкой расстояния до нулевого уровня в окрестности точки x_0, то по теореме о поправке существуют $K > 0$ и функция $\tilde{x}(\varepsilon) : (0, \varepsilon_0) \to X$ $(\varepsilon > 0)$ такие , что

$$g(x_0 + \varepsilon\bar{x} + \tilde{x}(\varepsilon)) = 0,$$

$$\|\tilde{x}(\varepsilon)\| \leqslant K\|g(x_0 + \varepsilon\bar{x})\| = o(\varepsilon).$$

Следовательно, $\bar{x} \in TM(x_0)$. □

Отметим, что теорема Люстерника доказана в предположении о строгой дифференцируемости g в точке x_0 вместо непрерывной дифференцируемости в этой точке, которая предполагалась в первоначальной формулировке теоремы Люстерника. Заметим также, что понятия: «накрывание», «метрическая регулярность» и связанные с ними теоремы, возникли из анализа доказательства теоремы Люстерника и так или иначе были с ней связаны.

6.1.3 Отделимость выпуклых множеств.

Пусть X — банахово пространство, $x_1, x_2 \in X$. Через $[x_1, x_2]$ мы обозначаем отрезок в X с концами x_1, x_2, т. е. множество точек, представимых в виде $x = \alpha x_1 + \beta x_2$, где $\alpha \geqslant 0$, $\beta \geqslant 0$, $\alpha + \beta = 1$.

Множество $M \subset X$ — называется *выпуклым*, если вместе с каждыми двумя точками оно содержит соединяющий их отрезок. Через X^* мы обозначаем сопряженное пространство к X, состоящее из всех линейных непрерывных функционалов $x^* : X \to \mathbb{R}$. По определению

$$\|x^*\| = \sup_{\|x\| \leqslant 1} \langle x^*, x\rangle$$

— норма в X^*. Пространство X^* является банаховым.

Пусть A,B — два множества в X.

Определение. Функционал $x^* \in X^*$, $x^* \neq 0$ *разделяет* A и B, если

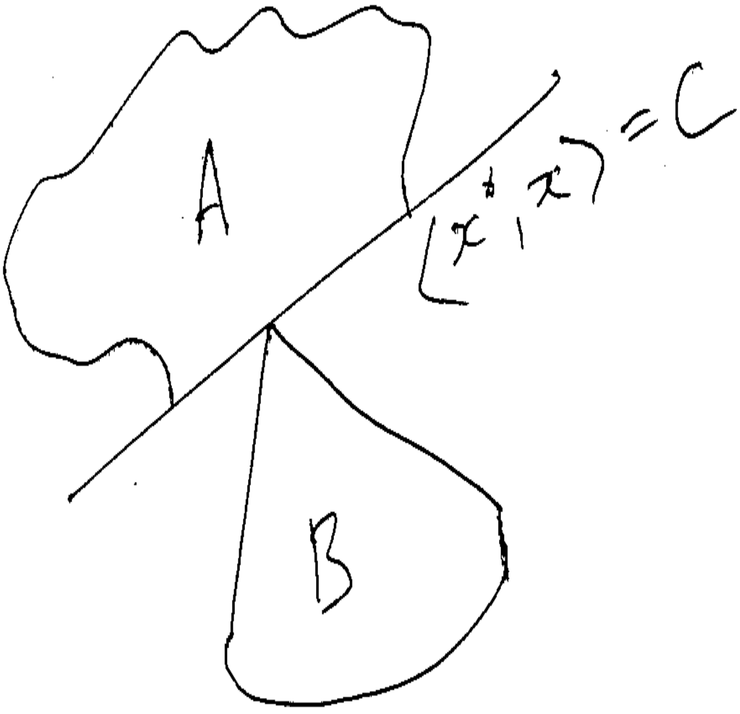
$$\sup_{x\in A}\langle x^*, x\rangle \leqslant \inf_{x\in B}\langle x^*, x\rangle \tag{6.1}$$

Пусть c — число, заключенное между левой и правой частями неравенства (6.1). Тогда гиперплоскость $\{x \mid \langle x^*, x\rangle = c\}$ *разделяет* A и B (или *отделяет* A и B друг от друга) в том смысле, что A лежит в полупространстве

$$\{x \mid \langle x^*, x\rangle \leqslant c\},$$

а B лежит в полупространстве

$$\{x \mid \langle x^*, x\rangle \geqslant c\},$$



(см. АТФ, стр. 124)

Теорема отделимости. *Пусть A — выпуклое открытое множество, B — выпуклое множество. Пусть $A \cap B = \varnothing$. Тогда существует ненулевой линейный непрерывный функционал, разделяющий A и B.*

Упражнение. *Докажите обратное.*

Пусть $M \subset X$ — множество, $x^* \in X^*$ — функционал.

Определение. Функционал x^* называется *опорным* к множеству M, если

$$\inf x^*(M) := \inf_{x \in M} \langle x^*, x \rangle > -\infty.$$

Это означает, что $M \subset \{x \mid \langle x^*, x \rangle \geqslant a\}$ при некотором a. Множество всех опорных к M обозначим через M^*.

Покажите, что M^* — выпуклый конус.

Напомним, что $K \subset X$ — конус, если из условий $x \in K, \lambda > 0 \Rightarrow \lambda x \in K$. Конус K является выпуклым $\Leftrightarrow$ из условий $x, y \in K \Rightarrow x + y \in K$ (покажите !!).

Пусть K — конус. Тогда

$$x^* \in K^* \Leftrightarrow \langle x^*, x \rangle \geqslant 0 \qquad \forall x \in K$$

(покажите !!)

Примеры. 1^0. *Пусть $K = \{x \in \mathbb{R}^n \mid x > 0\}$.*

Тогда $K^* = \{y \in \mathbb{R}^n \mid y \geqslant 0\}$.

2^0. *Пусть $M = \{x \in \mathbb{R}^2 \mid x_1 > 0, x_2 > 0, x_1 x_2 \geqslant 1\}$.*

Тогда $M^* = \{y \in \mathbb{R}^2 \mid y \geqslant 0\}$.

Пусть A и B два множества в X. Покажите, что A и B можно отделить $\Leftrightarrow$ $\exists$ ненулевые функционалы x^* и y^* такие, что

$$x^* \in A^*, \quad y^* \in B^*,$$
$$x^* + y^* = 0,$$
$$\inf \langle x^*, A \rangle + \inf \langle x^*, B \rangle \geqslant 0.$$

При такой формулировке понятия отделимости теорему отделимости можно обобщить на конечное число множеств.

Теорема (Дубовицкий и Милютин). *Пусть в банаховом пространстве X заданы непустые выпуклые открытые множества $M_1, M_2, \dots, M_n$ и непустое выпуклое множество M. Тогда условие*

$$M_1 \cap \ldots \cap M_n \cap M = \varnothing$$

эквивалентно существованию набора функционалов

$$x_1^*, \dots, x_n^*, x^*$$

такого, что

$$x_1^* + \ldots + x_n^* + x^* = 0, \tag{a}$$

$$\|x_1^*\| + \ldots + \|x_n^*\| + \|x^*\| > 0, \tag{b}$$

$$\inf\langle x_1^*, M\rangle + \ldots + \inf\langle x_n^*, M_n\rangle + \inf(x^*, M) \geqslant 0. \tag{c}$$

Из последнего условия вытекает, что

$$x_1^* \in M_1^*, \dots, x_n^* \in M_n^*, \quad x^* \in M^*.$$

Доказательство. Для доказательства нам понадобится следующий факт. Пусть $X_1, X_2, \dots, X_n$ — банаховы пространства, $\widehat{X} = X_1 \times \ldots \times X_n$ — их произведение. Тогда $\hat{x}^* \in \widehat{X}^* \Leftrightarrow \exists\, x_k^* \in X_k^*$ такие, что для любого

$$\hat{x} = (x_1, \dots, x_n) \in \widehat{X}$$

имеем

$$\langle \hat{x}^*, \hat{x}\rangle = \langle x_1^*, x_1\rangle + \ldots + \langle x_n^*, x_n\rangle$$

(покажите это !!).

Приступим к доказательству теоремы.

а) *Необходимость.*

Пусть

$$M_1 \cap \ldots \cap M_n \cap M = \varnothing.$$

В произведении

$$\widehat{X} = \underbrace{X \times X \ldots \times X}_{\text{n раз}} = X^n$$

рассмотрим множество A, состоящее из всех наборов вида

$$\hat{x} = (x_1 - x, x_2 - x, \dots, x_n - x),$$

где $x_1 \in M_1, \dots, x_n \in M_n, x \in M$. Тогда A не содержит нуля $\hat{0}$ пространства $\widehat{X}$ (в противном случае $x_1 = x_2 = \ldots = x_n = x$ — общий элемент множеств $M_1, M_2, \dots, M_n, M$). Множество A открыто, ибо открыто произведение

$$M_1 \times M_2 \ldots \times M_n,$$

а A есть объединение его «сдвигов» на элемент $-(x, \dots, x)$ по всем $x \in M$. Далее, из выпуклости $M_1, M_2, \dots, M_n, M$ вытекает выпуклость A (проверьте !!). Следовательно, $\exists$ ненулевой линейный непрерывный функционал $\hat{x^*} = (x_1^*, \dots, x_n^*)$, отделяющий A от нуля. Не ограничивая общности считаем, что $\hat{x^*}(A) \geqslant 0$, т. е.

$$\langle x_1^*, x_1 - x\rangle + \ldots + \langle x_n^*, x_1 - x\rangle \geqslant 0 \qquad \forall\, x_1 \in M_1, \dots, x_n \in M_n, \quad x \in M.$$

Положим $x^* = -x_1^* - \ldots - x_n^*$. Тогда $x_1^* + \ldots + x_n^* + x^* {=} 0$, не все функционалы $x_1^*, \dots, x_n^*$ — нулевые (набор нетривиален) и

$$\langle x_1^*, x_1\rangle + \ldots + \langle x_n^*, x_n\rangle + \langle x^*, x\rangle \geqslant 0 \qquad \forall\, x_1 \in M_1, \dots, x_n \in M_n, \quad x \in M.$$

Последнее условие равносильно тому, что

$$\inf x_1^*(M_1) + \ldots + \inf x_n^*(M_n) + \inf x^*(M) \geqslant 0.$$

Необходимость даказана.

б) *Достаточность.*

Пусть $\exists\ x_1^*, \ldots, x_n^*, x^*$, удовлетворяющие условиям (a), (b), (c) теоремы, но при этом пересечение $M_1 \cap \ldots \cap M_n \cap M$ непусто.

Пусть x_0 — общий элемент всех множеств $M_1, M_2, \ldots, M_n, M$.

Тогда
$$\langle x_1^*, x_0\rangle + \ldots + \langle x_n^*, x_0\rangle + \langle x^*, x_0\rangle = 0.$$

Выберем среди $x_1^*, \ldots, x_n^*$ ненулевой функционал. Пусть это будет x_1^*. Найдется элемент $\bar{x} \in X$ такой, что $\langle x_1^*, \bar{x}\rangle < 0$. Поскольку $x_0 \in M_1$ и M_1 открыто, то $x_0 + \varepsilon\bar{x} \in M_1$ при малом $\varepsilon > 0$. Тогда

$$\langle x_1^*, x_0 + \varepsilon\bar{x}\rangle + \langle x_2^*, x_0\rangle + \ldots + \langle x_n^*, x_0\rangle + \langle x^*, x_0\rangle < 0,$$

где $x_0 + \varepsilon\bar{x} \in M_1\ x_0 \in M_2, \ldots, x_0 \in M_n, x_0 \in M \Rightarrow$

$$\inf x_1^*(M_1) + \inf x_2^*(M_2) + \ldots + \inf x_n^*(M_n) + (x^*, M) < 0$$

— противоречие.

Достаточность доказана.

Теорема доказана.

$\square$

Мы уже отмечали, что если K — конус и $\inf x^*(K) > -\infty$, то $\inf x^*(K) = 0$. Последнее означает, что $\langle x^*, x\rangle \geqslant 0 \quad \forall\, x \in K$, т. е. $x^* \in K^*$. Поэтому для конусов теорема Дубовицкого–Милютина имеет следующую более простую формулировку.

Теорема (Дубовицкий и Милютин). *Пусть $\Omega_1, \ldots, \Omega_n$ непустые открытые выпуклые конусы в банаховом пространстве X, а $\Omega \subset X$ — непустой выпуклый конус. Тогда условие*

$$\Omega_1 \cap \ldots \cap \Omega_n \cap \Omega = \varnothing$$

равносильно существованию функционалов

$$x_1^* \in \Omega_1^*, \ldots, x_n^* \in \Omega_n^*, x^* \in \Omega^*,$$

таких, что
$$x_1^* + \ldots + x_n^* + x^* = 0$$

и при этом
$$\|x_1^*\| + \ldots + \|x_n^*\| > 0$$

(последнее означает, что набор $x_1^*, \ldots, x_n^*$ нетривиален).

6.1.4 Лемма о нетривиальности аннулятора.

Пусть $L \subset X$ — подпространство. Положим

$$L^\perp = \{x^* \in X^* \mid \langle x^*, x\rangle = 0 \quad \forall x \in L\}$$

и назовем $L^\perp$ *аннулятором* подпространства L. Ясно, что любой $x^* \in L^\perp$ является опорным к L, но верно и обратное (поясните!!). Таким образом, $L^\perp = L^*$. Очевидно, $L^\perp \subset X^*$ — подпространство.

Лемма. *Пусть $L \subset X$ — подпространство, не совпадающее с X. Тогда $L^\perp$ содержит ненулевой элемент.*

Доказательство. Пусть $x \in X \setminus L$. Тогда $B(x, \varepsilon) \subset X \setminus L$ при некотором $\varepsilon > 0$. Отделим $B(x, \varepsilon)$ от L ненулевым функционалом. Тогда $x^* \in L^\perp$, $x^* \neq 0$ (поясните почему!!). $\square$

(АТФ, стр. 127)

Глава 7

Лекция 7.

7.1 Тема: Аппарат теории экстремума (завершение)

7.1.1 Лемма о замкнутости образа.

Лемма. *Пусть X, Y, Z — банаховы пространства, $A : X \to Y$, $B : X \to Z$ — линейные операторы. Пусть $AX = Y$ и образ $B(Ker\, A)$ (ядра оператора A при отображении оператором B) замкнут в Z. Тогда образ TX пространства X при отображении оператором $T : X \to Y \times Z$ таким, что $Tx = (Ax, Bx)$ $\forall\, x \in X$ замкнут в произведении $Y \times Z$.*

Доказательство. Пусть последовательность $\{x_n\}$ в X такова, что

$$(Ax_n, Bx_n) \to (y, z) \in Y \times Z$$

сильно сходится, т. е.

$$\|Ax_n - y\| \to 0, \quad \|Bx_n - z\| \to 0 \qquad (n \to \infty).$$

Требуется доказать, что $\exists\, x \in X$ такой, что $Ax = y, Bx = z$. Тем самым лемма будет доказана.

Из условия $AX = Y$ следует, что $\exists\, x_y \in X$ такой, что $Ax_y = y$. Тогда

$$(A(x_n - x_y), B(x_n - x_y)) \to (0, z - Bx_y).$$

Поскольку $A(x_n - x_y) \to 0$ и $AX = Y$, то по лемме о правом обратном $\exists$ последовательность $\bar{x}_n$ такая, что

$$A(x_n - x_y) = A\bar{x}_n, \qquad \|\bar{x}_n\| \to 0.$$

Тогда

$$A(x_n - x_y - \bar{x}_n) = 0 \quad \forall\, n, \qquad \text{т. е. } (x_n - x_y - \bar{x}_n) \in Ker\, A \quad \forall\, n$$

$$B(x_n - x_y - \bar{x}_n) \to z - Bx_y \qquad (n \to \infty)$$

На ядре оператора A оператор B имеет замкнутый образ $\Rightarrow \exists\, x_0 \in Ker\, A$ такой, что

$$Bx_0 = z - Bx_y.$$

Тогда

$$B(x_0 + x_y) = z$$

$$A(x_0 + x_y) = Ax_y = y.$$

Лемма доказана. □

Ясно, что в этой лемме вместо условия $AX = Y$ можно предположить, что образ $Y_1 := AX$ замкнут в Y. Тогда $A : X \xrightarrow{HA} Y_1$ и, следовательно, образ T снова замкнут.

Обычно в оптимальном управлении используется следующее утверждение, вытекающее из леммы о замкнутости образа.

Следствие. *Пусть $A : X \to Y$ — линейный оператор «на», а $B : X \to Z$ — конечномерный линейный оператор ($dim\, Z < \infty$). Тогда оператор $T(x) = (Ax, Bx)$ имеет замкнутый образ в $Y \times Z$.*

Доказательство. Действительно, $dim\, B(Ker\, A) < \infty \Rightarrow B(Ker\, A)$ — замкнутое подпространство в Z. □

7.1.2 Лемма об аннуляторе ядра линейного оператора.

(АТФ, стр. 130)

Лемма. *Пусть $A : X \to Y$ — линейный оператор «на». Тогда*

$$(Ker\,A)^* = A^*Y^*,$$

где A^* — сопряженный оператор.

Доказательство. а) $A^*Y^* \subset (Ker\,A)^*$. Действительно, пусть $x^* = A^*y^*$, т. е.

$$\langle x^*, x\rangle = \langle y^*, Ax\rangle \qquad \forall\, x \in X.$$

Тогда

$$\langle x^*, x\rangle = \langle y^*, Ax\rangle = \langle y^*, 0\rangle = 0 \qquad \forall\, x \in Ker\,A,$$

т. е. $x^* \in Ker\,A^*$. Здесь условие $AX = Y$ не требуется.

б) Покажем, что $(Ker\,A)^* \subset A^*Y^*$. Пусть $x^* \in (Ker\,A)^*$, т. е.

$$\langle x^*, x\rangle = 0 \qquad \forall\, x \in Ker\,A.$$

Рассмотрим оператор

$$Tx = (Ax, \langle x^*, x\rangle),$$

отображающий X в произведение $Y \times \mathbb{R}$. Согласно следствию из леммы о замкнутости образа, образ TX замкнут в $Y \times \mathbb{R}$. При этом

$$TX \neq Y \times \mathbb{R}$$

поскольку, например, $(0, 1) \notin TX$, ибо если $Ax = 0$, то и $\langle x^*, x\rangle = 0$. Тогда по лемме о нетривиальности аннулятора на произведении $Y \times \mathbb{R}$ имеется ненулевой функционал равный нулю на образе TX, т. е. $\exists\, y^* \in Y^*$ и $c \in \mathbb{R}^1$ такие, что

$$\langle y^*, Ax\rangle + c\langle x^*, x\rangle = 0 \qquad \forall\, x \in X$$

и при этом

$$\|y^*\| + |c| > 0$$

Если $c = 0$, то $\langle y^*, Ax\rangle = 0\ \forall\, x \in X \Rightarrow \langle y^*, y\rangle = 0\ \forall\, y \in Y$ (ибо $AX = Y$) $\Rightarrow y^* = 0$.

Значит $c \neq 0$. Тогда

$$\langle x^*, x\rangle = -\langle \frac{1}{c} y^*, Ax\rangle \qquad \forall\, x \in X.$$

Таким образом,

$$x^* \in A^*Y^*.$$

Лемма доказана.

□

7.1.3 Производная по направлению. Производная Гато.

(АТФ, стр. 137)

Пусть X, Y — банаховы пространства, $U \subset X$ — открытое множество, $x_0 \in U$ — точка, $F : U \to Y$ — оператор, $h \in X$ — фиксированный элемент (направление).

Определение. Пусть существует предел

$$\lim_{\varepsilon \to +0} \frac{F(x_0 + \varepsilon h) - F(x_0)}{\varepsilon}.$$

Тогда он называется *производной* оператора F в точке x_0 *по напрвлению* h и обозначается $F'(x_0; h)$.

Определение. Пусть оператор F в точке x_0 имеет производную $F'(x_0, h)$ по любому направлению $h \in X$, и пусть отображение

$$h \in X \mapsto F'(x_0, h) \in Y$$

представляет собой линейный непрерывный оператор. Обозначим его $F'_\Gamma(x_0)$. Он переводит произвольный элемент $h \in Y$ в элемент

$$F'_\Gamma(x_0)h = F'(x_0, h) \in Y,$$

и называется *производной Гато* оператора F в точке x_0.

Если оператор *дифференцируем по Гато* в точке x_0, то есть имеет в этой точке производную Гато, то он имеет производную по любому направлению h, которая дается формулой

$$F'(x_0, h) = F'_\Gamma(x_0)h.$$

Далее, если оператор имеет в точке x_0 производную Фреше $F'(x_0)$, то он имеет и производную Гато $F'_\Gamma(x_0)$, и они совпадают

$$F'_\Gamma(x_0) = F'(x_0).$$

Обратное не имеет места, т. к., например, уже в $\mathbb{R}^2$ функция дифференцируемая по Гато в точке не обязана быть непрерывной в этой точке (см. АТФ стр. 141), в то время как из дифференцируемости по Фреше непрерывность, конечно, следует.

Итак, для точки x_0 мы имеем:

непрерывная дифференцируемость

$$\Downarrow$$

строгая дифференцируемость

$$\Downarrow$$

дифференцируемость по Фреше

$$\Downarrow$$

дифференцируемость по Гато

$$\Downarrow$$

дифференцируемость по каждому направлению.

Отметим также, что строгая дифференцируемость оператора g в точке x_0 (а значит, и непрерывня дифференцируемость в точке) влечет липшицевость в окрестности этой точки. Это вытекает, например, из представления

$$g(x) = g'(x_0)x + (g(x) - g'(x_0)x),$$

где $(g(x) - g'(x_0)x)$ имеет произвольную малую константу Липшица, если окрестность точки x_0 достаточно мала (см. Лекцию 5). Таким образом, константа Липшица оператора g в окрестности точки x_0 сколь угодно близка к величине $\|g'(x_0)\|$, если окрестность точки достаточно мала.

7.2 Тема: Условия минимума в абстрактных пространствах.

7.2.1 Минимизация функционала на множестве.

Пусть X — банахово пространство, $f : X \to \mathbb{R}$ — функционал, $M \subset X$ — множество, $x_0 \in X$ — точка. Ниже достаточно считать, что f определен в некоторой окрестности U точки x_0.

Определение. x_0 — точка *локального минимума* f *на* M, если $\exists\, \varepsilon > 0$ такое, что $f(x) \geqslant f(x_0)$ $\forall\, x \in M$ таких, что $\|x - x_0\| < \varepsilon$.

Нас будут интересовать необходимые условия локального минимума f на M. Чтобы их получить, нужно предположить определенные свойства f в окрестности точки x_0.

Теорема 1. *Пусть f имеет производную в точке $x_0 \in M$ по любому направлению $\bar{x} \in TM(x_0)$ и пусть f удовлетворяет условию Липшица в окрестности точки x_0. Пусть x_0 — точка локального минимума f на M. Тогда*

$$f'(x_0, \bar{x}) \geqslant 0 \qquad \forall\, \bar{x} \in TM(x_0).$$

Доказательство. Пусть выполнены условия теоремы и пусть $\bar{x} \in TM(x_0)$. Найдется $\tilde{x}(\varepsilon) : (0, \varepsilon_0) \to X$ $(\varepsilon_0 > 0)$ такая, что при $\varepsilon \in (0, \varepsilon_0)$

$$x_0 + \varepsilon\bar{x} + \tilde{x}(\varepsilon) \in M, \quad \|\tilde{x}(\varepsilon)\| = o(\varepsilon).$$

Из наличия локального минимума в x_0 на M вытекает, что

$$f(x_0 + \varepsilon\bar{x} + \tilde{x}(\varepsilon)) \geqslant f(x_0)$$

при всех достаточно малых $\varepsilon > 0$. Поскольку f удовлетворяет условию Липшица в окрестности точки x_0 с некоторой константой $L > 0$, то

$$f(x_0 + \varepsilon\bar{x} + \tilde{x}(\varepsilon)) \leqslant f(x_0 + \varepsilon\bar{x}) + L\|\tilde{x}(\varepsilon)\| = f(x_0 + \varepsilon\bar{x}) + o(\varepsilon).$$

Следовательно,

$$f(x_0) \leqslant f(x_0 + \varepsilon\bar{x}) + o(\varepsilon).$$

Отсюда вытекает, что

$$\lim_{\varepsilon \to +0} \frac{f(x_0 + \varepsilon\bar{x}) - f(x_0)}{\varepsilon} \geqslant 0,$$

т. е.

$$f'(x_0, \bar{x}) \geqslant 0.$$

□

Теорема 2. *Пусть f строго дифференцируема в точке $x_0 \in M$, и пусть $TM(x_0)$ — подпространство. Тогда необходимое условие локального минимума функции f на M в x_0 состоит в том, что*

$$f'(x_0)\bar{x} = 0 \qquad \forall\, \bar{x} \in TM(x_0).$$

Доказательство. Пусть $\bar{x} \in TM(x_0)$ — произвольный элемент. Поскольку $TM(x_0)$ — подпространство , то и $(-\bar{x}) \in TM(x_0)$. По теореме 1

$$f'(x_0)\bar{x} \geqslant 0 \quad \text{и} \quad f'(x_0)(-\bar{x}) \geqslant 0$$

Следовательно, $f'(x_0)\bar{x} = 0$. □

Пример. *В простейшей задаче вариационного исчисления*

$$\mathcal{J}(x) = \int_{t_0}^{t_1} F(t, x, \dot{x})\, dt \to \min,$$

$$x(t_0) = a, \quad x(t_1) = b,$$

рассматриваемой в пространстве C^1 (или $\mathbb{W}_1^1$), функционал $\mathcal{J}$ непрерывно дифференцируем, а множество M определено граничными условиями. Ясно, что $TM(x^0)$ в любой точке допустимой точке x^0 есть подпространство, определенное условиями:

$$\bar{x}(t_0) = \bar{x}(t_1) = 0.$$

Необходимое условие локального (т. е. слабого) минимума в точке x^0 состоит в том, что на этом подпространстве производная Фреше функционала (т. е. первая вариация)

$$\int_{t_0}^{t_1} F'_x\bar{x} + F'_{\dot{x}}\dot{\bar{x}}\, dt.$$

равна нулю (где F_x и $F_{\dot{x}}$ вычисляются вдоль траектории $(t, x^0(t), \dot{x}^0(t))$. Отсюда, как мы знаем, вытекает уравнение Эйлера.

Глава 8

Лекция 8

8.1 Тема: Условия минимума в абстрактных пространствах (продолжение)

8.1.1 Задачи с ограничением типа равенства.

Пусть X, Y - банаховы пространства, $U \subset X$ - открытое множество, $f : U \to \mathbb{R}$ - функционал, $g : U \to Y$ - оператор. Рассмотрим задачу:

$$f(x) \to min, \quad g(x) = 0, \quad x \in U. \tag{8.1}$$

Пусть x_0 — *допустимая точка*, т.е. $x_0 \in U,\ g(x_0) = 0$. Предположим, что f и g строго дифференцируемы в точке x_0 (например, достаточно считать, что f и g непрерывно дифференцируемы на U, т.е. $f \in C^1(U)$ $g \in C^1(U, Y)$).

Предположим также, что

$$g'(x_0)X = Y$$

— выполнено условие Люстерника, и следовательно, по теореме Люстерника касательный конус к множеству

$$M = \{x \in U \mid g(x) = 0\}$$

в точке x_0 есть подпространство

$$Ker\ g'(x_0).$$

Согласно теореме 2 прошлой лекции, необходимое условие локального минимума в точке x_0 состоит в том, что

$$\langle f'(x_0), \bar{x}\rangle = 0 \quad \forall \bar{x} \in Ker\ g'(x_0). \tag{8.2}$$

Пусть это условие выполнено. Согласно лемме об аннуляторе ядра линейного сюрьективного оператора существует $y^* \in Y^*$ такой, что

$$f'(x_0) + y^* \circ g'(x_0) = 0, \tag{8.3}$$

где $y^* \circ g'(x_0) = [g'(x_0)]^* y^*$ (т.е. $\langle y^* \circ g'(x_0), x\rangle = \langle y^*, g'(x_0)x\rangle \forall x \in X$). Итак, необходимое условие (8.2) мы представили в *двойственной форме* (8.3). Последняя носит название «правило множителей Лагранжа». Здесь единственным множителем Лагранжа является линейный функционал $y^* \in Y^*$. Это условие в столь общей бесконечномерной задаче с ограничением типа равенства впервые было получено Л.А.Люстерником в 1934 году.

Некоторое «неудобство» полученного условия состоит в том, что требуется выполнение условия Люстерника. Формально этого требования можно избежать в тех случаях, когда известно, что образ $g'(x_0)X$ замкнут. Последнее, как мы узнаем позже, верно для операторов ограничений равенства в задачах оптимального управления. Это верно и в случае, когда $\dim Y < +\infty$.

В предположении о замкнутости образа $g'(x_0)X$, формально более слабом, чем условие $g'(x_0)X = Y$, правило множителей Лагранжа, т.е. необходимое условие локального минимума, формулируют в следующем виде:

существуют $\alpha \in \mathbb{R}$ и $y^* \in Y^*$, не равные нулю одновременно, и такие, что

$$\alpha f'(x_0) + y^* \circ g'(x_0) = 0. \tag{8.4}$$

Если условие Люстерника выполнено, то (8.4) реализуется с $\alpha = 1$, т.е. в виде (8.3).

Если же условие Люстерника не выполнено, то по предположению образ $g'(x_0)X$ есть замкнутое подпространство в Y. Тогда по лемме о нетривиальности аннулятора существует $y^* \neq 0$ такой, что $\langle y^*, g'(x_0)X\rangle = 0$. Условие (8.4) реализовалось на этот раз с $\alpha = 0$. Ничего другого, кроме как констатации факта, что $g'(x_0)X \neq Y$, это условие в данном случае в себе не несет. По сути, при невыполненом условии Люстерника мы отказываемся от исследования задачи. Однако, подобная форма записи необходимого условия оказывается удобной.

Отметим также, что, не ограничивая общности, мы можем считать в необходимом условии, что $\alpha \geqslant 0$, ибо условие выдерживает умножение на (-1).

Введем теперь *функцию Лагранжа:*

$$L(x, \alpha, y^*) = \alpha f(x) + \langle y^*, g(x)\rangle.$$

Тогда условие (8.4) можно записать так

$$L'_x(x_0, \alpha, y^*) = 0,$$

где L'_x — частная производная Фреше (определяемая естественным образом).

Итак, доказана следующая теорема:

Теорема. *(Л.А.Люстерник) Пусть f и g строго дифференцируемы в точке $x_0 \in U$ такой, что $g(x_0) = 0$, и пусть образ $g'(x_0)X$ замкнут в Y. Пусть x_0 — точка локального минимума в задаче* (8.1)*. Тогда существуют множители Лагранжа $\alpha \in \mathbb{R}$ и $y^* \in Y^*$ такие, что*

$$\alpha \geqslant 0, \tag{8.5}$$

$$\alpha + \|y^*\| > 0, \tag{8.6}$$

$$L'_x(x_0, \alpha, y^*) = 0, \tag{8.7}$$

где $L(x_0, \alpha, y^*) = \alpha f(x) + \langle y^*, g(x)\rangle$.

Здесь
(8.5) — условие неотрицательности множителя α
(8.6) — условие нетривиальности набора (α, y^*)
(8.7) — условие стационарности по x функции Лагранжа.

Эта теорема реализует, так называемый, *принцип Лагранжа*, состоящий в том, что если x_0 — точка минимума в задаче с ограничениями, то она же стационарная точки функции Лагранжа (но не обязательно точка минимума последней!) в задаче без ограничений.

Задание. Сформулируйте теорему 1 в случае, когда $Y = \mathbb{R}^m$ и, следовательно, $g = (g_1, \dots g_m)$ - конечный набор функционалов над X. В этом случае образ $g'(x)X$ автоматически замкнут $\forall x \in U$.

8.1.2 Условия максимума в гладкой задаче математического программирования.

Пусть X — банахово пространство, $U \subset X$ - открытое множество, на котором заданы функционалы $f_i : U \to \mathbb{R}, \;\; i = 0, \dots, k$, и оператор $g : U \to Y$, действующий в другое банахово пространство Y. Предположим, что все f_i и g непрерывно диференцируемы на U, причем, образ $g'(x)X$ замкнут в Y $\forall x \in U$. Рассмотрим задачу (*):

$$f_0(x) \to min, \quad f_i(x) \leqslant 0, \quad i = 1, \dots, k, \quad g(x) = 0, \quad x \in U.$$

Назовем ее *гладкой задачей с ограничениями типа равенства и типа неравенства*, или *гладкой задачей математического программирования.*

Пусть $\alpha = (\alpha_0, \dots, \alpha_k) \in \mathbb{R}^{k+1}, \;\; y^* \in Y^*$. Введем функцию Лагранжа задачи (*) :

$$L(x_0, \alpha, y^*) = \sum_{i=0}^{k} \alpha_i f_i(x) + \langle y^*, g(x)\rangle.$$

Назовем $\alpha_0, \; \dots \;, \alpha_k, \; y^*$ *множителями Лагранжа*, а через $\lambda = (\alpha_0, \dots, \alpha_k, y^*) = (\alpha, y^*)$ будем обозначать произвольный набор множителей Лагранжа. Таким образом,

$$L = L(x, \lambda).$$

Теорема. *(правило множителей Лагранжа) Пусть x_0 — точка локального минимума в задаче (*). Тогда существуют набор*

$$\lambda = (\alpha_0, \ldots, \alpha_k, y^*)$$

такой, что выполнены следующие условия

$$\alpha_0 \geqslant 0, \ \alpha_1 \geqslant 0, \ldots, \alpha_k \geqslant 0, \tag{8.1}$$

$$\sum_0^k \alpha_i + \|y^*\| > 0, \tag{8.2}$$

$$\alpha_i f_i(x_0) = 0, \quad i = 1, \ldots, k, \tag{8.3}$$

$$L'_x(x_0, \lambda) = 0. \tag{8.4}$$

Говорят, что
(8.1) — условия *неотрицательности*,
(8.2) — условие *нетривиальности*,
(8.3) — условия *дополняющей нежесткости*,
(8.4) — условие *стационарности по x* функции Лагранжа.
Для доказательства теоремы нам понадобится

Лемма. *Пусть x_0 — точка локального минимума в задаче (*). Пусть $g'(x_0)X = Y$, т.е. для g в x_0 выполнено условие Люстерника. Тогда не существует $\bar{x} \in X$, удовлетворяющего условиям:*

$$\langle f'_0(x_0), \bar{x} \rangle < 0; \quad \langle f'_i(x_0), \bar{x} \rangle < 0, \quad i \in I, \tag{8.5}$$

$$g'(x_0)\bar{x} = 0, \tag{8.6}$$

где,

$$I = \{i \in \{1, \ldots, k\} \mid f_i(x_0) = 0\} \tag{8.7}$$

— множество активных индексов ограничений типа неравенства в точке x_0.

Доказательство. Пусть $\bar{x}$ удовлетворяет условиям (8.5) и (8.6). Согласно теореме Люстерника вектор $\bar{x}$ (в силу (8.6)) является касательным к ограничению $g(x) = 0$ в точке x_0. Следовательно, существует $\tilde{x}(\varepsilon) : (0, \varepsilon_0) \to X (\varepsilon_0 > 0)$ такая, что

$$g(x_0 + \varepsilon\bar{x} + \tilde{x}(\varepsilon)) = 0, \quad \|\tilde{x}(\varepsilon)\| = o(\varepsilon). \tag{8.8}$$

Положим $x_\varepsilon = x_0 + \varepsilon\bar{x} + \tilde{x}(\varepsilon)$. Покажем, что $\{x_\varepsilon\}$ «нарушает» минимум в точке x_0. Действительно,

$$g(x_\varepsilon) = 0, \quad \varepsilon \in (0, \varepsilon_0), \tag{8.9}$$

$$\|x_\varepsilon - x_0\| \to 0 \ \ (\varepsilon \to +0). \tag{8.10}$$

Пусть i не принадлежит I, $i \in \{1, \ldots, k\}$, т.е. $f_i(x_0) < 0$. Тогда из непрерывности f_i в точке x_0 в силу (8.10) следует, что

$$f_i(x_\varepsilon) < 0$$

для всех достаточно малых $\varepsilon > 0$.

Пусть теперь $i \in I$. Тогда $f_i(x_0) = 0$ и $\langle f'_i(x_0), \bar{x} \rangle < 0$ в силу (8.5). Следовательно,

$$f_i(x_\varepsilon) = f_i(x_0 + \varepsilon\bar{x} + \tilde{x}(\varepsilon)) = f_i(x_0) + \langle f'_i(x_0), \varepsilon\bar{x} + \tilde{x}(\varepsilon) \rangle + o(\varepsilon) = \varepsilon\langle f'_i(x_0), \bar{x} \rangle + o_1(\varepsilon).$$

Поскольку $\varepsilon > 0$, $\langle f'_i(x_0), \bar{x} \rangle < 0$, то $f_i(x_\varepsilon) < 0$ для всех достаточно малых $\varepsilon > 0$.

Аналогично устанавливается, что $f_0(x_\varepsilon) < f_0(x_0)$ для всех достаточно малых $\varepsilon > 0$. Но тогда $\{x_\varepsilon\}$ удовлетворяет всем ограничениям и «нарушает» локальный минимум. Лемма доказана. □

Нам понадобится еще одно

Предложение. *Пусть $l : X \to \mathbb{R}$ — ненулевой линейный функционал, т.е. $l \in X^*, l \neq 0$. Пусть $K = \{x \in X \mid \langle l, x \rangle > 0\}$ — открытое полупространство. Пусть $m \in K^*$, т.е. $\langle m, x \rangle \geqslant 0 \ \ \forall x \in K$. Тогда существует $\lambda \geqslant 0$ такое, что $m = \lambda l$.*

Доказательство. Покажем сначала, что

$$\langle m, x\rangle = 0 \quad \forall x \in Ker\ l.$$

Действительно, пусть $x_0 \in Ker\ l$, т.е. $\langle l, x_0\rangle = 0$ и пусть x_1 таков, что $\langle l, x_1\rangle > 0$. Тогда

$$x_1 + \beta x_0 \in K \quad \forall \beta \in \mathbb{R}.$$

Следовательно,

$$\langle m, x_1\rangle + \beta\langle m, x_0\rangle \geqslant 0 \ \ \forall \beta \in \mathbb{R}$$

Это возможно лишь при $\langle m, x_0\rangle = 0$. Итак,

$$\langle l, x\rangle = 0 \ \Rightarrow \ \langle m, x\rangle = 0.$$

Отсюда, как мы знаем, следует существование $\lambda \in \mathbb{R}$ такого, что $m = \lambda l$. Поскольку

$$\langle l, x\rangle > 0 \ \Rightarrow \ \langle m, x\rangle \geqslant 0,$$

то $\lambda \geqslant 0$. □

Доказательство. (доказательство теоремы) Пусть x_0 — точка локального минимума в задаче (*). Рассмотрим два случая.

а) *Невырожденный случай*: $g'(x_0)X = Y$. В этом случае согласно Лемме система условий (8.5) и (8.6) несовместна. Если среди функционалов $f_i(x_0), i \in I \cup \{0\}$ есть нулевой, то полагая соответствующий ему множитель α_i равным 1, а остальные множители (включая y^*) равными нулю, получае набор λ удовлетворяющий всем условиям (8.1)-(8.4) теоремы. Поэтому считаем, что

$$f_i(x_0) \neq 0, \forall i \in I \cup \{0\}.$$

Тогда все полупространства (8.5) непусты. Положим

$$\Omega_i = \{\bar{x} \mid \langle f_i'(x_0), \bar{x}\rangle < 0\}, \quad i \in I \cup \{0\},$$

$$\Omega = \{\bar{x} \mid g'(x_0) = 0\}.$$

Пересечение полупространств Ω_i, $i \in I \cup \{0\}$ и полупространства Ω пусто. При этом, согласно предложению

$$\Omega_i^* = \Big\{x^* \in X^* \mid \langle x^*, x\rangle = -\alpha_i\langle f_i'(x_0), x\rangle \quad \forall x \in X, \text{ где } \alpha_i \geqslant 0\Big\},$$

т.е. сопряженный конус к полупространству Ω_i есть луч (с вершиной в нуле), натянутый на $(-f_i'(x_0))$.

Согласно теореме Дубовицкого-Милютина существуют

$$x_i^* \in \Omega_i^*, \quad i \in I \cup \{0\}, \quad x^* \in \Omega^*$$

не все равные нулю и такие, что

$$\sum_{I \cup \{0\}} x_i^* + x^* = 0$$

Согласно лемме об ануляторе $x^* = [g'(x_0)]^*(-y^*) = -y^* \circ g'(x_0)$ (знак «-» удобен), где $y^* \in Y^*$, и как мы сказали выше

$$x_i^* = -\alpha_i f_i'(x_0), \quad \alpha_i \geqslant 0, \quad i \in I \cup \{0\}.$$

Следовательно,

$$-\sum_{I \cup \{0\}} \alpha_i f_i'(x_0) - y^* \circ g'(x_0) = 0.$$

Остается положить $\alpha_i = 0$ при i не принадлежащем I, $i \neq 0$. Тогда все условия (8.1)-(8.4) теоремы оказываются выполненными. (Набор α_i, y^* нетривиален, ибо в противном случае набор x_i^*, x^* тривиален.)

б) *Вырожденный случай*: $g'(x_0)X \neq Y$. В этом случае образ $L = g'(x_0)X$ есть замкнутое пространство в Y, не совпадаещее с Y. Тогда по лемме о нетривиальности аннулятора существует функционал $y^* \neq 0$, аннулирующийся на L. Следовательно, $\langle y^* g'(x_0), x\rangle = 0, \quad \forall x \in X$. Положим $\alpha_i = 0, \quad \forall i = 0, \ldots, k$. Тогда набор $\lambda = (0, \ldots, 0, y^*)$ удовлетворяет всем условиям (8.1)-(8.4) теоремы. (В этом случае правило множителей Лагранжа лишь констатирует тот факт, что $g'(x_0)X \neq Y$, и мы отказываемся от исследования задачи на минимум.) □

Замечание. *Из доказательства видно, что условия теоремы могут быть ослаблены. А именно, достаточно считать, что в точке x_0 все функционалы f_i дифференцируемы по Фреше, а оператор g строго дифференцируем, и, кроме того, образ $g'(x_0)X$ замкнут в Y.*

Глава 9

Лекция 9.

9.1 Тема: Оптимальное управление. Принцип максимума.

9.1.1 Пример задачи оптимального управления (А.А.Милютин)

Точка движется по плоскости. Скорость точки может иметь любое направление, но ограничена по величене. Какую скорость надо сообщить точке в каждый момент времени, чтобы за кратчайшее время точка перешла бы из положения А в положение В?

Формализуем задачу. Положение точки характеризуется ее координатами (x, y). Скорость точки в данный момент обозначим через $(u, v)(t)$. Если $u = u(t)$ и $v = v(t)$ заданы, то изменение состояния точки подчинено системе дифференциальных уравнений

$$\begin{cases} \dot{x} = u(t) \\ \dot{y} = v(t) \end{cases}$$

При этом $u^2 + v^2 \leqslant const$. Здесь t — независимая переменная — время, $t \in [t_0, t_1]$. Отрезок $[t_0, t_1]$ не фиксирован. Требуется минимизировать $t_1 - t_0$. Итак, задача имеет вид

$$t_1 - t_0 \to min$$

$$\begin{cases} \dot{x} = u(t) \\ \dot{y} = v(t) \end{cases}$$

$$u^2 + v^2 \leqslant const$$

$$x(t_0) = a_1 \qquad x(t_1) = b_1$$

$$y(t_0) = a_2 \qquad y(t_1) = b_2$$

где (a_1, a_2) и (b_1, b_2) — координаты А и В соответствено.

Приведенная задача имеет простое решение, и для ее решения не нужно знать теорию оптимального управления. Однако, если ограничение на управление

$$u^2 + v^2 \leqslant const$$

заменить на

$$\frac{u^2}{a^2} + \frac{v^2}{b^2} \leqslant 1,$$

то решение становится неочевидным. Наша цель - дать единый алгоритм, с помощью которого можно решать (но не обязательно решить!!) любую задачу оптимального управления.

9.1.2 Управляемая система.

Управляемой системой будем называть пару условий

$$\dot{x} = f(t, x, u), \quad u \in U. \tag{9.1}$$

Здесь $U \subset \mathbb{R}^n$ — произвольное множество, которому должно принадлежать *управление* $u = u(t)$ в каждый момент времени t. Таким образом, U задает *ограничение на управление.* Компоненты вектора $u = (u_1, \ldots, u_m)$ также называют *управлениями* или *управляющими параметрами.* Вектор $x = (x_1, \ldots, x_n)$ характеризует *состояние* управляемого объекта. В механических системах обычно x_i - это обобщенные координаты. В оптимальном управлении $x = x(t)$ называют *фазовой переменной*, или *фазой.* Уравнение $\dot{x} = f(t, x, u)$ описывает «динамику» объекта управления при выбранном управлении $u = u(t)$.

Если задано начальное состояние в момент t_0 $x(t_0) = a$ и задано управление $u(t)$ на отрезке $[t_0, t_1]$, то (в принципе) $x(t)$ можно определить как решение задачи Коши

$$\dot{x} = f(t, x, u(t)), \quad x(t_0) = a.$$

(Однако мы не будем интересоваться вопросом о существовании решения этой задачи.)

Относительно функции f мы будем предполагать, что *она сама и ее частные производные* f_x *и* f_t *непрерывны* по совокупности переменных.

Пара функций $x(t), u(t)$, определенных на отрезке $[t_0, t_1]$, называется *траекторией управляемой системы* (1), если

$$x(t) : [t_0, t_1] \to \mathbb{R}^n$$

— абсолютно непрерывная функция,

$$u(t) : [t_0, t_1] \to \mathbb{R}^m$$

— ограниченная измеримая функция, и при этом почти всюду (п.в.) на $[t_0, t_1]$ выполнены условия

$$\dot{x}(t) = f(t, x(t), u(t)), \quad u(t) \in U.$$

Заметим, что данное определение позволяет менять $u(t)$ на множестве меры нуль (называя функцию $u(t)$ ограниченной измеримой, мы имели в виду измеримость и существенную ограниченность).

С траекторией $(x(t), u(t) \mid t \in [t_0, t_1])$ мы будем связывать набор концевых значений $p = (t_0, x_0, t_1, x_1)$, где $x_0 = x(t_0),\ x_1 = x(t_1)$. Мы будем также полагать для краткости

$$w = (x, u).$$

Часто вводят открытое множество $Q \subset \mathbb{R}^{1+n+m}$ и предполагают, что f задана и непрерывна вместо с поизводными f_x и f_t на Q. В определении управляемой системы добавляют условие

$$(t, x, u) \in Q.$$

Предполагают, что это условие выполнено для траектории $x(t), u(t)$ п.в. на $[t_0, t_1]$. Более того предполагают, что для каждой траектории существует компакт $\mathcal{C} \subset Q$ (зависящий от траектории) такой, что

$$(t, x(t), u(t)) \in \mathcal{C}$$

п.в. на $[t_0, t_1]$. Это означает, что траектория проходит в Q на положительном расстоянии от границы ∂Q. Такое требование предполагается выполненным для любой траектории управляемой системы. Все дальнейшее изложение можно проводить с таким дополнительным открытым ограничением в виде Q. Однако, для простоты мы полагаем ниже

$$Q = \mathbb{R}^{1+m+n}.$$

9.1.3 Каноническая задача оптимального управления - задача А.

Эта задача ставится следующим образом: минимизировать функцию начального и конечного сосотояний, а также начального и конечного моментов времени

$$J = F_0(t_0, x_0, t_1, x_1) \to min,$$

где по определению

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

при ограничениях на t_0, x_0, t_1, x_1 заданных в виде неравенств и равенств

$$F(t_0, x_0, t_1, x_1) \leqslant 0, \quad K(t_0, x_0, t_1, x_1) = 0,$$

которым должны удовлетворять траектории

$$(x(t), u(t) \mid t \in [t_0, t_1])$$

управляемой системы

$$\dot{x} = f(t, x, u), \quad u \in U.$$

Отрезок $[t_0, t_1]$ *á priori* не фиксирован.

Полагая для краткости

$$p = (t_0, x_0, t_1, x_1),$$

перепишем задачу в виде:

$$J = F_0(p) \to min, \quad F(p) \leqslant 0, \quad K(p) = 0,$$
$$\dot{x} = f(t, x, u), \quad u \in U.$$

Здесь

$$F_0 : \mathbb{R}^{2+2n} \to \mathbb{R}$$
$$F : \mathbb{R}^{2+2n} \to \mathbb{R}^k, \quad F = (F_1, \dots, F_k)$$
$$K : \mathbb{R}^{2+2n} \to \mathbb{R}^s, \quad K = (K_1, \dots, K_s)$$
$$f : \mathbb{R}^{1+n+m} \to \mathbb{R}^n, \quad f = (f_1, \dots, f_n)$$

Предложение. *Предполагается, что F_0, F, K — функции класса $\mathcal{C}^1$, и, как было сказано, f непрерывна вместе с производными f_x и f_t. Никаких предположений относительно множества $U \subset \mathbb{R}^n$ не делается.*

Мы уже отметили, что свойства f можно было бы предполагать выполненными на некотором открытом множестве Q, соответственно уточнив понятие допустимой траектории. Аналогично можно было бы считать, что F_0, F, K определены и непрерывно дифференцируемы на некотором открытом множестве $P \subset \mathbb{R}^{2+2n}$. Мы этого не делаем, чтобы не отвлекаться на несущественные детали.

Если $F_0(p) = t_1 - t_0$, то есть требуется минимизировать время, то задача называется задачей *быстродействия*, при этом часто t_0 фиксируют, полагая $t_0 = 0$ (в этом случае $K_i(p) = t_0$ при некотором i). Может оказаться, что в задаче фиксирован левый конец $x(t_0) = a$ или правый $x(t_0) = a$, или оба. Такие условия задаются с помощью функции K: $K_i = x_0 - a = 0$, или $K_j = x_1 - b = 0$ при некоторых i, j. Бывают постановки, когда из точки $x(t_0) = a$ надо попасть на многообразие $\varphi(x(t_1)) = 0$ или наоборот, или с многообразия на многообразие. Все эти постановки охватываются равенством $K(p) = 0$.

С помощью равенства $K = 0$ можно фиксировать как t_0, так и t_1, или t_0 и t_1, одновременно. В последнем случае имеем задачу на фиксированном отрезке времени.

Наконец, пусть вместо концевого функционала в задаче присутствует интегральный функционал

$$J_\Phi = \int_{t_0}^{t_1} \Phi(t, x, u) dt,$$

где Φ непрерывна вместе со своими производными Φ_t и Φ_x. Тогда его можно переписать как концевой с помощью введения дополнительной фазовой переменой y:

$$\dot{y} = \Phi(t, x, u) \Rightarrow J_\Phi = y_1 - y_0,$$

где $y_1 = y(t_1),\ y_0 = y(t_0)$. Заметим, что при этом меняется и управляемая система, приобретая вид:

$$\dot{x} = f(t, x, u), \quad \dot{y} = \Phi(t, x, u), \quad u \in U.$$

Указаный прием существенно расширяет круг задач, охватываемый канонической задачей.

Траектория $(x(t), u(t) \mid t \in [t_0, t_1])$ называется *допустимой в канонической задаче A*, если $x(t)$ абсолютно непрерывна на $[t_0, t_1]$, $u(t)$ ограничена и измерима на $[t_0, t_1]$ и выполнены все ограничения задачи.

Наша цель сейчас - исследовать каноническую задачу А и получить в ней необходимое условие минимума - принцип максимума Понтрягина. Последний является обобщением условия Вейерштрасса на задачи оптимального управления.

Главное, что отличает задачи оптимального управления от задач вариационного исчисления - это присутствие ограничения $u \in U$. Как уже подчеркивалось, в качестве U может выступать *любое* множество, например двухточечное.

Поэтому нужны новые специальные приемы для исследования задач оптимального управления. Таких приемов имеется несколько, и они приводят к различным доказательствам принципа максимума. Мы будем следовать приему и доказательству, рассказанному А.А.Милютиным на одном из семинаров по оптимальному управлению. Это доказательство представляется нам одним из наиболее простых.

9.1.4 Каноническая задача B.

Оказывается, вовсе не обязательно рассматривать задачу A в полной общности, ибо имеется прием сведения ее к более простой задаче - задаче B. Последняя имеет вид :

$$J = F_0(x_0, x_1) \to min, \quad F(x_0, x_1) \leqslant 0, \quad K(x_0, x_1) = 0,$$

$$\dot{x} = f(x, u), \quad u \in U,$$

где по определению

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

Фунции F_0, F, K непрерывно дифференцируемы в $\mathbb{R}^{2n}$, f непрерывна вместе с производной f_u в $\mathbb{R}^{n+m}$.

Отрезок $[t_0, t_1]$ по-прежнему *не фиксирован.*

Теперь f не зависит от t явно, т.е. система $\dot{x} = f$ является *автономной.* Кроме того, F_0, F, K не зависят явно от t_0, t_1.

В задаче B траектории допускают сдвиг по времени с сохранением концов x_0, x_1. Действительно, если $(x(t), u(t) \mid t \in [t_0, t_1])$ — допустимая траектория задачи B, то при любом вещественном θ траектория

$$(x(t-\theta), u(t-\theta) \mid t \in [t_0 + \theta, t_1 + \theta])$$

— допустимая траектория той же задачи с теми же концами x_0, x_1. Это касается и оптимальных траекторий задачи B.

Переход от задачи A к задаче B достигается с помощью следующего приема. К системе $\dot{x} = f(t, x, u)$ добавим уравнение

$$\frac{dt}{d\tau} = 1,$$

считая, что $t = t(\tau)$ - новая фазовая переменая. Функции $x(\cdot)$ и $u(\cdot)$ также считаем зависящими от τ: $x = x(\tau), u = u(\tau)$. Положим

$$t_0 = t(\tau_0), \quad x_0 = x(\tau_0), \quad t_1 = t(\tau_1), \quad x_1 = x(\tau_1). \tag{9.2}$$

Приходим к следующей задаче

9.1.5 Задача A'.

$$J = F_0(t_0, x_0, t_1, x_1) \to min, \quad F(t_0, x_0, t_1, x_1) \leqslant 0, \quad K(t_0, x_0, t_1, x_1) = 0,$$

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

где $t(\tau), x(\tau)$ — фазовые переменные, $u(t)$ — управление, $\tau \in [\tau_0, \tau_1]$, $\tau_0 < \tau_1$ - не фиксированы, концы t_0, x_0, t_1, x_1 определены условиями (9.2).

Ясно, что задача A - того же типа, что и задача B, и следовательно в ней есть сдвиг траекторий по времени.

Какова связь между задачами A и A'?

Пусть $(x(t), u(t) \mid t \in [t_0, t_1])$ - допустимая траектория задачи A. Ее концы - это набор (t_0, x_0, t_1, x_1), где $x_0 = x(t_0), \quad x_1 = x(t_1)$.

Положим

$$\tau_0 = t_0, \quad \tau_1 = t_1, \quad t(\tau) = \tau.$$

Тогда

$$(t(\tau), x(\tau), u(\tau) \mid \tau \in [\tau_0, \tau_1])$$

— допустимая траектория задачт A' с теми же концами (t_0, x_0, t_1, x_1), определенными условиями (9.1).

Обратно, пусть имеется допустимая траектория

$$(t(\tau), x(\tau), u(\tau) \mid \tau \in [\tau_0, \tau_1])$$

в задаче A' с концами (t_0, x_0, t_1, x_1). Тогда

$$t(\tau) = \tau + const.$$

Сделав, если надо, сдвиг по времени, константу можно занулить, т.е. перейти к новой допустимой траектории в задаче A' с теми же концами, но с функцией $t(\tau) = \tau$. Для этой траектории положим $t_0 = \tau_0$, $t_1 = \tau_1$. Тогда

$$(x(t), u(t) \mid t \in [t_0, t_1])$$

— допустимой траектории задачи A с теми же концами (t_0, x_0, t_1, x_1), где $x_0 = x(t_0)$, $\quad x_1 = x(t_1)$.

Указаное соответствие между допустимыми траекториями задач A и A' имеется и для оптимальных траекторий. Поэтому, получив необходимое условие минимума в задаче B, мы легко перепишем его и как необходимое условие в задаче А - с помощью указаного приема.

Переписка задач, замена переменых играют в оптимальном управлении важную роль. Уже столь простой прием замены независимой переменной t принес заметную пользу. Более серьезная замена времени $t = t(\tau)$ позволит свести задачу к конечномерной задаче с неравенствами и равенствами, а затем воспользоваться правилом множителей Лагранжа. В конечном итоге это приведет к доказательству принципа максимума. Замена будет связана с уравнением $\frac{dt}{d\tau} = v$, где v неотрицательная функция и, следовательно, $t = t(\tau)$ - монотонная функция. Подобная замена, превращающая независимую переменную t в фазовую переменную, была предложена А.Я.Дубовицким и А.А.Милютиным, и была названа ими v-заменой.

9.1.6 v-замена.

Рассмотрим управляемую систему задачи B:

$$\dot{x} = f(x, u), \quad u \in U, \tag{9.3}$$

траекторией которой является пара

$$(x(t), u(t) \mid t \in [t_0, t_1]),$$

где $t_0 < t_1$ — произвольны; $x(t)$ абсолютно непрерывна, а $u(t)$ ограничена и измерима, причем условия (9.3) на траектории выполнены п.в.

Определим теперь *расширенную управляемую систему* задачи B:

$$\frac{dx}{d\tau} = vf(x, u), \quad \frac{dt}{d\tau} = v, \quad u \in U, \quad v \geqslant 0. \tag{9.4}$$

Здесь τ — независимая переменная. Траекторией системы (9.4) по определению является набор функций

$$(t(\tau), x(\tau), u(\tau), v(\tau) \mid \tau \in [\tau_0, \tau_1])$$

таких, что

$t(\tau), x(\tau)$ — абсолютно непрерывны,

$u(\tau)$ — ограничена и измерима,

$v(\tau)$ — кусочно постоянна, (!!)

(т.е. на $[\tau_0, \tau_1]$ имеется конечное число интервалов, примыкающих друг к другу, на которых $v(\tau)$ постоянна), условия (9.4) выполнены почти всюду на $[\tau_0, \tau_1]$, $\tau_0 < \tau_1$ - произвольны.

Концы траектории системы (9.3) - это t_0, t_1 и значения

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

Концы траектории системы (9.4) — это τ_0, τ_1 и значения

$$x_0 = x(\tau_0), \quad x_1 = x(\tau_1),$$

$$t_0 = t(\tau_0), \quad t_1 = t(\tau_1).$$

Лемма 1. *Для любой траектории системы (9.3) существует траектория системы (9.4) с теми же x_0, x_1.*

Доказательство. Пусть

$$(x(t), u(t) \mid t \in [t_0, t_1])$$

— траектория системы (9.3). Положим

$$\tau_0 = t_0, \quad \tau_1 = t_1, \quad v(\tau) = 1, \quad t(\tau) = \tau.$$

Тогда

$$(t(\tau), x(\tau), u(\tau), v(\tau) \mid \tau \in [\tau_0, \tau_1])$$

траектория системы (9.4) с теми же x_0, x_1. $\square$

Эта лемма элементарна, но в действительности она нам и не понадобится. Понадобится следующая, менее элементарная

Лемма 2. *Для любой траектории системы (9.4) существует траектория системы (9.3) с теми же* x_0, x_1.

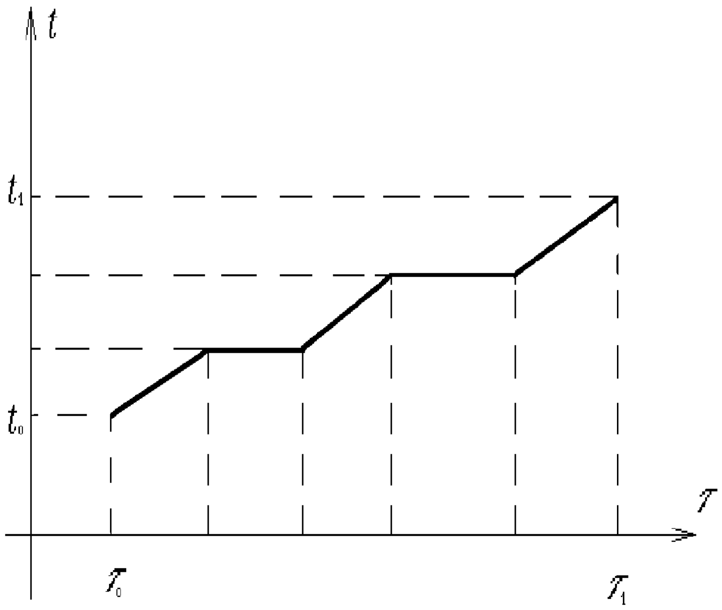
Для доказательства нам потребуются два предложения.

Пусть функция $\tilde{v}(\tau) : [\tau_0, \tau_1] \to \mathbb{R}$ — кусочно постоянна и неотрицательна, и пусть $\tilde{t}(\tau)$ удовлетворяет условиям

$$\frac{d\tilde{t}(\tau)}{\tau} = \tilde{v}(\tau), \quad \tilde{t}(\tau_0) = t_0,$$

то есть

$$\tilde{t}(\tau) = t_0 + \int_{\tau_0}^{\tau} \tilde{v}(s) ds.$$



Таким образом,

$$\tilde{t}(\tau) : [\tau_0, \tau_1] \to [t_0, t_1]$$

«отображение на» — кусочно-линейная неубывающая функция. Пусть

$$\Delta^1, \Delta^2, \ldots, \Delta^s$$

— набор всех отрезков постоянства функции $\tilde{t}(\tau)$, расположенных в естественном порядке. Положим

$$\mathcal{M}_0 = \bigcup_1^s \Delta^k, \quad \mathcal{M}_+ = [\tau_0, \tau_1] \backslash \mathcal{M}_0.$$

$\mathcal{M}_+$ состоит из конечного числа интервалов (или полуинтервалов, когда τ_0 или τ_1 служит одним из концов) линейного роста функции $\tilde{t}(\tau)$.

Пусть

$$\tilde{t}(\Delta^k) = \{t^k\}, \quad k = 1, \ldots, s$$

— значения функции $\tilde{t}$ на интервалах постоянства. Положим

$$\mathcal{N}_0 = \{t^1, \ldots, t^s\}, \quad \mathcal{N}_+ = [t_0, t_1] \backslash \mathcal{N}_0.$$

Ясно, что

$$\mathcal{N}_0 = \tilde{t}(\mathcal{M}_0), \quad \mathcal{N}_+ = \tilde{t}(\mathcal{M}_+).$$

Далее, пусть
$$\tilde{\tau}(t) : [t_0, t_1] \to [\tau_0, \tau_1],$$
такова, что
$$\tilde{t}(\tilde{\tau}(t)) = t \quad \forall t \in [t_0, t_1],$$
т. е. $\tilde{\tau}$ является правой обратной к $\tilde{t}(\tau)$. Поскольку $\tilde{t}(\tau)$ отображают «на», то правая обратная существует. Более того, она определена однозначно всюду, кроме конечного множества точек $t^1, \ldots, t^s$, состовляющих множество N_0. Значение в точке t^k можно выбрать произвольно из отрезка $\Delta^k := [\tau_0^k, \tau_1^k], \quad k = 1, \ldots, s$ являющегося полным прообразом $\tilde{t}^{-1}(t_k)$. Для определенности положим
$$\tilde{\tau}(t^k) = \tau_0^k, \quad k = 1, \ldots, s.$$
Тогда $\tilde{\tau}(t)$ непрерывна *слева* и $\tilde{\tau}(t_0) = \tau_0$.

Следующие два предложения будем называть предложениями о монотонных кусочно линейных функциях. В них используются функции $\tilde{t}(\tau)$, $\tilde{\tau}(t)$ и $\tilde{v}(\tau)$, определенные выше.

Предложение 1. *1. Пусть $\tilde{u}(\tau) : [\tau_0, \tau_1] \to \mathbb{R}^m$ — ограниченая измеримая функция такая, что $\tilde{u}(\tau) \in U$ п.в. на $[\tau_0, \tau_1]$*

Положим
$$u(t) = \tilde{u}(\tilde{\tau}(t)), \quad t \in [t_0, t_1].$$
Тогда $u(\tau) : [t_0, t_1] \to \mathbb{R}^m$ — ограниченная измеримая функция такая, что $u(t) \in U$ п.в. на $[t_0, t_1]$.

Доказательство. Достаточно проверить это утверждение на произвольном интервале множества N_+. Но на таком интервале $\tilde{\tau}(t)$ — линейная возрастающая функция и тогда измеримость, ограниченность (существенная) и условие $u(t) \in U$ п.в. - представляются очевидными. □

Второе, более сложное предложение мы сформулируем и докажем в начале следующей лекции.

Глава 10

Лекция 10

10.1 Тема: Принцип максимума, v-задача, теорема о неявной функции, уравнение в вариациях

10.1.1 О монотонных кусочно-линейных функциях (продолжение).

Предложение 2. *Пусть*

$$\tilde{x}(\tau) : [\tau_0, \tau_1] \to \mathbb{R}^n$$

— липшицева функция такая, что

$$\frac{d\tilde{x}(\tau)}{d\tau} = \tilde{v}(\tau)\tilde{\varphi}(\tau) \quad \text{п.в. на} \;\; [\tau_0, \tau_1], \tag{10.1}$$

где $\tilde{\varphi}(\tau) : [\tau_0, \tau_1] \to \mathbb{R}^n$ *ограниченная измеримая функция, а* $\tilde{v}(\tau)$ *— кусочно- постоянная неотрицательная функция. Пусть*

$$\tilde{t}(\tau) = t_0 + \int_{t_0}^{t} \tilde{v}(\tau) d\tau \qquad t \in [t_0, t_1] \tag{10.2}$$

и пусть $\tilde{\tau}(t) : [t_0, t_1] \to [\tau_0, \tau_1]$, *— правая обратная к* $\tilde{t}(\tau) : [\tau_0, \tau_1] \to [t_0, t_1]$, *непрерывная слева. Положим,*

$$x(t) = \tilde{x}(\tilde{\tau}(t)), \quad t \in [t_0, t_1].$$

Тогда $x(t) : [t_0, t_1] \to \mathbb{R}^n$ *— липшицева фунция такая, что*

$$x(t_0) = \tilde{x}(\tau_0), \quad x(t_1) = \tilde{x}(\tau_1). \tag{10.3}$$

Более того,

$$\frac{dx(t)}{dt} = \varphi(t) \;\; \text{п.в. на} \;\; [t_0, t_1], \tag{10.4}$$

где

$$\varphi(t) = \tilde{\varphi}(\tilde{\tau}(t)), \quad t \in [t_0, t_1].$$

Доказательство. Напомним, что на прошлой лекции через

$$\Delta^k = [\tau_0^k, \tau_1^k], \quad k = 1, \ldots, s$$

мы обозначили отрезки постоянства кусочно линейной функции $\tilde{t}(\tau)$, через t^k - ее значение на Δ^k, $k = 1, \ldots, s$. Мы положили

$$\mathcal{M}_0 = \bigcup_1^s \Delta^k, \quad \mathcal{M}_+ = [\tau_0, \tau_1] \backslash \mathcal{M}_0.$$

$$\mathcal{N}_0 = \{t^1, \ldots, t^s\}, \quad \mathcal{N}_+ = [t_0, t_1] \backslash \mathcal{N}_0.$$

а) Покажем, что $x(t)$ — липшицева функция. На любом интервале множества $\mathcal{N}_+$ функция $\tilde{\tau}(t)$ линейна и, следовательно, $x(t) = \tilde{x}(\tilde{\tau}(t))$ — липшицева (ибо $\tilde{x}(\cdot)$— липшицева). Следовательно, достаточно лишь показать, что $x(t)$ непрерывна в каждой точке t^k множества $\mathcal{N}_0$ (несмотря на то, что $\tilde{\tau}(t)$ в каждой такой точке разрывна). Отсюда последует липшицевость $x(t)$.

Заметим, что в силу уравнения (10.1) функция $\tilde{x}(\tau) = const$ на Δ^k ($k = 1, \dots, s$), ибо $\tilde{v}(\tau) = 0$ на Δ^k. Следовательно,

$$\tilde{x}(\tau_0^k) = \tilde{x}(\tau_1^k), \quad k = 1, \dots, s). \tag{10.5}$$

Возьмем произвольную точку $t^k \in \mathcal{N}_0$. Проверим непрерывность $x(t)$ в точке t^k. Имеем

$$x(t^k - 0) = \tilde{x}(\tilde{\tau}(t^k - 0)) = \tilde{x}(\tilde{\tau}(t^k)) = x(t^k), \tag{10.6}$$

так как $\tilde{\tau}$ непрерывна слева. При этом

$$\tilde{\tau}(t^k) = \tau_0^k \Rightarrow x(t^k - 0) = \tilde{x}(\tau_0^k). \tag{10.7}$$

Далее,

$$x(t^k + 0) = \tilde{x}(\tilde{\tau}(t^k + 0)) = \tilde{x}(\tau_1^k), \tag{10.8}$$

так как $\tilde{\tau}(t^k + 0) = \tau_1^k$. Из (10.5)-(10.8) следует, что

$$x(t^k - 0) = x(t^k) = \tilde{x}(\tau_0^k) = \tilde{x}(\tau_1^k) = x(t^k + 0).$$

Итак, $x(t)$ непрерывна в каждой точке t^k, а, значит, липшицева на $[t_0, t_1]$.

б) Докажем свойство (10.3) — совпадение концов.

б$_0$) По определению $\tilde{\tau}(t_0) = \tau_0$. Следовательно,

$$x(t_0) = \tilde{x}(\tilde{\tau}(t_0)) = \tilde{x}(\tau_0).$$

б$_1$) Если $\tau_1 \in M_+$, то есть τ_1 не является концом последнего отрезка постоянства функции $\tilde{t}(\tau)$ - отрезка Δ^s, то $\tilde{\tau}(t_1) = \tau_1$ и, следовательно,

$$x(t_1) = \tilde{x}(\tilde{\tau}(t_1)) = \tilde{x}(\tau_1).$$

Если же τ_1 является концом отрезка $\Delta^s = [\tau_0^s, \tau_1^s]$ (правым), то $\tilde{\tau}(t_1) = \tau_0^s$ (поскольку $\tilde{\tau}$ непрерывна слева). Следовательно,

$$x(t_1) = \tilde{x}(\tilde{\tau}(t_1)) = \tilde{x}(\tau_0^s).$$

Но в силу (10.5)

$$\tilde{x}(\tau_0^s) = \tilde{x}(\tau_1^s) = \tilde{x}(\tau_1).$$

Таким образом, и в этом случае $x(t_1) = \tilde{x}(\tau_1)$.

Вообще можно доказать, что $x(\tilde{t}(\tau)) = \tilde{x}(\tau) \quad \forall \tau \in [\tau_0, \tau_1]$, откуда (10.3) следует, ибо $\tilde{t}(\tau_0) = t_0$ и $\tilde{t}(\tau_1) = t_1$ (докажите это свойство).

в) Наконец, установим для $x(t) = \tilde{x}(\tilde{\tau}(t))$ свойство (10.4). Пусть $\mathcal{E} \subset [\tau_0, \tau_1]$ — множество полной меры, на котором:

(i) $\tilde{x}(t)$ имеет производную,

(ii) $\tilde{t}(\tau)$ имеет производную,

(iii) выполнено уравнение

$$\frac{d\tilde{x}(\tau)}{d\tau} = \frac{d\tilde{t}(\tau)}{d\tau}\tilde{\varphi}(\tau).$$

Пусть $\mathcal{E}_+ = \mathcal{E} \cap \mathcal{M}_+$. Поскольку $mes[\tilde{t}(M_+)] = t_1 - t_0$, то и $mes[\tilde{t}(\mathcal{E}_+)] = t_1 - t_0$

Действительно, $\mathcal{E}_+ \subset \mathcal{M}_+$ и если $m := \mathcal{M}_+ \backslash \mathcal{E}_+$, то $mes[m] = 0$, ибо $\mathcal{E}$ имеет полную меру в $[\tau_0, \tau_1]$. Тогда и $mes[\tilde{t}(m)] = 0$, ибо всякая липшицева функция, каковой является $\tilde{t}(\tau)$, переводит множество меры нуль в множество меры нуль (докажите). Тогда из включения

$$\tilde{t}(\mathcal{E}_+) = \tilde{t}(\mathcal{M}_+ \backslash m) \supset \tilde{t}(\mathcal{M}_+) \backslash \tilde{t}(m)$$

вытекает, что

$$mes[\tilde{t}(\mathcal{E}_+)] \geqslant mes[\tilde{t}(\mathcal{M}_+)] = t_1 - t_0 \Rightarrow mes[\tilde{t}(\mathcal{E}_+)] = t_1 - t_0.$$

Таким образом, $\tilde{t}(\mathcal{E}_+)$ имеет полную меру в $[t_0, t_1]$.

Пусть $t_* \in \tilde{t}(\mathcal{E}_+)$, то есть $t_* = \tilde{t}(\tau_*)$, где $\tau_* \in \mathcal{E}_+$ и тогда $\tilde{\tau}(t_*) = \tau_*$.

Покажем, что в точке t_* функция $x(t) = \tilde{x}(\tilde{\tau}(t))$ имеет производную. Действительно, в любой точке множества $\mathcal{N}_+$ функция $\tilde{\tau}$ имеет производную, значит, и в точке t_*. Поскольку $\tau_* \in \mathcal{E}_+$ и $\mathcal{E}_+ \in \mathcal{E}$, то $\tilde{x}$ имеет производную в точке τ_*. По правилу дифференцирования сложной функции

$$\frac{dx}{dt}(t_*) \stackrel{def}{=} \frac{d}{dt}\tilde{x}(\tilde{\tau}(t))\Big|_{t=t_*} = \frac{d\tilde{x}}{d\tau}(\tau_*) \cdot \frac{d\tilde{\tau}}{dt}(t_*) = (\text{т.к.}\tau_* \in \mathcal{E}) =$$

$$= \frac{d\tilde{t}}{d\tau}(\tau_*) \cdot \tilde{\varphi}(\tau_*) \cdot \frac{d\tilde{\tau}}{dt}(t_*) = \tilde{\varphi}(\tilde{\tau}(t_*)) \cdot \frac{d\tilde{t}}{d\tau}(\tau_*) \cdot \frac{d\tilde{\tau}}{dt}(t_*)$$

Но $\tilde{\varphi}(\tilde{\tau}(t_*)) = \varphi(t_*)$ по определению функции φ. Кроме того,

$$\frac{d\tilde{t}}{d\tau}(\tau_*) \cdot \frac{d\tilde{\tau}}{dt}(t_*) = 1,$$

поскольку

$$\tilde{t}(\tilde{\tau}(t)) = t \quad \forall t \in [t_0, t_1]$$

и обе производные существуют ($\tilde{t}$ дифференцируема в точке

$$\tilde{\tau}(t_*) = \tau_* \in \mathcal{E}_+ \subset \mathcal{M}_+,$$

а $\tilde{\tau}$ дифференцируема в точке t_*).

Итак, в любой точке $t_* \in \tilde{t}(\mathcal{E}_+)$ выполнено равенство

$$\frac{dx}{dt}(t_*) = \varphi(t_*).$$

Поскольку $\tilde{t}(\mathcal{E}_+)$ имеет полную меру в $[t_0, t_1]$, то равенство

$$\frac{dx}{dt} = \varphi$$

имеет место п.в. на $[t_0, t_1]$. □

10.1.2 Управляемая система задачи B и расширенная управляемая система задачи B, связь между ними.

Напомним формулировку леммы 2 прошлой лекции и докажем её.

Лемма 3. *Для любой траектории расширенной управляемой системы*

$$\frac{dx}{d\tau} = vf(x, u), \quad \frac{dt}{d\tau} = v, \quad u \in U, \quad v \geqslant 0, \tag{10.9}$$

где x и t абсолютно непрерывны, u — ограничена и измерима, а v — кусочно-постоянна, существует траектория системы

$$\frac{dx}{dt} = f(x, u), \quad u \in U \tag{10.10}$$

с теми же x_0, x_1.

Доказательство. Пусть

$$(\tilde{t}(\tau), \tilde{x}(\tau), \tilde{u}(\tau), \tilde{v}(\tau) \mid \tau \in [\tau_0, \tau_1])$$

— траектория решенной системы (10.9). Положим $t_0 = \tilde{t}(\tau_0),\ t_1 = \tilde{t}(\tau_1)$, и пусть

$$\tilde{\tau}(t) : [t_0, t_1] \to [\tau_0, \tau_1]$$

— правая обратная к $\tilde{t}(\tau)$, непрерывная слева. Положим

$$x(t) = \tilde{x}(\tilde{\tau}(t)), \quad u(t) = \tilde{u}(\tilde{\tau}(t)), \quad t \in [t_0, t_1]$$

Согласно Предложению 1 $u(t)$ - ограниченая измеримая функция такая, что $u(t) \in U$ п.в. на $[t_0, t_1]$.
Согласно Предложению 2 $x(t)$ — липшицева функция на $[t_0, t_1]$, причём,

$$x(t_0) = \tilde{x}(\tau_0), \quad x(t_1) = \tilde{x}(\tau_1)$$

$$\frac{dx(t)}{dt} = f(x(t), u(t)) \text{ п.в. на } [t_0, t_1].$$

Последнее, в силу Предложения 2, вытекает из условий

$$\frac{d\tilde{x}(\tau)}{d\tau} = \tilde{v}(\tau) f(\tilde{x}(\tau), \tilde{u}(\tau))$$

$$f(\tilde{x}(\tilde{\tau}(t)), \tilde{u}(\tilde{\tau}(t))) = f(x(t), u(t)).$$

Следовательно, $(x(t), u(t) \mid t \in [t_0, t_1])$ — траектория системы (10.10). Лемма доказана. □

10.1.3 Задача B и задача $\tilde{B}$, связь между ними.

Рассмотрим снова задачу B:

$$F_0(x_0, x_1) \to min, \quad F(x_0, x_1) \leqslant 0, \quad K(x_0, x_1) = 0,$$

$$\dot{x} = f(x, u), \quad u \in U,$$

где $x_0 = x(t_0)$, $x_1 = x(t_1)$, $x(t)$ — липшицева, а $u(t)$ - ограниченная измеримая на $[t_0, t_1]$, причем, t_0, t_1 - свободны.

Наряду с задачей B будем рассматривать задачу $\tilde{B}$:

$$F_0(x_0, x_1) \to min, \quad F(x_0, x_1) \leqslant 0, \quad K(x_0, x_1) = 0,$$

$$\frac{dx}{d\tau} = v f(x, u), \quad \frac{dt}{d\tau} = v, \quad v \geqslant 0, \quad u \in U,$$

где $x_0 = x(\tau_0)$, $x_1 = x(\tau_1)$, функции $x(\tau), t(\tau)$ — абсолютно непрерывны, $u(\tau)$ — ограничена и измерима, $v(\tau)$ — кусочно постоянна,
$\tau \in [\tau_0, \tau_1]$, τ_0, τ_1 - свободны.

Теорема. *(О связи задач B и $\tilde{B}$) Пусть*

$$\mathcal{T} = (x(t), u(t) \mid t \in [t_0, t_1])$$

— оптимальная траектория задачи B.
Пусть

$$\tilde{\mathcal{T}} = (\tilde{t}(\tau), \tilde{x}(\tau), \tilde{u}(\tau), \tilde{v}(\tau) \mid \tau \in [\tau_0, \tau_1])$$

— допустимая траектория задачи $\tilde{B}$ такая, что

$$\tilde{x}(\tau_0) = x(t_0), \quad \tilde{x}(\tau_1) = x(t_1).$$

Тогда $\tilde{\mathcal{T}}$ – оптимальная траектория задачи $\tilde{B}$.

Доказательство. Пусть $\tilde{\mathcal{T}}$ не оптимальна в задаче $\tilde{B}$. Тогда найдется допустимая траектория задачи $\tilde{B}$

$$\tilde{\mathcal{T}}' = (\tilde{t}'(\tau), \tilde{x}'(\tau), \tilde{u}'(\tau), \tilde{v}'(\tau) \mid \tau \in [\tau_0', \tau_1'])$$

с меньшим начением целевой функции:

$$F_0(\tilde{x}'(\tau_0'), \tilde{x}'(\tau_1')) < F_0(\tilde{x}(\tau_0), \tilde{x}(\tau_1)).$$

Расширенная управляемая система задачи B совпадает с управляемой системой задачи $\tilde{B}$. По лемме 2 найдется траектория управляемой системы задачи B

$$\mathcal{T}' = (x'(t), u'(t) \mid t \in [t_0', t_1'])$$

такая, что

$$x'(t_0') = \tilde{x}'(\tau_0'), \quad x'(t_1') = \tilde{x}'(\tau_1').$$

Тогда $\mathcal{T}'$ — допустимая траектория задачи B (поскольку $\tilde{\mathcal{T}}'$ допустима в задаче $\tilde{B}$) и при этом

$$F_0(x'(t_0'), x'(t_1')) < F_0(x(t_0), x(t_1)).$$

Значит, F_0 не оптимальна в задаче B. □

Мы провели важную подготовительную работу перед доказательством принципа максимума, связав между собой задачи B и $\tilde{B}$. Последнюю будем также называть *расширением задачи* B.

Перед доказательством принципа максимума, сформулируем и докажем еще один важный результат, относящийся к обыкновенным дифференциальным уравнениям: о гладкой зависимости решения от начальных условий и параметров. Для его доказательства нам понадобится теорема о неявной функции в абстрактной формулировке. Начнем с неё.

10.1.4 Теорема о неявной функции.

Теорема. *(АТФ, с. 166). Пусть X, Y, Z — банаховы пространства, W — окрестность точки $(x_0, y_0) \in X \times Y$.*

$$\Psi(x, y) : W \to Z$$

— оператор, непрерывно дифференцируемый по Фреше на W, причем

(a) $\Psi(x_0, y_0) = 0$,

(b) линейный оператор
$\Psi'_y(x_0, y_0) : Y \to Z$,

представляющий собой частную производную Фреше по переменной y оператора Ψ в точке (x_0, y_0), имеет обратный $[\Psi'_y(x_0, y_0)]^{-1} : Z \to Y$.

Тогда существует функция $y = \varphi(x)$, определенная в некоторой окрестности U точки x_0 и принимающая значения в Y, такая, что

(i) $\varphi(x_0) = y_0$;

(ii) $\Psi(x, \varphi(x)) = 0 \quad \forall x \in U$;

*(iii) в малой окрестности точки (x_0, y_0)
равенство $\Psi(x, y) = 0$ возможно
лишь при $y = \varphi(x)$;*

(iiii) функция φ — класса $C^1(U)$ и при этом
$\varphi'(x) = -[\Psi'_y(x, \varphi(x))]^{-1} \Psi'_x(x, \varphi(x))$.

Отметим, что последняя формула вытекает из (ii) и правила дифференцирования сложной функции.

Итак, равенство $\Psi(x, y) = 0$ задает неявную функцию $y = \varphi(x)$.

С доказательством можно ознакомиться в книге АТФ или известных руководствах по функциональному анализу.

Глава 11

Лекция

11.1 Тема 11: Принцип максимума.

11.1.1 Система уравнений в вариациях.

Пусть отрезок $[\tau_0, \tau_1]$ разбит на r последовательных интервалов, полуинтервалов или отрезков $\sigma_1, \ldots, \sigma_r$. Обозначим через $\chi_k(\tau)$ характеристическую функцию множества σ_k, $k = 1, \ldots, r$. Рассмотрим задачу Коши

$$\frac{dx}{d\tau} = f(x, \hat{u}(\tau)) \sum_{k=1}^{r} z_k \chi_k(\tau), \tag{11.1}$$

$$x(\tau_0) = x_0, \tag{11.2}$$

где z_k — числа, $k = 1, \ldots, r$, $x_0 \in \mathbb{R}^n$ - вектор, $\hat{u}(t)$ — фиксированная ограниченная измеримая функция.

Положим, $z = (z_1, \ldots, z_r)$. Как и выше, предполагаем, что f и f_x непрерывны по x, u.

Коротко задачу (11.1),(11.2) запишем в виде

$$\frac{dx}{d\tau} = vf(x, \hat{u}(\tau)), \quad x(\tau_0) = x_0, \tag{11.3}$$

где $v = \sum z_k \chi_k(\tau)$.

Теорема 1. *Пусть при некоторых $(\hat{x}_0, \hat{z}) \in \mathbb{R}^n \times \mathbb{R}^r$ задача Коши (11.1),(11.2) имеет решение $\hat{x}(\tau)$, определенное на $[\tau_0, \tau_1]$. Тогда в некоторой окрестности $\mathcal{U}$ точки $(\hat{x}_0, \hat{z}, v^0)$ задача (11.1),(11.2) определяет непрерывно дифференцируемый по Фреше оператор*

$$(x_0, z) \in \mathcal{U} \mapsto x(\cdot) \in W_1^1([\tau_0, \tau_1], \mathbb{R}^n), \tag{11.4}$$

причем производная Фреше этого оператора в точке $(\hat{x}_0, \hat{z}, v^0)$ есть линейный оператор, отображающий пару $(\bar{x}_0, \bar{z}) \in \mathbb{R}^n \times \mathbb{R}^r$ в функцию $\bar{x}(\cdot) \in W_1^1([\tau_0, \tau_1], \mathbb{R}^n)$, определяемую как решение задачи Коши

$$\frac{d\bar{x}}{d\tau} = f'_x(\hat{x}(\tau), \hat{u}(\tau))\bar{x} \sum_{k=1}^{r} \hat{z}_k \chi_k(\tau) + f(\hat{x}(\tau), \hat{u}(\tau)) \sum_{k=1}^{r} \bar{z}_k \chi_k(\tau), \tag{11.5}$$

$$\bar{x}(\tau_0) = \bar{x}_0. \tag{11.6}$$

Коротко задачу Коши (11.5),(11.6) запишем в виде

$$\frac{d\bar{x}}{d\tau} = \hat{v}\hat{f}_x\bar{x} + \bar{v}\hat{f}, \qquad \bar{x}(\tau_0) = \bar{x}_0, \tag{11.7}$$

где $\hat{v} = \sum \hat{z}_k \chi_k$, $\bar{v} = \sum \bar{z}_k \chi_k$, $\hat{f}_x = f_x(\hat{x}, \hat{u})$, $\hat{f} = f(\hat{x}, \hat{u})$. Система уравнений (11.5) называется *системой уравнений в вариациях* для системы (11.1).

В теореме 1 несущественно, что χ_k - характеристические функции множеств. Важно лишь то, что параметры z_k входят в правую часть системы (11.1) линейным образом (хотя и при нелинейном вхождении параметров теорема имеет соответствующий аналог).

Доказательство. (теоремы 1) Положим,

$$\Psi(x_0, z, x(\cdot)) = \Big(x(\tau_0) - x_0 \ , \ \frac{dx}{d\tau} - f(x, \hat{u}) \sum z_k \chi_k\Big) \tag{11.8}$$

Это равенство определяет оператор

$$\Psi : \mathbb{R}^n \times \mathbb{R}^r \times W_1^1 \mapsto \mathbb{R}^n \times L_1$$

такой, что

$$\Psi(\hat{x}_0, \hat{z}, \hat{x}(\cdot)) = (0, 0).$$

Последнее имеет место потому, что $\hat{x}(\cdot)$ есть решение задачи Коши (11.1),(11.2) при $x_0 = \hat{x}_0$, $z = \hat{z}$.

Нетрудно проверить (сделайте это сами), что Ψ непрерывно диференцируема по Фреше, и что его производная Фреше в точке $(\hat{x}_0, \hat{z}, \hat{x}(\cdot))$ есть линейный оператор

$$\hat{\Psi}' : \Big(\bar{x}_0, \bar{z}, \bar{x}(\cdot)\Big) \mapsto \Big(\bar{x}(\tau_0) - \bar{x}_0 \ , \ \frac{d\bar{x}}{d\tau} - \hat{f}_x \bar{x} \sum \hat{z}_k \chi_k - \hat{f} \sum \bar{z}_k \chi_k\Big) \tag{11.9}$$

$$\hat{\Psi}' : \mathbb{R}^n \times \mathbb{R}^r \times W_1^1 \mapsto \mathbb{R}^n \times L_1.$$

Частная производная оператора Ψ по переменной $x(\cdot)$ есть оператор

$$\hat{\Psi}'_{x(\cdot)} : \bar{x} \mapsto \Big(\bar{x}(\tau_0) \ , \ \frac{d\bar{x}}{d\tau} - \hat{f}_x \bar{x} \sum \hat{z}_k \chi_k\Big),$$

$$\hat{\Psi}'_{x(\cdot)} : W_1^1 \mapsto \mathbb{R}^n \times L_1.$$

Он получается из $\hat{\Psi}'$ при $\bar{x_0} = 0, \ \bar{z} = 0$. Покажем, что $\Psi'_{x(\cdot)}$ имеет обратный.

В образе этого оператора возьмем произвольный элемент $(\bar{\xi}, \bar{f})$, где

$$\bar{\xi} \in \mathbb{R}^n, \quad \bar{f}(\cdot) \in L_1,$$

и рассмотрим задачу Коши

$$\frac{d\bar{x}}{d\tau} = \hat{f}_x \bar{x} \sum \hat{z}_k \chi_k + \bar{f}, \quad \bar{x}(\tau_0) = \bar{\xi}. \tag{11.10}$$

Поскольку $\hat{f}_x(\cdot) \in L_\infty$, то при любых $\bar{\xi} \in \mathbb{R}^n$, $\bar{f} \in L_1$ эта задача имеет единственное решение $\bar{x}(\tau)$ на $[\tau_0, \tau_1]$, $\bar{x}(\cdot) \in W_1^1$ (см. АТФ, с. 191). Таким образом, задача (11.10) определяет линейный оператор

$$(\bar{\xi}, \bar{f}) \in \mathbb{R}^n \times L_1 \mapsto \bar{x}(\cdot) \in W_1^1,$$

который и является обратным к оператору $\hat{\Psi}'_{x(\cdot)}$.

Итак, для оператора Ψ выполнены условия теоремы о неявной функции. По этой теореме существует непрерыано диференцируемая по Фреше функция (оператор)

$$\varphi : (x_0, z) \mapsto x(\cdot),$$

определенная в некоторой окрестности $\mathcal{U}$ точки $(\hat{x}_0, \hat{z})$:

$$\varphi : \mathcal{U} \mapsto W_1^1,$$

такая, что

$$\varphi(\hat{x}_0, \hat{z}) = \hat{x}(\cdot)$$

и при этом значение $x(\cdot) = \varphi(x_0, z)$ определяется из условия

$$\Psi(x_0, z, x(\cdot)) = (0, 0),$$

т.е. как решение задачи Коши (11.1),(11.2).

Производная Фреше операторв φ в точке $(\hat{x}_0, \hat{z})$ есть линейный оператор $\hat{\varphi}'$ определяемый из условия

$$\hat{\Psi}'_{x_0 z} + \hat{\Psi}'_{x(\cdot)} \hat{\varphi}' = 0,$$

где $\hat{\Psi}'_{x_0 z}$ — частная производная оператора Ψ по x_0, z в точке $(\hat{x}_0, \hat{z}, \hat{x}(\cdot))$. Из приведенного равенства вытекает, что

$$\hat{\Psi}'_{x_0 z}(\bar{x}_0, \bar{z}) + \hat{\Psi}'_{x(\cdot)}\hat{\varphi}'(\bar{x}_0, \bar{z}) = 0, \quad \forall \bar{x}_0, \bar{z} \in \mathbb{R}^n \times \mathbb{R}^r.$$

Положим

$$\hat{\varphi}'(\bar{x}_0, \bar{z}) = \bar{x}(\cdot).$$

Тогда

$$\hat{\Psi}'_{x_0 z}(\bar{x}_0, \bar{z}) + \hat{\Psi}'_{x(\cdot)}\bar{x}(\cdot) = 0,$$

что эквивалентно равенству

$$\hat{\Psi}'(\bar{x}_0, \bar{z}, \bar{x}(\cdot)) = 0. \tag{11.11}$$

Итак, $\bar{x}(\cdot)$ определяется (посредством $\hat{\varphi}'$) по $(\bar{x}_0, \bar{z})$ из равенства (11.11). В силу (11.9) это означает, что $\bar{x}(\cdot)$ находится как решение задачи Коши (11.5),(11.6). Теорема доказана. □

Далее через p обозначим пару (x_0, x_1), где $x_0 \in \mathbb{R}^n$, $x_1 \in \mathbb{R}^n$, т.е. $p \in \mathbb{R}^{2n}$.

Введем конечномерный оператор $p = P(x_0, z)$ такой, что

$$P : (x_0, z) \in \mathcal{U} \mapsto p = (x_0, x(\tau_1)) \in \mathbb{R}^{2n}, \tag{11.12}$$

где $x(\cdot) = \varphi(x_0, z)$ — решение задачи Коши (11.1),(11.2), $\mathcal{U} \in \mathbb{R}^n \times \mathbb{R}^r$ — окрестность точки $(\hat{x}_0, \hat{z})$, указанная в теореме 1. Мы считаем, что условия этой теоремы выполнены. Непосредственно из теоремы 1 вытекает следующая теорема.

Теорема 2. *Конечномерный оператор (11.12) определен и непрерывно дифференцируем в некоторой окрестности $\mathcal{U}$ точки $(\hat{x}_0, \hat{z})$. Его производная в этой точке есть конечномерный линейный оператор*

$$\hat{P}' : (\bar{x}_0, \bar{z}) \in \mathbb{R}^n \times \mathbb{R}^r \mapsto \bar{p} = (\bar{x}_0, \bar{x}_1) \in \mathbb{R}^{2n},$$

где $\bar{x}_1 = \bar{x}(\tau_1)$, а $\bar{x}(\cdot)$ определяется как решение задачи Коши (11.5),(11.6).

11.2 Тема: принцип максимума задачи B.

11.2.1 Формулировка принципа максимума задачи B.

Напомним постановку задачи B:

$$J = F(x_0, x_1) \to min, \quad F(x_0, x_1) \leqslant 0, \quad K(x_0, x_1) = 0, \tag{11.1}$$

$$\dot{x} = f(x, u), \quad u \in U, \tag{11.2}$$

где $x_0 = x(t_0)$, $x_1 = x(t_1)$, а t_0, t_1 - свободны.

Здесь (11.1) — *концевой блок задачи*, (11.2) — *управляемая система* задачи. Положим

$$l = l(x_0, x_1, \alpha_0, \alpha, \beta) = \alpha_0 F_0(x_0, x_1) + \alpha F(x_0, x_1) + \beta K(x_0, x_1),$$

где α_0 — число, α и β - величины тех же размерностей, что и F и K соответственно, т.е. $\alpha \in \mathbb{R}^k, \beta \in \mathbb{R}^s$. Назовем l - *концевой функцией Лагранжа*. Удобно считать, что α и β — вектор-строки.

Далее положим,

$$H = H(t, x, u, \psi) = \psi f(t, x, u),$$

где $\psi \in \mathbb{R}^n$ — вектор-строка. Назовем H — *функцией Понтрягина.*

Пусть $\mathcal{T} = \{x(t), u(t) \mid t \in [t_0, t_1]\}$ — допустимая траектория задачи B. Будем говорить, что для нее выполнены условия *принципа максимума Понтрягина*, если существуют число α_0, векторы

$$\alpha \in \mathbb{R}^k, \quad \beta \in \mathbb{R}^s$$

и абсолютно непрерывная функция

$$\psi(t) : [t_0, t_1] \to \mathbb{R}^n,$$

такие, что

(i) $\alpha_0 \geqslant 0, \alpha \geqslant 0,$

(ii) $\alpha_0 + |\alpha| + |\beta| > 0$,

(iii) $\alpha_i F_i(x(t_0), x(t_1)) = 0, \quad i = 1, \ldots, k$,

(iv) $-\dot{\psi}(t) = H'_x(t, x(t), u(t), \psi(t))$ п.в. на $[t_0, t_1]$,

(v) $\psi(t_0) = l'_{x_0}(x(t_0), x(t_1), \alpha_0, \alpha, \beta)$,
$-\psi(t_1) = l'_{x_1}(x(t_0), x(t_1), \alpha_0, \alpha, \beta)$,

(vi) $H(t, x(t), u, \psi(t)) \leqslant 0, \quad \forall t \in [t_0, t_1] \ \ \forall u \in U$

(vii) $H(t, x(t), u(t), \psi(t)) = 0$, п.в. на $[t_0, t_1]$.

Из условий (vi) и (vii) следует, что почти для всех $t \in [t_0, t_1]$ функция

$$u \mapsto H(t, x(t), u, \psi(t))$$

достигает максимума по $u \in U$ в точке $u = u(t)$. Поэтому (vi) называют *условием максимума.*

Функцию

$$\mathcal{H}(t, x, \psi) = \max_{u \in U} H(t, x, u, \psi)$$

называют *гамильтонианом.*

Из сказанного вытекает, что

$$\mathcal{H}(t, x(t), \psi(t)) = H(t, x(t), u(t), \psi(t)) \text{ п.в. на } [t_0, t_1].$$

Поэтому условие (vii) называют *условием равенства нулю гамильтониана.*

Используют также следующие названия:
(i) — условия неотрицательности,
(ii) — условие нетривиальности,
(iii) — условия дополняющей нежесткости,
(iv) — сопряженная система (для ψ),
(v) — условия трансверсальности (для ψ).

Переменную ψ называют двойственой или сопряженной (adjoint) переменной, а α_0, α, β — *множителями Лагранжа концевого блока задачи.*

Набор $\lambda = (\alpha_0, \alpha, \beta, \psi(\cdot))$ называют *набором множителей Лагранжа* задачи B.

Теорема 3. *Пусть траектория*

$$\mathcal{T} = (x(t), u(t) \mid t \in [t_0, t_1])$$

— оптимальная в задаче B. Тогда для нее выполнены условия принципа максимума Понтрягина.

Мы переходим непосредственно к доказательству этой теоремы.

11.2.2 Индекс θ.

Пусть

$$\hat{\mathcal{T}} = (\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t_0}, \hat{t_1}])$$

— решение задачи B. С этим решением мы свяжем семейство конечномерных задач B^θ и их оптимальных решений (x_0^θ, z^θ), занумерованных индексом θ.

Под индексом будем понимать набор значений времени и управления

$$\theta = \{t^1, \ldots, t^N, u^1, \ldots, u^N\}$$

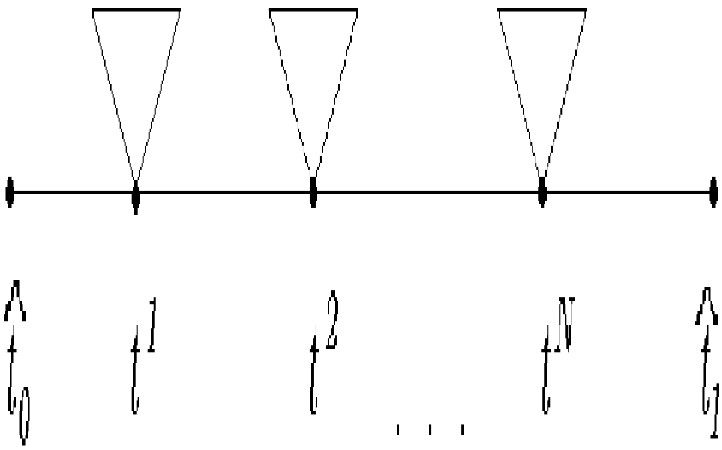
такой, что

$$\hat{t}_0 \leqslant t_1 \leqslant \ldots \leqslant t^N \leqslant \hat{t}_1$$

$$u^k \in U, \quad k = 1, \ldots, N.$$

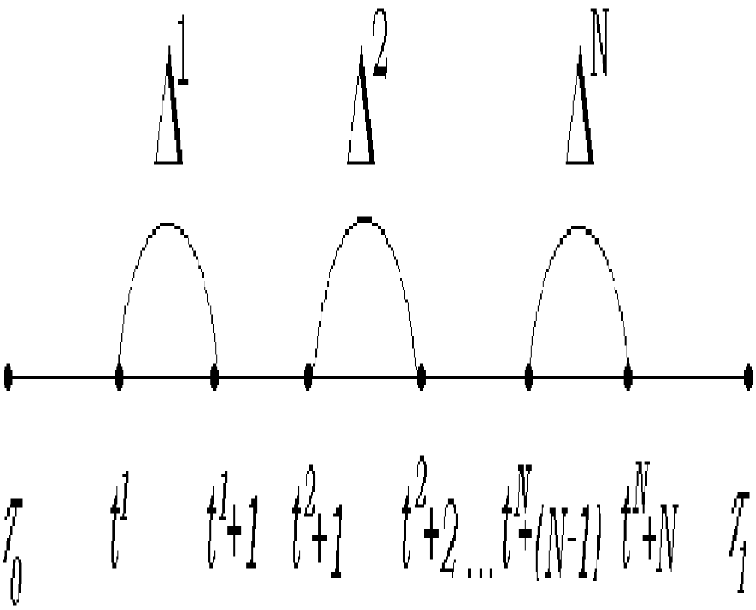
Здесь $N = N(\theta)$ — зависит от θ.

Определим отрезок $[\tau_0, \tau_1]$ следующим образом: берем отрезок $[\hat{t}_0, \hat{t}_1]$ и в точках $t^1, \ldots, t^N$ вставляем отрезки единичной длины , сохраняя всякий раз положение точки $\hat{t}_0$.



В результате получаем отрезок $[\tau_0, \tau_1]$ с концами

$$\tau_0 = \hat{t}_0, \quad \tau_1 = \hat{t}_1 + N.$$



Положим

$$\begin{array}{l} \Delta^1 = [t^1, t^1 + 1], \\ \Delta^2 = [t^2 + 1, t^2 + 2], \\ \dots\dots\dots\dots\dots\dots \\ \Delta^N = [t^N + (N-1), t^N + N], \\ \mathcal{M}_0 = \bigcup\limits_1^N \Delta^k, \quad \mathcal{M}_+ = [\tau_0, \tau_1] \backslash \mathcal{M}_0 \end{array}$$

Пусть

$$v^\theta(\tau) = \begin{cases} 0, & \tau \in M_0, \\ 1, & \tau \in M_+, \end{cases}$$

$$t^\theta(\tau) = \hat{t}_0 + \int\limits_{\tau_0}^{\tau} v^\theta(s) ds, \quad \tau \in [\tau_0, \tau_1].$$

Тогда

$$\frac{dt^\theta(\tau)}{d\tau} = v^\theta(\tau), \quad t^\theta(\tau_0) = \hat{t}_0.$$

Ясно, что $t^\theta(\tau_1) = \hat{t}_1$. Таким образом, $t^\theta(\tau)$ — кусочно-линейная неубывающая функция такая, что

$$t^\theta(\tau) : [\tau_0, \tau_1] \overset{HA}{\to} [\hat{t}_0, \hat{t}_1],$$

причем, Δ^k — её отрезки постоянства и

$$t^\theta(\Delta^k) = t^k, \quad k = 1, \dots, s.$$

Положим

$$u^\theta(\tau) = \begin{cases} \hat{u}(t^\theta(\tau)), & \tau \in M_+, \\ u^k, \quad \tau \in \Delta^k, & k = 1, \dots, N \end{cases}$$

Тогда $u^\theta(\tau) : [\tau_0, \tau_1] \to \mathbb{R}^m$ — ограниченная измеримая функция удовлетворяющая условию $u^\theta(\tau) \in U$ п.в. на $[\tau_0, \tau_1]$ (почему?).

Положим

$$x^\theta(\tau) = \hat{x}(t^\theta(\tau)), \quad \tau \in [\tau_0, \tau_1].$$

Предложение 1. *(a) Имеют место равенства*

$$x^\theta(\tau_0) = \hat{x}(\hat{t}_0), \quad x^\theta(\tau_1) = \hat{x}(\hat{t}_1); \tag{11.3}$$

(b) п.в. на $[\tau_0, \tau_1]$ имеет место равенство

$$\frac{dx^\theta(\tau)}{d\tau} = v^\theta(\tau) f(x^\theta(\tau), u^\theta(\tau)). \tag{11.4}$$

Доказательство. (a) — вытекает из условий $t^\theta(\tau_0) = \hat{t}_0, \ t^\theta(\tau_1) = \hat{t}_1$.

(b) — Пусть Δ^k — один из отрезков множества $\mathcal{M}_0$. Тогда из условия $t^\theta(\tau) = const$ на $\Delta^k \Rightarrow \frac{dx^\theta(\tau)}{d\tau} = 0$ на Δ^k.

Одновременно

$$v^\theta(\tau) = 0 \quad \text{на } \Delta^k.$$

Следовательно (11.4) имеет место на Δ^k, а, значит , п.в. на $\mathcal{M}_0$.

Пусть теперь Δ_+ — интервал множества $\mathcal{M}_+$. Согласно формуле замены переменной на Δ_+ имеем:

$$\frac{dx^\theta(\tau)}{d\tau} = \frac{d\hat{x}(t^\theta(\tau))}{dt} \cdot \frac{dt^\theta(\tau)}{d\tau} =$$

$$= f(\hat{x}(t^\theta(\tau)), \hat{u}(t^\theta(\tau)))v^\theta(\tau) = f(x^\theta(\tau), u^\theta(\tau))v^\theta(\tau)$$

Следовательно (11.4) имеет место на на Δ_+, а тогда и п.в. на $\mathcal{M}_+$. Но $\mathcal{M}_0 \cup \mathcal{M}_+ = [\tau_0, \tau_1]$.

Значит, (11.4) имеет место п.в. на $[\tau_0, \tau_1]$. □

Множество $\mathcal{M}_0$ состоит из конечного числа отрезков Δ^k , $k = 1, \ldots, N$. Множество $\mathcal{M}_+$ - из конечного числа интервалов или полуинтервалов. Все указанные отрезки, интервалы и полуинтервалы множеств $\mathcal{M}_0$ и $\mathcal{M}_+$ обьединим в набор, упорядочим и обозначим составляющие этого набора через $\sigma_k, \quad k = 1, \ldots, r$. Итак, $[\tau_0, \tau_1] = \bigcup_1^r \sigma_k$, причем, различные σ_k не пересекаются. Обозначим через $\chi_k(\tau)$ характеристическую функцию множества $\sigma_k, \quad k = 1, \ldots, r$.

Пусть $\hat{z}_k$ — значение $v^\theta(\tau)$ на $\sigma_k,\ k = 1, \ldots, r$, т.е.

$$v^\theta(\tau) = \sum_{k=1}^{r} \hat{z}_k \chi_k(\tau) \ \text{ п.в..}$$

Положим $\hat{z} = (\hat{z}_1, \ldots, \hat{z}_r),\ \hat{x}_0 = \hat{x}(\hat{t}_0)$. Напомним, что $\hat{z}_k = 0$, если $\sigma_k \subset \mathcal{M}_0$, и $\hat{z}_k = 1$, если $\sigma_k \subset \mathcal{M}_+$.

Глава 12

Лекция 12

12.1 Тема: Доказательство принципа максимума задачи B.

12.1.1 Задача B^θ индекса θ и ее оптимальное решение.

Итак, для оптимальной траектории $\hat{\mathcal{T}} = \{\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t}_0, \hat{t}_1]\}$ задачи B и индекса $\theta = \{t^1, \dots, t^N, u^1, \dots, u^N\}$ мы определим отрезок $[\tau_0, \tau_1]$ и функции $v^\theta, t^\theta, u^\theta, x^\theta$ на этом отрезке. Вместе они определяют траекторию

$$\mathcal{T}^\theta = \{t^\theta(\tau), x^\theta(\tau), u^\theta(\tau), v^\theta(\tau) \mid \tau \in [\tau_0, \tau_1]\} \tag{12.1}$$

в задаче $\tilde{B}$. Последняя была определена в лекции 10 и имела вид:

$$F_0(p) \to min, \;\; F(p) \leqslant 0 \;, K(p) = 0,$$

$$\frac{dx}{d\tau} = vf(x, u), \quad \frac{dt}{d\tau} = v,$$
$$u \in U, \quad v \geqslant 0,$$

где $p = (x_0, x_1) = (x(\tau_0), x(\tau_1))$,

$t(\tau), x(\tau)$ — абсолютно непрерывны,

$u(\tau)$ — ограничена и измерима, а $v(\tau)$ — кусочно постоянна, τ_0 и τ_1 — свободны.

Теорема 1. *Для любого индекса θ траектория $\mathcal{T}^\theta$ оптимальна в задаче $\tilde{B}$.*

Доказательство. Согласно предложению 1, пункт (b) прошлой лекции

$$\frac{dx^\theta}{d\tau} = v^\theta f(x^\theta, u^\theta) \text{ п.в..}$$

Кроме того,

$$\frac{dt^\theta}{d\tau} = v^\theta, \quad v^\theta \geqslant 0, \quad u^\theta \in U \text{ п.в.}$$

— согласно определениям функций v^θ,t^θ,u^θ. Наконец, в силу пункта (a) предложения 1 прошлой лекции $x^\theta(\tau_0) = \hat{x}(t_0)$, $x^\theta(\tau_1) = \hat{x}(\hat{t}_1)$. Отсюда и из оптимальности траектории $\hat{\mathcal{T}}$ в задаче B вытекает, согласно теореме 1 лекции 10 (о связи задач B и $\tilde{B}$) оптимальность траектории $\mathcal{T}^\theta$ в задаче $\tilde{B}$. ☐

Для индекса θ мы определили разбиение отрезка $[\tau_0, \tau_1]$ на отрезки $\sigma_1, \dots, \sigma_r$ (это отрезки Δ^k множества $\mathcal{M}_0$ и смежные с ними интервалы или полуинтервалы, принадлежащие $\mathcal{M}_+$). Рассмотрим оператор

$$P : (x_0, z) \mapsto p = (x_0, x(\tau_1)),$$

где $x(\tau) : [\tau_0, \tau_1] \to \mathbb{R}^n$ определяется как решение задачи Коши

$$\frac{dx}{d\tau} = f(x, u^\theta(\tau)) \sum z_k \chi_k(\tau) \tag{12.2}$$

$$x(\tau_0) = x_0$$

(через χ_k мы обозначили характеристические функции множеств σ_k)

Поскольку

$$v^\theta(\tau) := \sum \hat{z}_k \chi_k(\tau)$$

$$\frac{dx^\theta}{d\tau} = v^\theta f(x^\theta, u^\theta)$$

$$\hat{x}_0 := \hat{x}(\hat{t}_0) = x^\theta(\tau_0),$$

то при $x = \hat{x}_0, z = \hat{z}$ задача Коши (1) имеет решение $x^\theta(\tau)$, определенное на $[\tau_0, \tau_1]$, тогда по теореме 2 прошлой лекции оператор P определен в некоторой окрестности $\mathcal{U}$ точки $(\hat{x}_0, \hat{z})$ в пространстве $\mathbb{R}^n \times \mathbb{R}^r$.

Теперь мы можем поставить конечномерную задачу B^θ. Она рассматривается в пространстве переменных

$$(x_0, z, p) \in \mathbb{R}^n \times \mathbb{R}^r \times \mathbb{R}^{2n}.$$

12.1.2 Задача B^θ.

$$F_0(p) \to min, \quad F(p) \leqslant 0 \quad, K(p) = 0$$

$$-z \leqslant 0, \quad p - P(x_0, z) = 0, \quad (x_0, z) \in U,$$

где $P(x_0, z) = (x_0, x(\tau_1))$, а функция $x(\cdot) = \varphi(x_0, z)$ определяется как решение задачи Коши (12.2).

Положим $\hat{x}_1 = \hat{x}(\hat{t}_1)$, $\hat{p} = (\hat{x}_0, \hat{x}_1)$. Напомним, что $\hat{z}$ определяется индексом θ, т.е. $\hat{z} = \hat{z}(\theta)$.

Теорема 2. *Для любого индекса θ точка $(\hat{x}_0, \hat{z}, \hat{p})$ есть решение задачи B^θ.*

Доказательство. Предположим противное: пусть для некоторого индекса θ в задаче B^θ имеется допустимая точка (x_0, t, p) такая, что

$$F_0(p) < F_0(\hat{p}) \tag{12.3}$$

Пусть $x(\cdot) = \varphi(x_0, z)$, $v(\cdot) = \sum z_k \chi_k(\cdot)$, а $t(\cdot)$ найдена из условий

$$\frac{dt}{d\tau} = v, t(\tau_0) = t_0,$$

t_0 выбрано произвольно. Тогда из определений следует, что

$$\tilde{\mathcal{T}} = \{t(\tau), x(\tau), u^\theta(\tau), v(\tau) \mid \tau \in [\tau_0, \tau_1]\}$$

— допустимая траектория задачи $\tilde{B}$ такая, что

$$(x(\tau_0), x(\tau_1)) = p.$$

Согласно теореме 1 траектория $\mathcal{T}^\theta$ (12.2) оптимальна в задаче $\tilde{B}$. Однако в силу (12.3) $\tilde{\mathcal{T}}$ доставляет меньшее значение целевой функции F_0, чем траектория $\mathcal{T}^\theta$. Противоречие доказывает теорему. □

12.1.3 Условие стационарности в задаче B^θ.

Фиксируем индекс θ. Согласно теореме 1 $(\hat{x}_0, \hat{z}, \hat{p})$ — точка минимума в задаче B^θ. Следовательно она стационарная точка в той же задаче. Выпишем условие стационарности в этой точке.

Функция Лагранжа задачи B^θ имеет вид

$$L = \alpha_0 F_0(p) + \alpha F(p) + \beta K(p) - \mu z + \nu(p - P(x_0, z)),$$

где $\mu \in (\mathbb{R}^r)^*$, $\nu \in (\mathbb{R}^{2n})^*$ — вектор-строки. Пусть $l = \alpha_0 F_0 + \alpha F + \beta K$. Тогда

$$l = l(p, \alpha_0, \alpha, \beta),$$

$$L = l - \mu z + \nu(p - P(x_0, z)),$$

$$L = L(x_0, z, p, \alpha_0, \alpha, \beta, \mu, \nu).$$

Производные функции Лагранжа по переменым x_0,z и p в точке $(\hat{x}_0, \hat{z}, \hat{p})$ имеют вид

$$L_{x_0} = -\nu P_{x_0}(\hat{x}_0, \hat{z}),$$

$$L_z = -\mu - \nu P_z(\hat{x}_0, \hat{z}),$$

$$L_p = l_p(\hat{p}, \alpha_0, \alpha, \beta) + \nu.$$

Следовательно условия стационарности в точке $(\hat{x}_0, \hat{z}, \hat{p})$ имеют вид:

$$\alpha_0 \geqslant 0, \alpha \geqslant 0, \mu \geqslant 0, \tag{12.4}$$

$$\alpha_0 + |\alpha| + |\beta| + |\mu| + |\nu| > 0, \tag{12.5}$$

$$\alpha F(\hat{p}) = 0, \quad \mu\hat{z} = 0, \tag{12.6}$$

$$\nu P_{x_0}(\hat{x}_0, \hat{z}) = 0, \tag{12.7}$$

$$\nu P_z(\hat{x}_0, \hat{z}) + \mu = 0, \tag{12.8}$$

$$l_p(\hat{p}, \alpha_0, \alpha, \beta) + \nu = 0. \tag{12.9}$$

Если $\alpha_0 + |\alpha| + |\beta| = 0$, то $l_p(\hat{p}, \alpha_0, \alpha, \beta) = 0$, а тогда в силу (12.9) $\nu = 0$ и в силу (12.8) $\mu = 0$. Следовательно, условие нетривиальности (12.5) эквивалентно условию $\alpha_0 + |\alpha| + |\beta| > 0$. Не ограничивая общности, можно считать, что $\alpha_0 + |\alpha| + |\beta| = 1$. Окончательно получаем следующую систему условий стационарности:

$$\begin{cases} \alpha_0 \geqslant 0, \alpha \geqslant 0, \mu \geqslant 0, \\ \alpha_0 + |\alpha| + |\beta| = 1, \\ \alpha F(\hat{p}) = 0, \quad \mu\hat{z} = 0, \\ \nu\hat{P}_{x_0} = 0, \ \nu\hat{P}_z + \mu = 0, \ \nu = -\hat{l}_p, \end{cases} \tag{12.10}$$

где $\hat{P}_{x_0} = P_{x_0}(\hat{x}_0, \hat{z}), \ \hat{P}_z = P_z(\hat{x}_0, \hat{z}), \ \hat{l}_p = l(\hat{p}, \alpha_0, \alpha, \beta)$.

Найдем частные производные $\hat{P}_{x_0}, \hat{P}_z$ оператора P по переменым x_0 и z в точке $(\hat{x}_0, \hat{z})$.

Напомним, что согласно Теореме 2 прошлой лекции производная оператора P в точке $(\hat{x}_0, \hat{z})$ есть линейный оператор

$$\hat{P}' : (\bar{x}_0, \bar{z}) \mapsto (\bar{x}_0, \bar{x}(\tau_1)),$$

где $\bar{x}(\cdot)$ есть решение задачи Коши для уравнениях в вариациях

$$\frac{d\bar{x}}{d\tau} = v^\theta f_x^\theta \bar{x} + \sum \bar{z}_k \chi_k f^\theta, \quad \bar{x}(\tau_0) = \bar{x}_0. \tag{12.11}$$

Мы обозначили здесь

$$f_x^\theta = f_x(x^\theta, u^\theta), \quad f^\theta = f(x^\theta, u^\theta). \tag{12.12}$$

Следовательно,

$$\hat{P}_{x_0} : \bar{x}_0 \mapsto (\bar{x}_0, \bar{x}(\tau_1)),$$

где $\bar{x}(\cdot)$ определяется как решение задачи Коши

$$\frac{d\bar{x}}{d\tau} = v^\theta f_x^\theta \bar{x}, \quad \bar{x}(\tau_0) = \bar{x}_0 \tag{12.13}$$

(В условиях (12.11) мы положили $\bar{z} = 0$). Аналогично

$$\hat{P}_z : \bar{z} \mapsto (0, \bar{x}(\tau_1)) \in \mathbb{R}^{2n},$$

где $\bar{x}(\cdot)$ есть решение задачи Коши

$$\frac{d\bar{x}}{d\tau} = v^\theta f_x^\theta \bar{x} + \sum \bar{z}_k \chi_k f^\theta, \quad \bar{x}(\tau_0) = 0. \tag{12.14}$$

(В условиях (12.11) мы положили $\bar{x}_0 = 0$)

12.1.4 Сопряженное уравнение и условия трансверсальности.

В условиях стационарности (12.10) проанализируем условие $\nu\hat{P}_{x_0}=0$. Оно эквивалентно условию

$$\nu\hat{P}_{x_0}\bar{x}_0=0 \quad \forall\bar{x}_0\in\mathbb{R}^n$$

Последнее можно представить в виде:

$$\nu_0\bar{x}_0+\nu_1\bar{x}_1=0 \tag{12.15}$$

для любых $\bar{x}_0\in\mathbb{R}^n$, $\bar{x}_1\in\mathbb{R}^n$ таких, что

$$\bar{x}_0=\bar{x}(\tau_0),\quad \bar{x}_1=\bar{x}(\tau_1), \tag{12.16}$$

а $\bar{x}(\tau)$ удовлетворяет условию

$$\frac{d\bar{x}}{d\tau}=v^\theta f_x^\theta\bar{x}. \tag{12.17}$$

Здесь $\nu=(\nu_0,\nu_1)$, $\nu_0\in\mathbb{R}^n$, $\nu_1\in\mathbb{R}^n$. Итак (12.17),(12.16)$\Rightarrow$ (12.15). Проанализируем это условие.

Пусть $A=A(\tau)$ - квадратная матрица порядка n с ограничеными измеримыми коэффициентами, определенными на $[\tau_0,\tau_1]$. Рассмотрим линейную систему

$$\frac{d\bar{x}}{d\tau}=A(\tau)\bar{x}. \tag{12.18}$$

Систему

$$-\frac{d\psi}{d\tau}=\psi A(\tau) \tag{12.19}$$

называют сопряженной к системе (12.18). Здесь $\psi=(\psi_1,\dots,\psi_n)$ — вектор-строка, в то время как x — вектор-столбец.

Пусть заданы ν_0 и ν_1 - n-мерные вектор-строки.

Лемма 4. *Пусть для любой $\bar{x}(\cdot)\in W^1_\infty([\tau_0,\tau_1],\mathbb{R}^n)$, являющейся решением на $[\tau_0,\tau_1]$ системы (12.18), выполнено условие*

$$\nu_0\bar{x}(\tau_0)+\nu_1\bar{x}(\tau_1)=0. \tag{12.20}$$

Тогда на отрезке $[\tau_0,\tau_1]$ существует липшицева функция $\psi(\tau)$, являющаяся решением (на этом отрезке) сопряженой системы (12.19) и такая, что

$$\psi(\tau_0)=-\nu_0,\quad \psi(\tau_1)=\nu_1. \tag{12.21}$$

Доказательство. Определим ψ из условий

$$-\dot{\psi}=\psi A,\quad \psi(\tau_1)=\nu_1.$$

Покажем, что $\psi(\tau_0)=-\nu_0$. Пусть $\bar{x}_0\in\mathbb{R}^n$ - произвольный вектор, а $\bar{x}(\cdot)$ найдена из условий

$$\dot{\bar{x}}=A\bar{x},\quad \bar{x}(\tau_0)=\bar{x}_0.$$

Тогда

$$\frac{d}{d\tau}(\psi\bar{x})=\frac{d\psi}{d\tau}\bar{x}+\psi\frac{d\bar{x}}{d\tau}=-\psi A\bar{x}+\psi A\bar{x}=0.$$

Следовательно,

$$\int_{\tau_0}^{\tau_1}\frac{d}{d\tau}(\psi\bar{x})d\tau=(\psi\bar{x})\Big|_{\tau_0}^{\tau_1}=\nu_1\bar{x}(\tau_1)-\psi(\tau_0)\bar{x}_0\overset{(12.20)}{=}-\nu_0\bar{x}_0-\psi(\tau_0)\bar{x}_0.$$

Итак, $(\nu_0+\psi(\tau_0))\bar{x}_0=0$ $\forall\bar{x}_0\in\mathbb{R}^n$, следовательно, $\psi(\tau_0)=-\nu_0$. □

Вернемся к условиям (12.15), (12.16), (12.17). Согласно Лемме 1 на $[\tau_0,\tau_1]$ существует липшицева функция $\psi^\theta(\tau)$ такая, что

$$-\frac{d\psi^\theta}{d\tau}=\psi^\theta v^\theta f_x^\theta, \tag{12.22}$$

$$\psi^\theta(\tau_0)=-\nu_0,\quad \psi^\theta(\tau_1)=\nu_1. \tag{12.23}$$

Но из условия $\nu=-\hat{l}_p$ вытекает, что

$$\nu_0=-\hat{l}_{x_0},\quad \nu_1=-\hat{l}_{x_1},$$

Следовательно

$$\psi^\theta(\tau_0)=\hat{l}_{x_0},\quad \psi^\theta(\tau_1)=-\hat{l}_{x_1}. \tag{12.24}$$

Мы получили для сопряженое уравнение (12.22) и условие трансверсальности (12.24).

12.1.5 Анализ условия $\nu\hat{P}_z + \mu = 0$.

Обратимся теперь к условию $\nu\hat{P}_z + \mu = 0$. Оно эквивалентно условию

$$\nu\hat{P}_z\bar{z} + \mu\bar{z} = 0 \quad \forall\bar{z} \in \mathbb{R}^r. \tag{12.25}$$

Поскольку $\hat{P}_z\bar{z} = (0, \bar{x}(\tau_1))$, где $\bar{x}(\cdot)$ — решение задачи Коши

$$\frac{d\bar{x}}{d\tau} = \bar{v}f^\theta + v^\theta f_x^\theta\bar{x}, \quad \bar{v} = \sum \bar{z}_k\chi_k, \quad \bar{x}(\tau_0) = 0, \tag{12.26}$$

то

$$\nu\hat{P}_z\bar{z} = \nu_1\bar{x}(\tau_1) \overset{(12.23)}{=} \psi^\theta(\tau_1)\bar{x}(\tau_1) = \psi^\theta\bar{x}\Big|_{\tau_0}^{\tau_1} = \int_{\tau_0}^{\tau_1} \frac{d}{d\tau}(\psi^\theta\bar{x})d\tau =$$

$$= \int_{\tau_0}^{\tau_1} (\frac{d\psi^\theta}{d\tau}\bar{x} + \psi^\theta\frac{d\bar{x}}{d\tau})d\tau \overset{(12.22),(12.26)}{=} \int_{\tau_0}^{\tau_1} (-\psi^\theta v^\theta f_x^\theta\bar{x} + \psi^\theta(\bar{v}f^\theta + v^\theta f_x\bar{x}))d\tau =$$

$$= \int_{\tau_0}^{\tau_1} \psi^\theta f^\theta\bar{v}d\tau \overset{(12.26)}{=} \int_{\tau_0}^{\tau_1} \psi^\theta f^\theta \sum \bar{z}_k\chi_k d\tau.$$

Итак, условие (12.25) равносильно условию

$$\int_{\tau_0}^{\tau_1} \psi^\theta f^\theta \sum \bar{z}_k\chi_k d\tau + \sum \mu_k\bar{z}_k = 0, \quad \forall\bar{z} \in \mathbb{R}^r.$$

Отсюда вытекает, что

$$\int_{\tau_0}^{\tau_1} \psi^\theta f^\theta\chi_k d\tau + \mu_k = 0, \qquad k = 1, \ldots, r.$$

— в силу произвольности $\bar{z}_k$, или

$$\int_{\sigma_k} \psi^\theta f^\theta d\tau = -\mu_k, \qquad k = 1, \ldots, r.$$

Что отсюда следует? Вспомним, что $\mu \geqslant 0$ и $\mu\hat{z} = 0$, т.е. $\mu_k \geqslant 0$, причем $\mu_k = 0$ при $\hat{z}_k > 0 \;\; \forall k$. Поскольку $\sum \hat{z}_k\chi_k = v^\theta$, то условие $\hat{z}_k > 0$ означает, что $\sigma_k \subset \mathcal{M}_+$. Итак, мы получили, что

$$(a) \quad \int_{\sigma_k} \psi^\theta f^\theta d\tau = 0, \;\; \text{если } \sigma_k \subset \mathcal{M}_+, \tag{12.27}$$

$$(b) \quad \int_{\sigma_k} \psi^\theta f^\theta d\tau \leqslant 0, \;\; \text{если } \sigma_k \subset \mathcal{M}_0, \tag{12.28}$$

$$k = 1, \ldots, r.$$

Итог сказанному подводит следующая

Теорема 3. *Для любого индекса θ существует набор $\lambda^\theta = (\alpha_0, \alpha, \beta, \psi^\theta)$ такой, что выполнены условия*

$$\begin{cases} \alpha_0 \geqslant 0, \alpha \geqslant 0, \alpha F(\hat{p}) = 0, \\ \alpha_0 + \sum\alpha_i + |\beta| = 1, \\ -\frac{\psi^\theta}{d\tau} = v^\theta\psi^\theta f_x^\theta, \\ \psi(\tau_0) = \hat{l}_0, \psi(\tau_1) = -\hat{l}_1, \\ \int_{\sigma_k} \psi^\theta f^\theta d\tau = 0, \;\; \textit{при} \;\; \sigma_k \subset \mathcal{M}_+, \\ \int_{\sigma_k} \psi^\theta f^\theta d\tau \leqslant 0, \;\; \textit{при} \;\; \sigma_k \subset \mathcal{M}_0, \\ \qquad k = 1, \ldots, r \end{cases} \tag{12.29}$$

Здесь $\psi^\theta(\tau) : [\tau_0, \tau_1] \to \mathbb{R}^n$ - липшицева функция.

Глава 13

Лекция 13.

13.0.6 Представление условий стационарности в задаче B^θ с помощью функции $\psi(t)$. Принцип максимума индекса θ.

Пусть θ — фиксированный индекс. Ему соответствует отрезок $[\tau_0, \tau_1]$, отрезки $\Delta^k \subset [\tau_0, \tau_1], \quad k = 1, \dots, N$, множества $\mathcal{M}_0 = \bigcup\limits_{k=1}^{N} \Delta^k, \quad \mathcal{M}_+ = [\tau_0, \tau_1] \backslash \mathcal{M}_0$.

Вместе отрезки Δ^k множества $\mathcal{M}_0$ и интервалы (полуинтервалы) множества $\mathcal{M}_+$ образуют набор $\sigma_1, \dots, \sigma_k$.

Функция $v^\theta(\tau)$ равна нулю на $\mathcal{M}_0$ и единице на $\mathcal{M}_+$. Мы пользуемся представлением

$$v^\theta(\tau) = \sum \hat{z}_k \chi_k(\tau),$$

где χ_k — характеристическая функция отрезка σ_k, $\hat{z}_k$ принимает значения 0 или 1. Если $\hat{z}_k = 0$, то $\sigma_k \subset \mathcal{M}_0$; если $\hat{z}_k = 1$, то $\sigma_k \subset \mathcal{M}_k$. Наконец, $t^\theta(\tau)$ таковы, что

$$\frac{dt^\theta}{d\tau} = v^\theta, \qquad t^\theta(\tau_0) = \hat{t}_0 \Rightarrow t^\theta(\tau_1) = \hat{t}_1.$$

Обозначим через $\tau^\theta(t) : [\hat{t}_0, \hat{t}_1] \to [\tau_0, \tau_1]$ правую обратную функцию к функции $t^\theta(\tau)$, непрерывную слева. Тогда

$$t^\theta(\tau^\theta(t)) = t \qquad \forall t \in \hat{\Delta}, \quad \text{где} \quad \hat{\Delta} = [\hat{t}_0, \hat{t}_1].$$

Пусть $\lambda^\theta = (\alpha_0, \alpha, \beta, \psi^\theta)$ — набор, удовлетворяющий условиям (29) теоремы 3 лекции 12.

Положим

$$\psi(t) = \psi^\theta(\tau^\theta(t)) \qquad \forall t \in \hat{\Delta}.$$

Согласно предложению 2 лекции 10 из уравнения

$$-\frac{d\psi^\theta}{d\tau} = v^\theta \psi^\theta f_x^\theta, \quad \text{где} \quad f_x^\theta = f_x(x^\theta, u^\theta)$$

вытекает, что

$$-\frac{d\psi(t)}{dt} = \psi(t) f_x(x^\theta(\tau^\theta(t)), u^\theta(\tau^\theta(t))).$$

Пусть $\mathcal{N}_+ = [\hat{t}_0, \hat{t}_1] \backslash \{t^1, \dots, t^N\}$. Поскольку $u^\theta(\tau) = \hat{u}(t^\theta(\tau))$ на $\mathcal{M}_+$ и $\tau^\theta(t)$ переводит $\mathcal{N}_+$ в $\mathcal{M}_+$, то на $\mathcal{N}_+$ имеем:

$$u^\theta(\tau^\theta(t)) = \hat{u}(t^\theta(\tau^\theta(t))) = \hat{u}(t).$$

Следовательно п.в. на $[\hat{t}_0, \hat{t}_1]$ имеем:

$$u^\theta(\tau^\theta(t)) = \hat{u}(t).$$

Далее,

$$x^\theta(\tau^\theta(t)) = \hat{x}(t^\theta(\tau^\theta(t))) = \hat{x}(t) \qquad \forall t \in [\hat{t}_0, \hat{t}_1].$$

Следовательно,

$$-\frac{d\psi}{dt} = \psi \hat{f}_x \quad \text{п.в. на} [\hat{t}_0, \hat{t}_1] \quad \text{где} \hat{f}_x = f_x(\hat{x}, \hat{u}).$$

При этом в силу того же предложения 2 лекции 10:

$$\psi(\hat{t}_0) = \psi^\theta(\tau_0), \quad \psi(\hat{t}_1) = \psi^\theta(\tau_1).$$

Следовательно,

$$\psi(\hat{t}_0) = \hat{l}_{x_0}, \quad -\psi(\hat{t}_1) = \hat{l}_{x_1},$$

где

$$\hat{l}_{x_0} = l_{x_0}(\hat{p}, \alpha_0, \alpha, \beta),$$

$$\hat{l}_{x_1} = l_{x_1}(\hat{p}, \alpha_0, \alpha, \beta),$$

$$\hat{p} = (\hat{x}_0, \hat{x}_1), \quad \hat{x}_0 = \hat{x}(\hat{t}_0), \quad \hat{x}_0 = \hat{x}(\hat{t}_0).$$

Проанализируем теперь условие (27) лекции 12:

$$\int\limits_{\sigma_k} \psi^\theta(\tau) f(x^\theta(\tau), u^\theta(\tau)) d\tau = 0, \tag{13.1}$$

если $\sigma_k \subset \mathcal{M}_+$. Пусть $\sigma_k \subset \mathcal{M}_+$. Тогда $t^\theta(\sigma_k) \subset \mathcal{N}_+$. В интеграле (13.1) произведем замену переменной $\tau = \tau^\theta(t)$. Поскольку

$$\frac{d\tau^\theta(t)}{dt} = 1 \text{ на } \mathcal{N}_+$$

$$\psi^\theta(\tau^\theta(t)) \stackrel{def}{=} \psi(t) \text{ на } \hat{\Delta},$$

$$x^\theta(\tau^\theta(t)) = x(t) \text{ на } \hat{\Delta},$$

$$u^\theta(\tau^\theta(t)) = u(t) \text{ на } \mathcal{N}_+,$$

и следовательно, те же равенства верны на $t^\theta(\sigma_k)$, то по формуле замены переменной из (13.1) получаем:

$$\int\limits_{t^\theta(\sigma_k)} \psi(t) f(\hat{x}(t), \hat{u}(t)) dt = 0.$$

Здесь $t^\theta(\sigma_k)$ один из интервалов (полуинтервалов), составляющих множество $\mathcal{N}_+$. Обозначим его черезΔ_+. Тогда

$$\int\limits_{\Delta_+} \psi \hat{f} dt = 0, \quad \text{где} \quad \hat{f} = f(\hat{x}, \hat{u}).$$

Это условие выполнено для любого $\Delta_+ \subset \mathcal{N}_+$.

Проанализируем теперь условие (28) лекции 12:

$$\int\limits_{\Delta_k} \psi^\theta(\tau) f(x^\theta(\tau), u^\theta(\tau)) d\tau \leqslant 0, \tag{13.2}$$

где $\Delta_k \subset \mathcal{M}_0$. Напомним, что

$$\Delta_k := [\tau_0^k, \tau_1^k], \quad \tau^\theta(t^k) = \tau_0^k.$$

Поскольку

$$t^\theta(\tau) = t^k \text{ на } \Delta^k,$$

то

$$x^\theta(\tau) \stackrel{def}{=} \hat{x}(t^\theta(\tau)) = \hat{x}(t^k) \text{ на } \Delta_k.$$

Далее, из условий

$$-\frac{d\psi^\theta}{d\tau} = v^\theta \psi^\theta f_x^\theta = 0 \quad \text{на } \Delta_k$$

вытекает, что на Δ_k

$$\psi^\theta(\tau) = \psi^\theta(\tau_0^k) = \psi^\theta(\tau^\theta(t^k)) \stackrel{def}{=} \psi(t^k).$$

Наконец, по определению функции u^θ имеем на Δ^k

$$u^\theta(\tau) = u^k.$$

Следовательно, из условия (13.2) получаем, что

$$\psi(t^k) f(\hat{x}(t^k), u^k) \leqslant 0, \quad k = 1, \dots, N.$$

Таким образом, из теоремы 3 лекции 12 вытекает:

Теорема 4 (Принцип максимума индекса θ). *Для любого индекса θ существует набор $\lambda = (\alpha_0, \alpha, \beta, \psi)$ такой, что*

$$\begin{array}{ll} (a) & \alpha_0 \geqslant 0, \quad \alpha \geqslant 0; \quad \alpha_0 + |\alpha| + |\beta| = 1; \\ (b) & \alpha F(\hat{p}) = 0; \\ (c) & -\dot{\psi} = \psi \hat{f}_x; \\ (d) & \psi(\hat{t}_0) = \hat{l}_{x_0}; \quad -\psi(\hat{t}_1) = \hat{l}_{x_1}; \\ (e) & \psi(t^k) f(\hat{x}(t^k), u^k) \leqslant 0, \quad k = 1, \dots, N; \\ (f) & \int\limits_{\Delta_+} \psi \hat{f} dt = 0, \end{array} \tag{13.3}$$

для любого интервала или полуинтервала Δ_+, из которых состоит $\mathcal{N}_+$.

Здесь

$$\hat{f} = f(\hat{x}, \hat{u}); \qquad \hat{f}_x = f_x(\hat{x}, \hat{u});$$

$$\hat{l}_{x_0} = l_{x_0}(\hat{p}, \alpha_0, \alpha, \beta); \quad \hat{l}_{x_1} = l_{x_1}(\hat{p}, \alpha_0, \alpha, \beta);$$

$$\hat{p} = (\hat{x}_0, \hat{x}_1) = (\hat{x}(\hat{t}_0), \hat{x}(\hat{t}_1)).$$

Множество наборов λ, удовлетворяющих условиям (13.3), обозначим M^θ.

Предложение. *M^θ — конечномерный компакт, причем, проектор*

$$\lambda = (\alpha_0, \alpha, \beta, \psi) \mapsto (\alpha_0, \alpha, \beta) \tag{13.4}$$

инъективен на M^θ.

Доказательство. Действительно, набор $(\alpha_0, \alpha, \beta)$ однозначно определяет вектор $\hat{l}_{x_0} = l_{x_0}(\hat{p}, \alpha_0, \alpha, \beta)$, а значит и ψ - с помощью условий

$$-\dot{\psi} = \psi \hat{f}_x, \qquad \psi(\hat{t}_0) = \hat{l}_{x_0}.$$

Следовательно, проектор (13.4) инъективен на M^θ.

Ограниченность M^θ вытекает из условия нормировки $\alpha_0 + |\alpha| + |\beta| = 1$. Замкнутость M^θ очевидна. Следовательно, M^θ — конечный компакт. □

13.0.7 Лемма о равенстве нулю измеримой функции.

Лемма. *Пусть $h(t) : [\hat{t}_0, \hat{t}_1] \to \mathbb{R}$ — ограниченная измеримая функция. Пусть для любого отрезка $\Delta \subset [\hat{t}_0, \hat{t}_1]$ имет место равенство:*

$$\int\limits_\Delta h(t) dt = 0.$$

Тогда $h(t) = 0$ п.в. на $\hat{\Delta} := [\hat{t}_0, \hat{t}_1]$.

Доказательство. Предположим, что данное утверждение не верно. Тогда существует множество положительной меры на котором $h(t) \neq 0$. Следовательно, существует множество положительной меры на котором $h(t)$ сохраняет знак. Обозначим его A. Пусть для определенности $h(t) > 0$ на A. Поскольку A — измеримое множество, то $\forall\, \varepsilon > 0$ найдется покрытие множества A не более, чем счетной системой отрезков $\{\Delta_k\}$ ($\Delta_k \subset \hat{\Delta} \quad \forall\, k$), не имеющих общих внутренних точек, такое, что множество

$$m = (\bigcup \Delta_k) \backslash A$$

имеет меру не превосходящую ε. Поскольку

$$\bigcup \Delta_k = A \cup m, \quad A \cap m = \varnothing,$$

то

$$\int\limits_{\bigcup \Delta_k} h(t)dt = \int\limits_A h(t)dt + \int\limits_m h(t)dt. \tag{13.5}$$

В силу σ-аддитивности интеграла и условия леммы левая часть этого равенства равна нулю. Интеграл по m в правой части удовлетворяет оценке

$$\left|\int\limits_m h(t)dt\right| \leqslant \|h\|_\infty \mathrm{mes}(m) \leqslant \|h\|_\infty \cdot \varepsilon$$

Поскольку ε — произвольно мало, то из (13.5) получаем:

$$\int\limits_A h(t)dt = 0.$$

Последнее противоречит условиям

$$h(t) > 0 \quad \text{на} \quad A, \quad \mathrm{mes}(A) > 0.$$

Лемма доказана. □

13.0.8 Организация принципов максимумов индексов θ. Завершение доказательства принципа максимума задачи B.

Ниже нам будет удобно понимать индекс θ как множество пар (t^k, u^k), $\quad k = 1, \ldots, s$, где t^k упорядочены, $t^k \in [\hat{t}_0, \hat{t}_1]$, $\quad u^k \in U \quad \forall k$.

Введем на множестве индексов θ отношение частичного порядка: $\theta_1 < \theta_2$ (θ_1 предшествует θ_2 или θ_2 следует за θ_1), если для любого значения t^k индекса θ_1 и соответствующего значения u^k этого индекса найдутся значения $t^{k'}$ и $u^{k'}$ индекса θ_2 такие, что $t^k = t^{k'}$, $\quad u^k = u^{k'}$ (т.е. любая пара (t^k, u^k) первого индекса принадлежит второму индексу).

Пусть имеется пара ирдексов θ^1 и θ^2, возможно не сравнимых. Возьмем все пары (t^k, u^k) первого индекса и все пары $(t^{k'}, u^{k'})$ второго индекса и объединим их. Из объединения пар составим новый индекс θ_3. Положим

$$\theta_3 = \theta_1 \vee \theta_2,$$

где операция $\vee$ соответствует объединению пар t^k, u^k. Ясно, что $\theta_1 < \theta_3$ и $\theta_2 < \theta_3$. Итак, за любыми двумя индексами следует третий.

Отметим еще одно важное свойство следования, вытекающее из определений:

$$\theta_1 < \theta_2 \Rightarrow M^{\theta_2} \subset M^{\theta_1}.$$

Положим

$$M = \bigcap_\theta M^\theta,$$

где пересечение берется по всем индексам θ.

Лемма. *Множество M непусто.*

Доказательство. Фиксируем индекс θ_0 и рассмотрим все индексы θ следующие за $\theta_0 : \theta_0 < \theta$. Тогда

$$M = \bigcap_{\theta_0 < \theta} M^\theta$$

(Действительно, включение

$$M \subset \bigcap_{\theta_0 < \theta} M^\theta$$

очевидно. Обратно, для любого индекса θ_1 найдется θ такой, что $\theta_0 < \theta$ и $\theta_1 < \theta$. Следовательно, $M^\theta \subset M^{\theta_1}$. Отсюда вытекает, что

$$\bigcap_{\theta_0 < \theta} M^\theta \subset M.$$

Значит, имеет место равенство.)

Множества M^θ такие, что $\theta_0 < \theta$, образуют центрированную систему замкнутых подмножеств компакта M^{θ_0}. Действительно, если имеется конечная подсистема

$$M^{\theta_1}, \ldots, M^{\theta_s}, \quad \text{где} \quad \theta_0 < \theta_1, \ldots, \theta_0 < \theta_s,$$

то полагая

$$\theta = \theta_1 \vee \theta_2 \vee \ldots \vee \theta_s,$$

получаем непустой компакт M^θ, содержащийся в каждом M^{θ_k} (ибо $\theta_k < \theta$), а значит, и в их пересечении. Поскольку центрированная система замкнутых подмножеств компакта имеет непустое пересечение, то множество M непусто. Лемма доказана. □

Итак, компакт M непуст. Возьмем $\lambda = (\alpha_0, \alpha, \beta, \psi) \in M$. Для него выполнены все условия (13.3) принципа максимума любого индекса θ. Для любых $t = t^k \in [\hat{t}_0, \hat{t}_1], \quad u = u^k \in U$ пара (t^k, u^k) входит в некоторый индекс θ.

Для этого индекса из (13.3),(e) получаем:

$$\psi(t) = f(\hat{x}(t), u) \leqslant 0.$$

Это условие выполнено для любого $t \in [\hat{t}_0, \hat{t}_1]$ и любого $u \in U$. Это и есть условие максимума. Далее, для любого интервала $\Delta \subset [\hat{t}_0, \hat{t}_1]$ найдется индекс θ такой, что $\Delta = \Delta_+$ — один из интервалов множества $\mathcal{N}_+$ данного индекса. Следовательно, в силу (13.3),(f)

$$\int_\Delta \psi \hat{f} dt = 0$$

для любого $\Delta \subset [\hat{t}_0, \hat{t}_1]$. Отсюда в силу леммы 1 получаем:

$$\psi \hat{f} = 0 \quad \text{п.в. на} \quad [\hat{t}_0, \hat{t}_1],$$

т.е.

$$\psi(t) f(\hat{x}(t), \hat{u}(t)) = 0 \quad \text{п.в. на} \quad [\hat{t}_0, \hat{t}_1].$$

Это условие равенства нулю гамильтониана. Остальные условия (неотрицательности, нетривиальности, дополняющей нежесткости, сопряженное уравнение и условие трансверсальности) совпадают для всех индексов. Принцип максимума задачи B полностью доказан. Таким образом, завершено доказательство теоремы 3 лекции 11.

13.0.9 Принцип максимума в задаче A.

Напомним постановку задачи A:

$$F_0(p) \to \min, \quad F(p) \leqslant 0, \quad K(p) = 0, \quad \dot{x} = f(t, x, u), \quad u \in U,$$

где

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

t_0, t_1 — нефиксированы, $F \in \mathbb{R}^k,\ K \in \mathbb{R}^s,\ x \in \mathbb{R}^n,\ u \in \mathbb{R}^m,\ f \in \mathbb{R}^n$.

Пусть траектория

$$\mathcal{T} = (x(t), u(t), \mid t \in [t_0, t_1])$$

допустима в задаче B. Сформулируем для нее условия принципа максимума. Положим

$$l = \alpha_0 F_0 + \alpha F + \beta K,$$

где $\alpha_0 \in \mathbb{R}, \quad \alpha \in (\mathbb{R}^k)^*, \quad \beta \in (\mathbb{R}^s)^*$,

$$H = \psi_x f + \psi_t,$$

где $\psi_x \in (\mathbb{R}^n)^*$, $\psi_t \in \mathbb{R}$. Тогда

$$l = l(p, \alpha_0, \alpha, \beta)$$

— *концевая функция Лагранжа*

$$H = H(t, x, u, \psi_t, \psi_x)$$

— *функция Понтрягина.*

Отметим, что здесь ψ_x и ψ_t - двойственные переменные, связанные с переменными x и t соответственно (ранее нижний индекс x всегда означал вычисление градиента по x, например, f_x; здесь индексы x и t имеют другой смысл; надеемся, что это не приведет к путанице).

Определение. Будем говорить, что для траектории $\mathcal{T}$ имеет место принцип максимума, если существуют число α_0, векторы

$$\alpha \in (\mathbb{R}^k)^*, \beta \in (\mathbb{R}^s)^*$$

и липшицевы функции

$$\psi_x(t) : [t_0, t_1] \to (\mathbb{R}^n)^*, \quad \psi_t(t) : [t_0, t_1] \to \mathbb{R}$$

такие, что выполнены условия:

(I) неотрицательности

$$\alpha_0 \geqslant 0, \quad \alpha \geqslant 0;$$

(II) нетривиальности

$$\alpha_0 + |\alpha| + |\beta| > 0;$$

(III) дополняющей нежесткости

$$\alpha_i F_i(p) = 0, \quad i = 1, \ldots, k, \quad \text{где} \quad p = (t_0, x(t_0), t_1, x(t_1))$$

$$\alpha_i \text{ и } F_i \quad \text{компоненты векторов } \alpha \text{ и } F \quad \text{соответственно;}$$

(IV) сопряженные уравнения

$$-\dot{\psi}_x(t) = H_x(t, x(t), u(t), \psi_t(t), \psi_x(t)),$$

$$-\dot{\psi}_t(t) = H_t(t, x(t), u(t), \psi_t(t), \psi_x(t));$$

(V) трансверсальности

$$\psi_t(t_0) = l_{t_0}, \quad -\psi_t(t_1) = l_{t_1},$$

$$\psi_x(t_0) = l_{x_0}, \quad -\psi_x(t_1) = l_{x_1},$$

где градиенты $l_{t_0}, l_{t_1}, l_{x_0}, l_{x_1}$ вычисляются в точке

$$(t_0, x(t_0), t_1, x(t_1), \alpha_0, \alpha, \beta);$$

(VI) максимума

$$H(t, x(t), u, \psi_t(t), \psi_x(t)) \leqslant 0$$

для любых $t \in [t_0, t_1]$, $u \in U$;

(VII) равенство нулю гамильтониана

$$H(t, x(t), u(t), \psi_t(t), \psi_x(t)) = 0 \quad \text{п.в. на} \quad [t_0, t_1].$$

Теорема. *Если $\mathcal{T}$ — решение задачи A, то для $\mathcal{T}$ имеет место принцип максимума.*

13.0.10 Доказательство принципа максимума задачи A.

Доказательство. Пусть

$$\mathcal{T} = (x(t), u(t) | t \in [t_0, t_1])$$

— решение задачи A. Положим

$$\tau_0 = t_0, \;\; \tau_1 = t_1, \;\; t(\tau) = \tau.$$

Тогда $\mathcal{T}' = (t(\tau), x(\tau), u(\tau) | \tau \in [\tau_0, \tau_1])$ — решение следующей задачи A':

$$F_0(p) \to \min, \quad F(p) \leqslant 0, \quad K(p) = 0$$

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

где $p = (t(\tau_0), x(\tau_0), t(\tau_1), x(\tau_1))$.

Это было объяснено в лекции 11. Задача A' того же типа, что и задача B. Выпишем условия принципа максимума для траекторий $\mathcal{T}'$ в задаче B':

существуют число α_0, векторы α и β, липшицевы функции

$$\psi_x(\tau) : [\tau_0, \tau_1] \to (\mathbb{R}^n)^*, \quad \psi_t(\tau) : [\tau_0, \tau_1] \to \mathbb{R}$$

такие, что

$$\begin{gathered}
\alpha_0 \geqslant 0, \; \alpha \geqslant 0, \;\; \alpha_0 + |\alpha| + |\beta| > 0, \;\; \alpha F(p) = 0 \\
-\frac{d\psi_x}{d\tau} = H_x, \quad -\frac{d\psi_t}{d\tau} = H_t, \\
\psi_x(\tau_0) = l_{x_0}, \quad -\psi_x(\tau_1) = l_{x_1}, \\
\psi_t(\tau_0) = l_{t_0}, \quad -\psi_t(\tau_1) = l_{t_1}, \\
H(t(\tau), x(\tau), u(\tau), \psi_t(\tau), \psi_x(\tau)) \leqslant 0, \quad \forall \tau \in [\tau_0, \tau_1] \quad \forall u \in U; \\
H(t(\tau), x(\tau), u(\tau), \psi_t(\tau), \psi_x(\tau)) = 0, \quad \text{п.в. на} \quad [\tau_0, \tau_1],
\end{gathered} \tag{13.6}$$

где $l = \alpha_0 F_0 + \alpha F + \beta K$, $\qquad H = \psi_x f + \psi_t$, градиенты функции l вычисляются в точке $(t(\tau_0), x(\tau_0), t(\tau_1), x(\tau_1))$, а градиенты функции H — в точке

$$(t(\tau), x(\tau), u(\tau); \psi_t(\tau), \psi_x(\tau)), \; \tau \in [\tau_0, \tau_1].$$

Рассмотрим набор

$$\alpha_0, \alpha, \beta, \psi_t(t), \psi_x(t).$$

Элементарно проверяется, что он удовлетворяет всем условиям принципа максимума для траектории $\mathcal{T}$ в задаче A. Теорема доказана. $\square$

Глава 14

Лекция 14

14.0.11 Экстремаль управляемой системы.

Рассмотрим управляемую систему
$$\dot{x} = f(t, x, u), \quad u \in U, \tag{14.1}$$
где $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $U \subset \mathbb{R}^m$ — произвольное множество, f и f_x непрерывны по всем переменным. Введем для нее *функцию Понтрягина*
$$H = H(t, x, u, \psi_x) = \psi_x f(t, x, u).$$
(Ранее мы пользовались несколько иным определением функции Понтрягина, полагая $H = \psi_x f + \psi_t$. Так ее определили Дубовицкий и Милютин. Мы видели, что эта функция возникает естественно в том процессе доказательства принципа максимума, которому мы следовали. Однако подобное определение функции H не стало общепринятым. Позднее сам А.А. Милютин вернулся к первоначальному определению H в виде $\psi_x f$, данному Л.С. Понтрягиным).

Определение. Четверку функций $(x(t), u(t), \psi_t(t), \psi_x(t) \mid t \in [t_0, t_1])$, определенных на произвольном отрезке $\Delta = [t_0, t_1]$, будем называть *экстремалью* управляемой системы (14.1), если

(a) $(x(t), u(t) \,|\, t \in [t_0, t_1])$ траектория управляемой системы (14.1) т.е. $x(t)$ — абсолютно непрерывна на Δ, $u(t)$ - ограничена и измерима на Δ и
$$\dot{x} = f(t, x(t), u(t)), \quad u(t) \in U \quad \text{п.в. на} \quad \Delta \tag{14.2}$$

(b) функции $\psi_t(t)$ и $\psi_x(t)$ абсолютно непрерывны на Δ и выполнены условия
$$-\dot{\psi}_t(t) = H_t(t, x(t), u(t), \psi_x(t)) \quad \text{п.в. на} \quad \Delta \tag{14.3}$$
$$-\dot{\psi}_x(t) = H_x(t, x(t), u(t), \psi_x(t)) \quad \text{п.в. на} \quad \Delta \tag{14.4}$$
$$H(t, x(t), u(t), \psi_x(t)) + \psi_t(t) \leqslant 0, \quad \forall u \in U, \ \forall t \in \Delta \tag{14.5}$$
$$H(t, x(t), u(t), \psi_x(t)) + \psi_t(t) = 0 \quad \text{п.в. на} \quad \Delta. \tag{14.6}$$

Итак, экстремаль определяется условиями (14.2)-(14.6). Нетрудно видеть, что (14.3)-(14.6) это условия принципа максимума (за исключением условий неотрицательности, нетривиальности, дополняющей нежесткости и трансверсальности).

Экстремали, определеные условиями (14.2)-(14.6), называют также *понтрягинскими экстремалями*, но мы будем говорить просто - экстремаль.

Наша цель будет состоять в том, чтобы показать, что условия экстремальности можно переписать эквивалентным образом так, что переменная ψ_t в новых условиях участвовать не будет, т.е. ψ_t можно из условий экстремальности исключить вместе с сопряженным уравнением (14.3). Мы сделаем это, предположив, что множество U *замкнуто.* Ниже это предположение считаем выполненным.

14.0.12 Гамильтониан.

Рассмотрим функцию
$$\mathcal{H}(t, x, \psi_x) = \max_{u \in U} H(t, x, u, \psi_x) \tag{14.7}$$

Обозначим через $\operatorname{dom}\mathcal{H}$ множество троек (t,x,ψ_x) таких, что максимум по $u\in U$ в правой части равенства (14.7) достигается. Мы покажем, что для экстремали выполнены условия

$$(t,x(t),\psi_x(t))\in \operatorname{dom}\,\mathcal{H},\quad \forall t\in\Delta, \tag{14.8}$$
$$\mathcal{H}(t,x(t),\psi_x(t))+\psi_t(t)=0,\quad \forall t\in\Delta. \tag{14.9}$$

Другими словами,
$$\max_{u\in U} H(t,x(t),u,\psi_x)+\psi_t(t)=0,\quad \forall t\in\Delta \tag{14.10}$$

причем, для каждого $t\in\Delta$ максимум в левой части формулы (14.10) достигается. С этой целью мы докажем две леммы.

Лемма 1. *Пусть $h(t,u)$ непрерывна по t,u , где $t\in\Delta,\ u\in U$. Пусть существует множество $\mathcal{E}\subset\Delta$ и компакт $K\subset U$ такие, что замыкание $\mathcal{E}$ совпадает с Δ, $\bar{\mathcal{E}}=\Delta$, и*

$$\sup_{u\in U} h(t,u)=\max_{u\in K} h(t,u)\quad \forall t\in\mathcal{E}. \tag{14.11}$$

Тогда
$$\sup_{u\in U} h(t,u)=\max_{u\in K} h(t,u)\quad \forall t\in\Delta. \tag{14.12}$$

Доказательство. Пусть $\tilde t\in\Delta,\ u\in U$. Поскольку $\bar{\mathcal{E}}=\Delta$, то найдется последовательность $t_n\in\mathcal{E}$ такая, что $t_n\to\tilde t$. Для последовательности t_n укажем последовательность $u_n\in K$ точек максимума функции $h(t_n,\cdot)$ по компакту K. Тогда в силу (14.11)
$$h(t_n,u)\leqslant h(t_n,u_n) \tag{14.13}$$

Не ограничивая общности, считаем, что $u_n\to\tilde u\in K$ (иначе перейдем к подпоследовательности). Переходя в (14.13) к пределу по n, получаем:
$$h(\tilde t,u)\leqslant h(\tilde t,\tilde u) \tag{14.14}$$

Итак, для любого $u\in U$ найдется $\tilde u\in K$ такой, что выполнено (14.14) (при этом $\tilde t\in\Delta$ -произвольно). Пусть последовательность $u_n\in U$ такова, что

$$h(\tilde t,u_n)\to\sup_{u\in U} h(\tilde t,u). \tag{14.15}$$

Тогда в силу (14.14) найдется последовательность $\tilde u_n\in K$ такая, что

$$h(\tilde t,u_n)\leqslant h(\tilde t,\tilde u_n) \tag{14.16}$$

Не ограничивая общности, считаем, что $\tilde u_n\to\tilde u\in K$ (иначе перейдем к подпоследовательности). Переходя к пределу в (14.16) и учитывая (14.15) получаем:

$$\sup_{u\in U} h(\tilde t,u)\leqslant h(\tilde t,\tilde u)\leqslant\max_{u\in K} h(t,u)$$

Неравенство
$$\sup_{u\in U} h(\tilde t,u)\geqslant\max_{u\in K} h(\tilde t,u)$$

вытекает из условия $K\subset U$. Значит
$$\sup_{u\in U} h(\tilde t,u)=\max_{u\in K} h(\tilde t,u).$$

Равенство имеет место для любого $\tilde t\in\Delta$. □

Лемма 2. *Пусть $K\subset\mathbb{R}^m$ — компакт, Δ — отрезок, и пусть $h(t,u)$ непрерывна на $\Delta\times K$. Тогда функция*

$$\varphi(t)=\max_{u\in K} h(t,u)$$

непрерывна на Δ.

Доказательство. Пусть $\tilde{t} \in \Delta$ и пусть $t_n \in \Delta$ - произвольная последовательность такая, что $t_n \to \tilde{t}$. Покажем, что найдется подпоследовательность t_{n_k}такая, что $\varphi(t_{n_k}) \to \varphi(\tilde{t})$. (Отсюда последует непрерывность φ в точке $\tilde{t}$.)

Действительно, для последовательности t_n найдется последовательность $u_n \in K$ такая, что

$$\varphi(t_n) = h(t_n, u_n).$$

Из последоваиельности u_n извлечем сходящуюся подпоследовательность. Пусть это - сама u_n. Тогда $u_n \to \tilde{u} \in K$. Пусть $u \in K$ — произвольный элемент. Тогда

$$h(t_n, u) \leqslant h(t_n, u_n).$$

Переходя в этом неравенстве к пределу по n, получаем

$$h(\tilde{t}, u) \leqslant h(\tilde{t}, \tilde{u}).$$

Поскольку это верно для любого $u \in K$, то

$$h(\tilde{t}, \tilde{u}) = \max_{u \in K} h(\tilde{t}, u) = \varphi(\tilde{t}).$$

Итак,

$$\varphi(\tilde{t}) = h(\tilde{t}, \tilde{u}) = \lim_n h(t_n, u_n) = \lim_n \varphi(t_n).$$

Мы показали, что $\varphi(t_n) \to \varphi(\tilde{t})$. (В действительности это имеет место для подпоследовательности). Непрерывность φ в произвольной точке $\tilde{t} \in \Delta$ доказана. □

Обратимся теперь к условиям (14.2)-(14.6), определяющим экстремаль. Поскольку $u(t)$ существенно ограничена, то найдется замкнутый шар $B \subset \mathbb{R}^m$ такой, что $u(t) \in B$ п.в. на Δ. Положим $K = U \cap B$. Тогда K — компакт, $K \subset U$ и $u(t) \in K$ п.в. на Δ. Пусть $\mathcal{E}$ множество полной меры в Δ такое, что $u(t) \in K$ на $\mathcal{E}$ и всюду на $\mathcal{E}$ выполнено условие (14.6) в определении экстремали. Отметим, что $\bar{\mathcal{E}} = \Delta$. Положим

$$h(t, u) = H(t, x(t), u, \psi_x(t)) \tag{14.17}$$

Тогда в силу (14.5) и (14.6) имеем

$$h(t, u) + \psi_t(t) \leqslant 0 \quad \forall t \in \Delta \quad \forall u \in U \tag{14.18}$$

$$h(t, u(t)) + \psi_t(t) = 0 \quad \forall t \in \mathcal{E}. \tag{14.19}$$

Кроме того

$$u(t) \in K, \quad \forall t \in \mathcal{E}. \tag{14.20}$$

Из (14.18)-(14.20) получаем:

$$\sup_{u \in U} h(t, u) = \max_{u \in K} h(t, u) = -\psi_t(t), \quad \forall t \in \mathcal{E} \tag{14.21}$$

Отсюда в силу леммы 1 вытекает, что

$$\sup_{u \in U} h(t, u) = \max_{u \in K} h(t, u), \quad \forall t \in \Delta \tag{14.22}$$

Поскольку $K \subset U$, то из (14.22) получаем:

$$\sup_{u \in U} h(t, u) = \max_{u \in U} h(t, u), \quad \forall t \in \Delta \tag{14.23}$$

т.е. максимум реализуется на всех отрезках Δ. Из (14.21) и (14.22) вытекает, что

$$\max_{u \in K} h(t, u) = -\psi_t(t), \quad \forall t \in \mathcal{E} \tag{14.24}$$

Функция ψ_t липшицева, а функция в левой части равенства (14.24) непрерывна - в силу леммы 2. Поскольку $\bar{\mathcal{E}} = \Delta$, то из (14.24) следует, что

$$\max_{u \in K} h(t, u) = -\psi_t(t), \quad \forall t \in \Delta. \tag{14.25}$$

Из (14.22), (14.23) и (14.25) вытекает, что

$$\max_{u\in U} h(t,u) = -\psi_t(t), \quad \forall t \in \Delta, \tag{14.26}$$

причем, максимум в левой части равенства реализуется $\forall t \in \Delta$. Вспоминая определение (14.17) функции $h(t,u)$, получаем отсюда

$$\max_{u\in U} H(t,x(t),u,\psi_x(t)) + \psi_t(t) = 0, \quad \forall t \in \Delta \tag{14.27}$$

т.е. имеет место (14.9):

$$\mathcal{H}(t,x(t),u,\psi_x(t)) + \psi_t(t) = 0, \quad \forall t \in \Delta$$

Тем самым доказано и условие (14.8):

$$(t,x(t),\psi_x(t)) \in \mathrm{dom}\mathcal{H}, \quad \forall t \in \Delta$$

Из (14.6) и (14.9) получаем:

$$H(t,x(t),u(t),\psi_x(t)) = \mathcal{H}(t,x(t),\psi_x(t)) \text{ п.в. на } \Delta. \tag{14.28}$$

Последнее условие можно представить в виде:

$$u(t) \in \arg\max_{u\in U} H(t,x(t),u(t),\psi_x(t)) \text{ п.в. на } \Delta \tag{14.29}$$

где в правой части (14.29) стоит множество тех $u \in U$, на которых реализуется максимум функции $H(t,x(t),u(t),\psi_x(t))$ по $u \in U$.

Итак, доказана

Теорема 1. *Пусть множество U замкнуто. Тогда для всякой экстремали*

$$\Big(x(t),u(t),\psi_x(t),\psi_t(t) \mid t \in \Delta\Big)$$

системы (14.1) выполнены условия:

(a) $(t,x(t),\psi_x(t)) \in dom\ \mathcal{H}, \quad \forall t \in \Delta$

(b) $H(t,x(t),u(t),\psi_x(t)) = \mathcal{H}(t,x(t),\psi_x(t))$ *п.в. на* Δ

(c) $\mathcal{H}(t,x(t),\psi_x(t)) + \psi_t(t) = 0, \quad \forall t \in \Delta$

(d) Существует компакт $K \subset U$ такой, что $\max_{u\in K} H(t,x(t),u,\psi_x(t)) = \mathcal{H}(t,x(t),\psi_x(t)), \quad \forall t \in \Delta.$

Итак согласно теореме 1 и определению экстремали, для всякой экстремали выполнены условия:

(I) $\dot{x}(t) = f(t,x(t),u(t))$ п.в. на Δ

(II) $u(t) \in U$ п.в. на Δ

(III) $-\dot{\psi}_x(t) = H_x(t,x(t),u(t),\psi_x(t))$ п.в. на Δ

(IV) $(t,x(t),\psi_x(t)) \in \mathrm{dom}\mathcal{H}, \quad \forall t \in \Delta$

(V) $H(t,x(t),u(t),\psi_x(t)) = \mathcal{H}(t,x(t),\psi_x(t))$ п.в. на Δ

Последнее условие, как уже отмечалось эквивалентно условию (14.29).

В условиях (I)-(V) компонента ψ_t отсутствует. Покажем, что эти условия эквивалентным образом определяют экстремаль. Точнее докажем следующую теорему.

Теорема 2. *Пусть тройка*

$$(x(t), u(t), \psi_x(t) | t \in \Delta) \tag{14.30}$$

такова, что $x(t), \psi_x(t)$ абсолютно непрерывны на Δ, $u(t)$ — ограничена и измерима на Δ и выполнены условия (I)-(V) ($\Delta = [t_0, t_1]$). Пусть U замкнуто. Тогда функция

$$\psi_t(t) \overset{def}{=} -\mathcal{H}(t, x(t), \psi_x(t)), \quad t \in \Delta$$

удовлетворяет условию Липшица и уравнению

$$-\dot{\psi}_t(t) = H_t(t, x(t), u(t), \psi_x(t)) \text{ п.в. на } \Delta,$$

а четверка

$$\Big(x(t), u(t), \psi_x(t), \psi_t(t) \mid t \in \Delta\Big)$$

является экстремалью системы (14.1).

Итак, под экстремалью системы (14.1) можно понимать тройку (14.30) удовлетворяющую условиям (I)-(V). При этом выполнено равенство

$$\frac{d}{dt}\mathcal{H}(t, x(t), \psi_x(t)) = H_t(t, x(t), u(t), \psi_x(t)) \text{ п.в. на } \Delta, \tag{14.31}$$

вытекающее из условий (I)-(V). Это утверждает теорема 2.

Для доказательства теоремы 2 понадобятся две леммы.

Лемма 3. *Пусть $K \subset \mathbb{R}^m$ — компакт, Δ — отрезок, $h(t,u) : \Delta \times K \mapsto \mathbb{R}$ непрерывна на $\Delta \times K$. Пусть существует константа $L > 0$ такая, что $|h(\tau_2, u) - h(\tau_1, u)| \leqslant L|\tau_2 - \tau_1|$ для любых $\tau_1, \tau_2 \in \Delta$, $u \in U$. Тогда функция*

$$\varphi(t) = \max_{u \in K} h(t, u)$$

удовлетворяет на Δ условию Липшица с константой L.

Доказательство. Пусть $\tau_1, \tau_2 \in \Delta$. Выберем $u_1 \in K$ и $u_2 \in K$ так, что

$$\varphi(\tau_k) = h(\tau_k, u_k), \quad k = 1, 2.$$

Тогда

$$\varphi(\tau_2) - \varphi(\tau_1) = h(\tau_2, u_2) - h(\tau_1, u_1).$$

Поскольку u_1 — точка максимума $h(\tau_1, u)$ по K, то

$$h(\tau_1, u_1) \geqslant h(\tau_1, u_2).$$

Следовательно,

$$\varphi(\tau_2) - \varphi(\tau_1) \leqslant h(\tau_2, u_2) - h(\tau_1, u_2) \leqslant L|\tau_2 - \tau_1|.$$

Аналогично,

$$\varphi(\tau_1) - \varphi(\tau_2) \leqslant L|\tau_1 - \tau_2|.$$

Следовательно,

$$|\varphi(\tau_2) - \varphi(\tau_1)| \leqslant L|\tau_2 - \tau_1|.$$

Лемма доказана. □

Лемма 4. *Пусть $K \subset \mathbb{R}^m$ — компакт, Δ — отрезок, $h(t,u) : \Delta \times K \mapsto \mathbb{R}$ непрерывная функция. Положим*

$$\varphi(t) = \max_{u \in K} h(t, u).$$

Пусть $\tilde{t} \in int\Delta$, $\tilde{u} \in K$ — точка максимума функции $h(\tilde{t}, u)$ по $u \in K$, т.е.

$$\varphi(\tilde{t}) = h(\tilde{t}, \tilde{u}).$$

Пусть существуют производные $\varphi'(\tilde{t})$ и ${h'}_t(\tilde{t}, \tilde{u})$. Тогда $\varphi'(\tilde{t}) = {h'}_t(\tilde{t}, \tilde{u})$.

Доказательство. Положим
$$\psi(t) = h(t, \tilde{u}).$$
Тогда функции $\psi(t)$ и $\varphi(t)$ удовлетворяют условиям:

(a) $\varphi(\tilde{t}) = \psi(\tilde{t})$ т.к. $\psi(\tilde{t}) = h(\tilde{t}, \tilde{u}) = \varphi(\tilde{t})$;

(b) $\varphi(t) \geqslant \psi(t)$ т.к. $\varphi(t) = \max\limits_{u \in K} h(t, u) \geqslant h(t, \tilde{u}) = \psi(t)$.

Из (a) и (b) с очевидностью следует, что
$$\varphi'(\tilde{t}) = \psi'(\tilde{t}).$$
Но $\psi'(t) = h'_t(t, \tilde{u})$. Следовательно, $\varphi'(\tilde{t}) = h'_t(\tilde{t}, \tilde{u})$.

Лемма доказана. $\square$

Глава 15

Лекция 15

15.1 Тема: Эквивалентные формулировки принципа максимума. Понтрягинский минимум.

15.1.1 Доказательство теоремы 2 лекции 14.

Доказательство. Пусть тройка $(x(t), u(t), \psi(t) \mid t \in \Delta)$, где $\Delta = [t_0, t_1]$ удовлетворяет условиям (I)-(V) лекции 14. Положим, как и в доказательстве теоремы 1 лекции 14

$$h(t, u) = H(t, x(t), u, \psi_x(t)).$$

В силу (V) существует множество $\mathcal{E} \subset \Delta$ полной меры такое, что

$$h(t, u(t)) = \max_{u \in U} h(t, u) \quad \forall t \in \mathcal{E}$$

Пусть K — компакт в U такой, что $u(t) \in K$ п.в. на Δ. Считаем, не ограничивая общности, что

$$u(t) \in K \quad \forall t \in \mathcal{E}.$$

Тогда

$$\max_{u \in K} h(t, u) = \max_{u \in U} h(t, u) \quad \forall t \in \mathcal{E}$$

Поскольку $\bar{\mathcal{E}} = \Delta$, то в силу леммы 1

$$\max_{u \in K} h(t, u) = \max_{u \in U} h(t, u) \quad \forall t \in \Delta. \tag{15.1}$$

Далее, для любого фиксированного u п.в. на Δ имеем

$$\begin{aligned} h_t(t, u) = H_t(t, x(t), u, \psi_x(t)) + H_x(t, x(t), u, \psi_x(t))\dot{x}(t) + \\ + H_{\psi_x}(t, x(t), u, \psi_x(t))\dot{\psi}_x(t). \end{aligned} \tag{15.2}$$

Отсюда вытекает, что

$$\sup_{u \in K} \|h_t(\cdot, u)\|_\infty < +\infty.$$

Обозначим эту величину через L. Тогда ясно, что

$$|h(\tau_2, u) - h(\tau_1, u)| \leqslant L|\tau_2 - \tau_1|.$$

для всех $\tau_1, \tau_2 \in \Delta, u \in K$. Согласно лемме 3 лекции 14

$$\varphi(t) = \max_{u \in K} h(t, u)$$

— липшицева функция на Δ. При этом из определения функции h и равенства (15.1) вытекает, что

$$\varphi(t) = \max_{u \in U} H(t, x(t), u, \psi_x(t)) \overset{def}{=} \mathcal{H}(t, x(t), \psi_x(t)) \quad \forall t \in \Delta. \tag{15.3}$$

Таким образом, $\mathcal{H}(t, x(t), \psi_x(t))$ — липшицева функция на Δ.

Пусть теперь $\mathcal{E}_1 \subset \Delta$ — множество полной меры, такое, что в каждой точке $t \in \mathcal{E}_1$ выполнены пять условий:

(a) существуют $\dot{\varphi}(t), \dot{x}(t), \dot{\psi}_x(t)$;

(b) $\varphi(t) = h(t, u(t))$ (см.условие (v) лекции 14);

(c) $\dot{x}(t) = f(t, x(t), u(t))$;

(d) $-\dot{\psi}_x(t) = H_x(t, x(t), u(t), \psi_x(t))$;

(e) $u(t) \in K$.

Ясно, что такое множество существует, поскольку каждое из условий выполнено п.в. на Δ

Пусть $t \in \mathcal{E}_1$. Тогда существует $h_t(t, u(t))$. Действительно, в силу (15.2) имеем

$$h_t(t, u(t)) = H_t(t, x(t), u(t), \psi_x(t)) + H_x(t, x(t), u(t), \psi_x(t))\dot{x}(t) +$$

$$+ H_{\psi_x}(t, x(t), u(t), \psi_x(t))\dot{\psi}_x(t).$$

В силу (c), (d) и условия $H_{\psi_x} = f$ сумма двух последних слагаемых равна нулю. Следовательно,

$$h_t(t, u(t)) = H_t(t, x(t), u(t), \psi_x(t)) \tag{15.4}$$

Согласно лемме 4 предыдущей лекции

$$\dot{\varphi}(t) = h_t(t, u(t)).$$

Отсюда и из (15.3), (15.4) получаем, что

$$\frac{d}{dt}\mathcal{H}(t, x(t), \psi_x(t)) = H_t(t, x(t), u(t), \psi_x(t)).$$

Это равенство имеет место всюду на $\mathcal{E}_1$, т.е. п.в. на Δ. Положим

$$-\psi_t(t) = \mathcal{H}(t, x(t), \psi_x(t)) \quad \forall t \in \Delta. \tag{15.5}$$

Тогда

$$-\dot{\psi}_t(t) = H_t(t, x(t), u(t), \psi_x(t)) \quad \text{п.в. на } \Delta.$$

Итак, для четверки

$$(x(t), u(t), \psi_t(t), \psi_x(t) | t \in \Delta)$$

условия (2)-(4) из определения экстремали, данного в лекции 14, выполнены. Условие (5) лекции 14:

$$H(t, x(t), u, \psi_x(t)) + \psi_t(t) \leqslant 0 \quad \forall u \in U \quad \forall t \in \Delta$$

вытекает из равенства (15.5). Наконец, из условия (V) лекции 14:

$$\mathcal{H}(t, x(t), \psi_x(t)) = H(t, x(t), u(t), \psi_x(t)) \quad \text{п.в. на } \Delta$$

и равенства (15.5) вытекает, что

$$H(t, x(t), u(t), \psi_x(t)) + \psi_t(t) = 0 \quad \text{п.в. на } \Delta$$

т.е. имеет место условие (6) лекции 14. Таким образом, все условия (2)-(6) лекции 14, определяющие экстремаль, для четверки

$$(x(t), u(t), \psi_t(t), \psi_x(t) \mid t \in \Delta)$$

оказались выполнены.

Теорема доказана. $\square$

Замечание. *Обратимся к условию максимума:*

$$H(t, x(t), u(t), \psi_x(t)) = \mathcal{H}(t, x(t), \psi_x(t)) \tag{15.6}$$

Мы можем изменить $u(t)$ на множестве меры нуль так, что равенство (15.6) будет выполняться всюду на Δ, а не почти всюду, равно как и условие $u(t) \in U$.

Действительно, пусть $\mathcal{E} \subset \Delta$ -множество полной меры, на котором имеют место условие (15.6) и условие $u(t) \in U$. Пусть $K \subset U$ — компакт такой, что

$$\max_{u \in K} H(t, x(t), u, \psi_x(t)) = \mathcal{H}(t, x(t), \psi_x(t))$$

На множестве $\Delta \backslash \mathcal{E}$, имеющем нулевую меру, переопределим $u(t)$ так, чтобы были выполнены условия

$$u(t) \in K$$

и

$$H(t, x(t), u(t), \psi_x(t)) = \max_{u \in K} H(t, x(t), u, \psi_x(t))$$

Новая функция $u(t)$ окажется ограниченной измеримой (почему?) и условие (15.6) будет выполнено всюду на Δ. Тогда и условие (6) лекции 14 выполнено всюду на Δ.

Итак, определяя экстремаль с помощью условий (2)-(6) лекции 14 мы можем считать, что условия (15.5) и (15.6) выполнены для всех $t \in \Delta$. Аналогично, определяя экстремаль с помощью условий (I)-(V) лекции 14, мы можем считать, что условие (V) выполнено всюду на Δ. Мы можем также считать, что $u(t) \in U$ всюду на Δ и $u(t)$ ограничена всюду на Δ.

15.1.2 Уточнение условий принципа максимума.

Вернемся к задаче A:

$$F_0(p) \to \min, \quad F(p) \leqslant 0, \quad K(p) = 0,$$

где $p = (t_0, x(t_0), t_1, x(t_1))$;

$$\dot{x} = f(t, x, u) \quad u \in U.$$

Пусть сначала множество U замкнуто.

Пусть пара

$$(x(t), u(t) | t \in \Delta), \quad \Delta = [t_0, t_1]$$

— решение задачи A. Тогда для нее найдется число α_0, векторы α, β и абсолютно непрерывные на Δ функции $\psi_t(t), \psi_x(t)$ такие, что выполнены все условия принципа максимума (см. лекцию 13, условия (I)-(VII)).

Теорема 1 лекции 14 позволяет уточнить эти условия в случае когда U замкнуто. Условия (I)-(V) неотрицательности, нетривиальности, дополняющей нежесткости, сопряженные уравнения на ψ_t и ψ_x, а также условия трансверсальности сохраняются в прежнем виде:

$$\begin{aligned} \alpha_0 \geqslant 0, \quad \alpha \geqslant 0, \quad \alpha_0 + |\alpha| + |\beta| > 0, \quad \alpha F(p) = 0, \\ -\dot{\psi}_t = H_t, \quad -\dot{\psi}_x = H_x, \quad \text{п.в. на } \Delta, \\ \psi_t(t_0) = l_{t_0}, \quad -\psi_t(t_1) = l_{t_1}, \\ \psi_x(t_0) = l_{x_0}, \quad -\psi_x(t_1) = l_{x_1}, \end{aligned} \tag{15.7}$$

где

$$l = \alpha_0 F_0 + \alpha F + \beta K,$$

$$H = \psi_x f(t, x, u) = H(t, x, u, \psi_x).$$

Условия (VI) и (VII) усиливаются согласно теореме 1 лекции 14, до следующих условий

$$\begin{aligned} (t, x(t), \psi_x(t)) \in \operatorname{dom}\mathcal{H} \quad \forall t \in \Delta, \\ \mathcal{H}(t, x(t), \psi_x(t)) + \psi_t = 0 \quad \forall t \in \ \Delta, \\ \mathcal{H}(t, x(t), \psi_x(t)) = H(t, x(t), u(t), \psi_x(t)) \quad \text{п.в. на } \Delta, \end{aligned} \tag{15.8}$$

причем $u(t)$ можно изменить на множестве меры нуль так, что последнее равенство вместе с условием ограниченности $u(t)$ и условием $u(t) \in U$ окажутся выполненными всюду на Δ.

Ясно, что условиями (15.7) и (15.8) на практике пользоваться удобней.

Теорема 2 лекции 14 позволяет в случае, когда U замкнуто, дать следующую эквивалентную формулировку принципа максимума в задаче A : существует число α_0, векторы α, β и абсолютно непрерывные функции $\psi_t(t)$, $\psi_x(t)$ такие, что выполнены все условия (15.7) и (15.8) за исключением сопряженного уравнения на ψ_t

$$-\dot{\psi}_t = H_t.$$

Согласно теореме 2 лекции 14 сопряженное уравнение на ψ_t является следствием остальных условий принципа максимума. Фактически это позволяет компоненту ψ_t из условий принципа максимума исключить воспользовавшись равенством

$$\psi_t(t) = -\mathcal{H}(t, x(t), \psi_x(t)).$$

Согласно той же теореме 2 лекции 14 функция $\mathcal{H}(t, x(t), \psi_x(t))$ абсолютно непрерывна. С ее помощью можнозаписать и условие трансверсальности на ψ_t:

$$-\mathcal{H}(t_0, x(t_0), \psi_x(t_0)) = l_{x_0},$$

$$\mathcal{H}(t_1, x(t_1), \psi_x(t_1)) = l_{x_1}.$$

Рассмотрим теперь случай когда U — произвольное множество. В этом случае условие максимума (V) в формулировке принципа максимума можно, очевидно, распространить с множества U на его замыкание $\bar{U}$:

$$H(t, x(t), u, \psi_x(t)) + \psi_t(t) \leqslant 0 \ \forall u \in \bar{U}, \ \forall t \in \Delta.$$

(мы пользуемся определением H в виде $\psi_x f$), после чего мы снова можем применить теоремы 1 и 2 лекции 14. В результате мы получим те же уточнения в формулировках принципа максимума с той лишь разницей, что в определении гамильтониана $\mathcal{H}$ множество U заменится на его замыкание $\bar{U}$:

$$\mathcal{H}(t, x, \psi_x) = \max_{u \in \bar{U}} H(t, x, u, \psi_x),$$

причем, $\mathrm{dom}\mathcal{H}$ есть множество троек (t, x, ψ_x) таких, что максимум в правой части этого равенства достигается.

Наконец присутствие ограничений

$$p \in P, \quad (t, x, u) \in Q$$

в задаче A, где P и Q — открытые множества, никак не влияет ни на справедливость принципа максимума, ни на его формулировку. Напомним лишь, что траектория $(\hat{x}(t), \hat{u}(t), \mid t \in \hat{\Delta})$, $\hat{\Delta} = [\hat{t}_0, \hat{t}_1]$, должна удовлетворять этим ограничениям, причем, допустимость ее по ограничению Q означает существование компакта $\mathcal{C} \subset Q$ такого, что $(t, \hat{x}(t), \hat{u}(t)) \in \mathcal{C}$ п.в. на $\hat{\Delta}$.

На этом мы завершаем анализ условий принципа максимума.

15.1.3 Задача на фиксированном отрезке времени.

Рассмотрим задачу

$$\begin{aligned} J = F_0(p) \to \min, \quad & F(p) \leqslant 0, \quad K(p) = 0, \\ & \dot{x} = f(t, x, u), \quad u \in U, \end{aligned} \tag{15.9}$$

где $p = (t_0, x(t_0), t_1, x(t_1))$, отрезок $\Delta = [t_0, t_1]$ — фиксирован. При сведении ее к задаче A в концевой блок добавляются условия

$$t_0 - \hat{t}_0 = 0, \quad t_1 - \hat{t}_1 = 0,$$

после чего концевая функция Лагранжа приобретает вид:

$$\bar{l} = \beta_0(t_0 - \hat{t}_0) + \beta_1(t_1 - \hat{t}_1) + l,$$

где $l = \alpha_0 F_0 + \alpha F + \beta K$.

Условия трансверсальности для ψ_t дают

$$\psi_t(t_0) = \bar{l}_{t_0} = \beta_0,$$

$$-\psi_t(t_1) = \bar{l}_{t_1} = \beta_1.$$

Покажем, что условие

$$\alpha_0 + |\alpha| + |\beta| > 0 \tag{15.10}$$

эквивалентно условию нетривиальности.

Действительно, пусть

$$\alpha_0 + |\alpha| + |\beta| = 0.$$

Тогда $\alpha_0 = 0, \quad \alpha = 0, \quad \beta = 0$ и, следовательно,

$$\bar{l}_{x_0} = l_{x_0} = 0.$$

Из условий

$$\psi_x(t_0) = \bar{l}_{x_0} = 0, \; -\dot{\psi}_x = \psi_x f_x$$

вытекает, что $\psi_x \equiv 0$, а тогда в силу условия

$$\psi_x f + \psi_t = 0 \quad \text{п.в.}$$

получаем $\psi_t = 0$ и следовательно $\beta_0 = \beta_1 = 0$.

Итак, в качестве условия нетривиальности можно выбрать условие (15.10). Это позволяет убрать условия трансверсальности для ψ_t вместе с компонентами β_0 и β_1 из условий принципа максимума. Уравнение на ψ_t, $-\dot{\psi}_t = H_t$, как мы знаем, вытекает из остальных условий принципа максимума, и его также можно исключить. Выпишем кратко систему условий принципа максимума в задаче (15.10), к которой мы приходим в результате проведенного анализа.

Пусть $(x(t), u(t) | t \in \Delta)$ — решение задачи. Тогда существуют число α_0, векторы α, β и абсолютно непрерывная на Δ функция $\psi(t)$ такие, что

$$\alpha_0 \geqslant 0, \quad \alpha \geqslant 0, \quad \alpha_0 + |\alpha| + |\beta| > 0, \quad \alpha F(p) = 0 \tag{15.11}$$

$$-\dot{\psi} = H_x, \quad \text{п.в. на } \Delta, \tag{15.12}$$

$$\psi(t_0) = l_{x_0}, \quad -\psi(t_1) = l_{x_1}, \tag{15.13}$$

$$\max_{u \in \bar{U}} H(t, x(t), u, \psi(t)) = H(t, x(t), u(t), \psi(t)) \quad \text{п.в. на } \Delta, \tag{15.14}$$

причем, максимум в левой части последнего равенства достигается при каждом $t \in \Delta$. Здесь

$$H(t, x, u, \psi) = \psi f(t, x, u)$$

(мы заменили ψ_x на ψ).

Если обозначить левую часть равенства (15.14) через $(-\psi_t(t))$, то, как мы знаем, из условий принципа максимума последует, что ψ_t — липшицева функция и для нее выполнено сопряженное уравнение

$$-\dot{\psi}_t(t) = H_t(t, x(t), u(t), \psi(t)) \quad \text{п.в. на } \Delta.$$

15.1.4 Автономный случай.

Рассмотрим управляемую систему

$$\dot{x} = f(x, u), \quad u \in U.$$

Здесь f не зависит явно от t. Такая система называется *автономной*. Для нее

$$H = \psi_x f(x, u) = H(x, u, \psi_x)$$

— не зависит от t. Следовательно, сопряженное уравнение на ψ_t имеет вид

$$-\dot{\psi}_t = H_t = 0,$$

откуда вытекает, что $\psi_t = const$ для любой экстремали. Следовательно, на экстремали

$$\max_{u \in \bar{U}} H(x(t), u, \psi_x(t)) = const = H(x(t), u(t), \psi_x(t)).$$

Первое равенство выполняется всюду, а второе — почти всюду на Δ.

15.1.5 Задача с интегральным функционалом на фиксированном отрезке времени.

Рассмотрим задачу

$$J=\int_{t_0}^{t_1}F(x,u)dt\to\min,$$
$$\dot x=f(x,u),\quad u\in U, \tag{15.15}$$
$$x(t_0)=a,\quad x(t_1)=b,$$

где отрезок $[t_0,t_1]$ фиксирован. Предполагается, что F,F_x,f,f_x непрерывны по x,u.

Перепишем задачу в виде:

$$J=y_1-y_0\to\min,\quad x_0-a=0,\quad x_1-b=0,$$
$$\dot x=f(x,u),\quad \dot y=F(x,u),\quad u\in U,$$

$[t_0,t_1]$ — фиксирован.

Здесь

$$l=\alpha_0(y_1-y_0)+\beta_0(x_0-a)+\beta_1(x_1-b),$$
$$H=\psi_x f(x,y)+\psi_y F.$$

Поскольку t_0,t_1 фиксированы, и управляемая система автономна, то условия принципа максимума имеют имеют следующий вид:

$$\alpha_0\geqslant 0,\quad \alpha_0+|\beta_0|+|\beta_1|>0, \tag{15.16}$$
$$-\dot\psi_y=0,\quad \psi_y(t_0)=-\alpha_0,\quad \psi_\alpha(t_1)=\alpha_0 \tag{15.17}$$
$$-\dot\psi_x=H_x,\quad \psi_x(t_0)=\beta_0,\quad -\psi_\alpha(t_1)=\beta_1 \tag{15.18}$$
$$\max_{u\in\bar U}H(x(t),u,\psi_x(t),\psi_y(t))=const=H(x(t),u(t),\psi_x(t),\psi_y(t)), \tag{15.19}$$

причем первое из равенств (15.19) выполняется всюду, а второе — почти всюду на Δ.

Из условий (15.17) следует, что

$$\psi_y\equiv-\alpha_0.$$

Тогда

$$H=\psi_x f(x,y)+\alpha_0F(x,u).$$

Предположим, что $\alpha_0=0,\beta_0=0$. Тогда из условий

$$\psi_x(t_0)=0,\ -\dot\psi_x=\psi_x f_x$$

вытекает, что $\psi_x\equiv 0$ и, следовательно $\psi_x(t_1)=-\beta_1=0$. Поэтому условие нетривиальности эквивалентно условию

$$\alpha_0+|\beta_0|>0$$

или условию

$$\alpha_0+|\psi_x(t_0)|>0.$$

Положим теперь $\psi_x=\psi$ и

$$H=\psi f(x,u)-\alpha_0F(x,u)=H(x,u,\psi,\alpha_0). \tag{15.20}$$

Тогда принцип максимума для допустимой пары

$$(x(t),u(t)|t\in\Delta)$$

в задаче (15.15) заключается в следующем: существует число α_0 и абсолютно непрерывная на Δ функция $\psi(t)$ такие, что

$$\alpha_0\geqslant 0,\alpha_0+|\psi(t_0)|>0, \tag{15.21}$$
$$-\dot\psi=H_x(x(t),u(t),\psi(t),\alpha_0)\text{п.в. }, \tag{15.22}$$
$$\max_{u\in\bar U}H(x(t),u,\psi(t),\alpha_0)=const=H(x(t),u(t),\psi(t),\alpha_0), \tag{15.23}$$

причем, первое из равенств (15.23) выполняется всюду на Δ, а второе — почти всюду.

15.1.6 Задача быстродействия.

Рассмотрим задачу:

$$T \to \min, \quad x(0) = a, \quad x(T) = b,$$
$$\dot{x} = f(x, u), \quad u \in U. \tag{15.24}$$

Перепишем ее в виде:

$$t_1 \to \min, \quad t_0 = 0, \quad x_0 - a = 0, \quad x_1 - b = 0,$$

$$\dot{x} = f(x, u), \quad u \in U.$$

Здесь

$$l = \alpha_0 t_1 + \beta_t t_0 + \beta_0 (x_0 - a) + \beta_1 (x_1 - b),$$

$$H = \psi_x f(x, y) = H(x, u, \psi_x).$$

Условия принципа максимума имеют вид:

$$\alpha_0 \geqslant 0, \quad \alpha_0 + |\beta_t| + |\beta_0| + |\beta_1| > 0, \tag{15.25}$$
$$-\dot{\psi}_x = H_x, \quad \psi_x(t_0) = \beta_0, \quad -\psi_x(t_1) = \beta_1, \tag{15.26}$$
$$-\dot{\psi}_t = 0, \quad \psi_t(t_0) = \beta_t, \quad -\psi_t(t_1) = \alpha_0, \tag{15.27}$$
$$\max_{u \in \bar{U}} H(x(t), u, \psi_x(t)) + \psi_t(t) = 0, \quad \forall t \in \Delta, \tag{15.28}$$
$$H(x(t), u(t), \psi_x(t)) + \psi_t(t) = 0, \quad \text{п.в. на } \Delta. \tag{15.29}$$

Из (15.27) следует, что

$$\psi_t \equiv -\alpha_0.$$

Предположим, что $\psi_x(t_0) = 0$. Тогда из уравнения $-\dot{\psi}_x = \psi_x f_x(x, u)$ следует, что $\psi_x \equiv 0$. Следовательно, $\beta_0 = \beta_1 = 0$ и $\psi_x f = 0$. Из последнего равенства и условия (15.28) следует, что $\psi_t = 0$.Тогда $\alpha_0 = 0$ и $\beta_t = 0$ в силу (15.27). Мы пришли к противоречию с (15.25). Таким образом,

$$|\psi_x(t_0)| > 0$$

— эквивалентное условие нормировки.

Итак, окончательная формулировка принципа максимума для траектории

$$(x(t), u(t) | t \in [0, T])$$

в задаче (15.24) такова:
существует абсолютно непрерывная на $\Delta = [0, T]$ функция $\psi(t)$ такая, что

$$|\psi(0)| > 0, \tag{15.30}$$
$$-\dot{\psi} = H_x(x(t), u(t), \psi(t)), \quad \text{п.в. на } \Delta, \tag{15.31}$$
$$0 \leqslant \max_{u \in \bar{U}} H(x(t), u, \psi_x(t)) = H(x(t), u(t), \psi_x(t)), \quad \text{п.в. на } \Delta. \tag{15.32}$$

Из этих условий вытекает, что

$$\max_{u \in \bar{U}} H(x(t), u, \psi_x(t)) \equiv const \quad \text{на} \;\; \Delta. \tag{15.33}$$

Поэтому неравенство в (15.32) достаточно проверять хотя бы в одной точке $t \in \Delta$.

Последние две задачи, (15.15) и (15.24), являются основными в книге Понтрягина, Болтянского, Гамкрелидзе и Мищенко «Математическая теория оптимальных процессов» , где был получен принцип максимума для этих задач. Первое полное доказательство принципа максимума было дано В.Г. Болтянским.

15.1.7 Понтрягинский минимум.

Мы получили принцип максимума как необходимое условие стационарности в каждой конечномерной задаче B^θ (и можно доказать, что он есть эквивалент стационарности в каждой конечномерной задаче B^θ). Это дает некоторое представление о принципе максимума как о необходимом условии первого порядка. Но возникает вопрос: для какого типа минимума (наиболее слабого) принцип максимума является необходимым условием? Мы доказали лишь, что для абсолютного. Это произошло потому, что между основной и конечномерной задачами была установлена довольно грубая связь: из абсолютного минимума в основной задаче вытекает абсолютный минимум в каждой конечномерной задаче, а значит, стационарность в каждой конечномерной задаче. Однако, затратив несколько больше усилий, мы могли бы увидеть, к нарушению какого типа минимума приводит отсутствие стационарности в одной из конечномерных задач. Тогда бы мы дали точный ответ на поставленный вопрос. Ниже, не приводя больше никаких доказательств, мы сформулируем этот ответ в задаче A.

Как и условие Вейерштрасса, принцип максимума вытекает, например, из сильного минимума. В литературе так и принято квалифицировать его как необходимое условие (первого порядка) для сильного минимума. Напомним, что сильный минимум связан лишь с малыми вариациями фазовой переменной, а при этом вариации управления могут быть любыми. Однако уже наш первый опыт получения условия Вейерштрасса с помощью игольчатой вариации, позволяет заключить, что интересущий нас тип минимума связан как раз с игольчатыми вариациями управления, или, возможно, с такими вариациями управления, которые на множестве малой меры принимают не малые значения. Минимум в данном классе вариаций А.А.Милютин предложил назвать *понтрягинским* в честь первооткрывателя принципа максимума Льва Семеновича Понтрягина.

Дадим теперь точное определение понтрягинского минимума в канонической задаче A. Будем говорить, что траектория

$$\hat{\mathcal{J}} = (\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t}_0, \hat{t}_1]),$$

удовлетворяющая всем ограничениям канонической задачи A, доставляет понтрягинский минимум в этой задаче, если не существует последовательности траекторий

$$\mathcal{J}^N = (x^N(t), u^N(t) \mid t \in [t_0^N, t_1^N]), \quad N = 1, 2, \dots,$$

удовлетворяющей всем ограничениям задачи и такой что при $N \to \infty$ выполнены условия:

(i) $t_0^N \to \hat{t}_0, \quad t_1^N \to \hat{t}_1;$

(ii) $\max |x^N(t) - \hat{x}(t)| \to 0,$
где максимум берется по отрезку $[\hat{t}_0, \hat{t}_1] \bigcap [t_0^N, t_1^N]$;

(iii) $\int |u^N(t) - \hat{u}(t)| \to 0,$
где интеграл берется по отрезку $[\hat{t}_0, \hat{t}_1] \bigcap [t_0^N, t_1^N]$;

(iv) существует компакт $\mathcal{C} \in Q$ такой что для каждого N
$(t, x^N(t), u^N(t)) \in \mathcal{C}$ п.в. на $[t_0^N, t_1^N]$;

(v) для каждого N
$F_0(t_0^N, x^N(t_0^N), t_1^N, x^N(t_1^N)) < F_0(\hat{t}_0, \hat{x}(\hat{t}_0), \hat{t}_1, \hat{x}(\hat{t}_1)).$

Задание. Сформулируйте понятие понтрягинского минимума

а) в случае, когда в задаче A отрезок $[t_0, t_1]$ фиксирован, множество U компактно, а множество Q есть все пространство $\mathbb{R}^1 \times \mathbb{R}^n \times \mathbb{R}^m$,

б) для простейшей задачи вариационного исчисления.

Справедлива теорема: принцип максимума есть необходимое условие понтрягинского минимума в канонической задаче A. Но как уже было сказано, ее доказательство потребовало бы от нас некоторых дополнительных усилий, и мы его опускаем.

В заключение отметим, что понтрягинский минимум, занимающий промежуточное положение между слабым и сильным минимумами, не является локальным минимумом в смысле какой-либо топологии. Можно показать, что сходимость последовательностей, соответствуюших понятию понтрягинского минимума, такова, что второе замыкание множества в смысле этой сходимости не совпадает с первым. (Однако в частном случае, когда U компактно, Q есть все пространство, а отрезок $[t_0, t_1]$ фиксирован, понтрягинский минимум определяется следующей «малостью» вариаций: $\|\delta x\|_C < \varepsilon, \qquad \|\delta u\|_{L_1} < \varepsilon.$).

Несмотря на столь странное обстоятельство, связанное с понятием понтрягинского минимума и, на первый взгляд, затрудняющее его исследование традиционными методами анализа, именно понтрягинский минимум обладает наиболее богатой теорией условий первого и второго порядков, во многих отношениях более полной и содержательной, чем, скажем, теория слабого минимума классического вариационного исчисления. Это показали исследования последних десятилетий.